Comparative analysis of stochastic models for simulating leveraged ETF price paths

Kartikay Goyle[1]

[1] Department of Applied Mathematics, University of Washington, Washington, United States.
goylekar@uw.edu

Abstract:
This paper compares stochastic models for simulating leveraged Exchange-Traded Funds (LETFs) price paths, focusing on their applications in risk management and option pricing. Using TQQQ (a 3x leveraged ETF tracking NASDAQ-100) as our case study, we evaluate Geometric Brownian Motion (GBM), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), Heston stochastic volatility, Stochastic Volatility with Jumps (SVJD), and propose a novel Multi-Scale Volatility with Jumps (MSVJ) model that captures both fast and slow volatility components. Furthermore, we develop a comprehensive evaluation framework that examines both price and volatility characteristics of the simulated paths against the actual TQQQ data. Our analysis spans different market conditions, including the COVID-19 crash and the 2022 market drawdown. While our proposed MSVJ model excels in capturing volatility dynamics and price range estimation, we find that each model exhibits unique strengths in different aspects of LETFs' behavior. The choice of most appropriate model depends on specific considerations for different applications, such as risk assessment, options pricing, or portfolio management.

Keywords: Leveraged ExchangeTraded Funds (LETFs), Volatility Modeling and Forecasting, Path Forecasting and Simulation, Stochastic Modeling
Classification: 91G50, 91G60, 91G70, 91G99

1 Introduction

Leveraged exchange-traded funds (LETFs) have gained significant popularity among investors due to their ability to amplify returns and provide exposure to specific market sectors. These financial instruments are designed to deliver multiples of the daily returns of their underlying indices, making them attractive for traders seeking to capitalize on short-term market movements [12].

Two of the most prominent LETFs are the ProShares UltraPro QQQ (TQQQ) and the ProShares UltraPro S&P500 (UPRO). TQQQ aims to provide three times the daily return of the NASDAQ-100 Index, while UPRO seeks to offer three times the daily return of the S&P 500 Index. In particular, TQQQ has attracted significant attention from both retail and institutional investors, with its assets under

management (AUM) reaching \$22.95 billion as of September 2024 [31]. The fund's focus on the technology sector and its potential for amplified returns have contributed to its growing popularity. Similarly, UPRO's AUM has experienced substantial growth over the last five years, increasing from \$1.3 billion in September 2019 to \$4 billion in September 2024 [30].

The academic literature on LETFs has evolved significantly over the past decade. Avellaneda and Zhang (2010) [5] established a fundamental framework demonstrating that LETF values depend on both the underlying index movement and the accumulated variance over time. Cheng and Madhavan (2009) [12] provided early insights into LETFs' behavior, showing how their daily rebalancing mechanism can affect market volatility and create path-dependent returns that may lead to value destruction for long-term investors. Building on these foundations, Leung and Sircar (2015) [20] developed methods for analyzing implied volatility of LETFs options by first modeling the underlying ETF using stochastic volatility models and then deriving LETFs dynamics through Itô's lemma for option pricing purposes, while Leung and Santoli (2016) [19] provided a comprehensive treatment of LETFs price dynamics and options valuation. These theoretical frameworks have been further developed by Ahn et al. (2015) [1], who proposed consistent pricing approaches for LETFs options.

While these theoretical and empirical frameworks provide valuable insights, LETFs trade as independent instruments with their own market dynamics, including dedicated options markets and specific trading patterns. This market reality motivates our approach of directly modeling LETFs price dynamics for practical applications in risk management and option pricing.

Using TQQQ as our case study, we evaluate various stochastic models including Geometric Brownian Motion (GBM), Heston stochastic volatility, Stochastic Volatility with Jumps (SVJD), and a proposed Multi-Scale Volatility with Jumps (MSVJ) model. We compare both price and volatility characteristics of the simulated paths with actual TQQQ data. For price dynamics, we evaluate path trajectories using Dynamic Time Warping distance, price range capture using multi-band metrics, and return distributions using Kolmogorov-Smirnov tests. For volatility dynamics, we analyze realized volatility regression, volatility persistence, jump capture, and volatility of volatility similarity through specific statistical measures. Our analysis includes periods of significant market events, such as the COVID-19 crash and the 2022 market drawdown, to assess each model's ability to generate paths that reflect both price and volatility behavior observed in the market. This comprehensive evaluation provides insights for applications in risk management, option pricing, and trading strategy development.

The remainder of this paper is organized as follows. Section 2 presents the theoretical framework for each model under consideration. Section 3 describes the data sources, preprocessing techniques, and TQQQ data analysis. Section 4 outlines the model calibration methodology and optimization techniques. Section 5 presents the

evaluation framework and performance metrics used to assess the models focusing on price path and volatility. Section 6 presents the results of our comparative analysis, including the models' performance in path pricing, volatility modeling, and capturing distributional properties of returns. In Subsection 6.4, we also assess the models' capability to capture and simulate both short-term market crashes, such as the COVID-19 crash, and long-term drawdowns, like the 2022 market decline driven by recession fears. Section 7 discusses the practical implications and applications of our findings in context of risk management, option pricing, and portfolio optimization. Finally, Section 8 summarizes the key findings and contributions and suggests directions for future research.

2   Theoretical Framework

Existing approaches to modeling LETFs have primarily focused on their theoretical relationship with underlying indices. Leung and Santoli (2016) [19] model the reference index using stochastic volatility to derive LETF option prices that are arbitrage-free across different leverage ratios. Other studies by Dobi and Avellaneda (2012) [14] and Tang and Xu (2013) [29] have analyzed tracking error and return deviations through the lens of the underlying index relationship.

We propose a different approach by directly modeling the LETF as an independent financial instrument, focusing on practical applications in risk management and trading. We evaluate the following stochastic models:

2.1   Geometric Brownian Motion (GBM)

Geometric Brownian Motion (GBM) models asset price dynamics with the assumption that the logarithm of the asset price follows a Brownian motion with drift, as shown by Black and Scholes (1973) [9] and Merton (1973) [23]. The dynamics of an asset price under GBM are given by the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1}$$

where S_t is the asset price at time t, μ is the drift term (expected return), σ is the volatility, and W_t is a standard Wiener process.

The GBM model is implemented using a discretized version of the stochastic differential equation:

$$S_{t+\Delta t} = S_t \exp\left((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}Z_t \right) \tag{2}$$

where Δt is the time step size and Z_t is a random variable drawn from a t-distribution with df degrees of freedom, scaled to have the same variance as the standard normal distribution. This modification allows for the simulation of returns with heavier tails, capturing extreme events observed in financial markets [25].

However, the GBM model with t-distributed random variables still assumes constant drift and volatility, which is unrealistic for many financial assets. It does not capture other important stylized facts of asset returns, such as volatility clustering and asymmetric volatility. More sophisticated models, such as stochastic volatility models (e.g., Heston model) [18] and Stochastic Volatility with Jump-Diffusion (SVJD) models (e.g., Bates model) [7], have been developed to address these limitations, allowing for both time-varying volatility and discontinuous price movements which we incorporate and are discussed in the sections below.

2.2 Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Bollerslev (1986) [10], is an extension of the Autoregressive Conditional Heteroskedasticity (ARCH) model proposed by Engle (1982) [15]. The GARCH model is widely used to capture the time-varying volatility of asset returns. In this implementation, we consider a GARCH(1,1) model with t-distributed innovations, defined as follows:

$$r_t = \mu \Delta t + \sqrt{h_t} \varepsilon_t \tag{3}$$

$$h_t = \omega + \alpha (r_{t-1} - \mu \Delta t)^2 + \beta h_{t-1} \tag{4}$$

where r_t is the return at time t, μ is the drift term, Δt is the time step size, h_t is the conditional variance at time t, ε_t is the innovation term following a t-distribution with ν degrees of freedom, and ω, α, and β are non-negative parameters of the GARCH model.

The GARCH(1,1) model captures key features of asset returns, such as volatility clustering and persistence. The innovations are drawn from a t-distribution, allowing for heavier tails compared to the standard normal distribution. The conditional variance is modeled as a linear function of the previous squared innovation and the previous conditional variance, with the initial volatility determined based on the long-term average variance.

The GARCH model can be viewed as an extension of the Geometric Brownian Motion (GBM) model, as it incorporates time-varying volatility. While the GBM model assumes constant volatility, the GARCH model allows for the volatility to evolve overtime based on past returns and conditional variances. This enables the GARCH model to better capture the dynamic nature of asset return volatility.

2.3 Heston Stochastic Volatility Model

The Heston Stochastic Volatility Model, introduced by Heston (1993) [18], is a widely used continuous-time model in financial mathematics for capturing the dy-

namics of asset prices and their volatilities. The Heston model is defined by the following system of stochastic differential equations:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_t^S \tag{5}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^v \tag{6}$$

$$dW_t^S dW_t^v = \rho dt \tag{7}$$

where S_t is the asset price at time t, v_t is the instantaneous variance (squared volatility) at time t, μ is the drift term representing the expected return of the asset, κ is the mean reversion speed determining how quickly the variance reverts to its long-term mean, θ is the long-term mean of the variance process representing the average level of volatility over the long run, σ is the volatility of the variance process controlling the speed at which the volatility changes, W_t^S and W_t^v are correlated Wiener processes, and ρ is the correlation coefficient between the asset price and the variance process, capturing the leverage effect.

The Heston model improves the Black-Scholes model and the GBM model by incorporating stochastic volatility, allowing for more realistic modelling of asset price dynamics. While the GBM model assumes that the asset price follows a log-normal distribution with constant volatility, the Heston model allows for time-varying volatility that follows a Cox-Ingersoll-Ross (CIR) process. This enables the Heston model to capture the volatility clustering and the leverage effect observed in financial markets. Compared to the GARCH model, which is a discrete-time model for capturing volatility clustering and time-varying volatility, the Heston model provides a continuous-time framework. The GARCH model models the conditional variance of asset returns as a function of past squared returns and past conditional variances, while the Heston model treats the instantaneous variance as a separate stochastic process following a CIR process. The continuous-time nature of the Heston model allows for more flexible modeling of volatility dynamics and enables the derivation of closed-form solutions for European option prices.

2.4   Stochastic Volatility with Jump-Diffusion (SVJD) Model

The Stochastic Volatility with Jump-Diffusion (SVJD) model, also known as the Bates model, introduced by Bates (1996) [7], aims to capture both the continuous and discontinuous components of asset price dynamics, as well as the time-varying volatility for financial modeling and option pricing. The SVJD model is defined by the following system of stochastic differential equations:

$$\frac{dS_t}{S_t} = (\mu - \lambda \bar{k})dt + \sqrt{v_t}dW_t^S + J_t dN_t \tag{8}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^v \tag{9}$$

$$dW_t^S dW_t^v = \rho dt \tag{10}$$

$$J_t \sim N(\mu_{jump}, \sigma_{jump}^2) \tag{11}$$

$$N_t \sim \text{Poisson}(\lambda t) \tag{12}$$

where S_t is the asset price at time t, v_t is the instantaneous variance (squared volatility) at time t, μ is the drift term, κ is the mean reversion speed of the variance process, θ is the long-term mean of the variance process, σ is the volatility of the variance process, W_t^S and W_t^v are correlated Wiener processes, ρ is the correlation coefficient between the Wiener processes (capturing the leverage effect), J_t is the jump size following a normal distribution with mean μ_{jump} and standard deviation σ_{jump}, N_t is a Poisson process with intensity λ representing the arrival of jumps, and $\bar{k} = \mathbb{E}[e^{J_t}] - 1$ is the compensator term for the jump process.

The SVJD model extends the Heston model by incorporating discontinuous jumps in the asset price process, allowing for the modelling of sudden and significant price movements that are not captured by the continuous component alone. The jump component is modeled using a compound Poisson process, where the jump sizes follo a normal distribution, and the arrival of jumps modeled by the intensity parameter λ.

2.5   Multi-Scale Volatility with Jumps (MSVJ) Model

The Multi-Scale Volatility with Jumps (MSVJ) model incorporates both fast and slow volatility components to better capture the complex volatility dynamics observed in financial markets, particularly in the context of LETFs. This model is inspired by the work of Fouque et al. (2000) [17] on multi-scale stochastic volatility models. The proposed model is defined as follows:

$$dS_t = \mu S_t dt + \sqrt{V_t}S_t dW_t^S + S_{t-}dJ_t \tag{13}$$

$$V_t = V_t^f + V_t^s \tag{14}$$

$$dV_t^f = \kappa_f(\theta_f - V_t^f)dt + \sigma_f\sqrt{V_t^f}dW_t^f \tag{15}$$

$$dV_t^s = \kappa_s(\theta_s - V_t^s)dt + \sigma_s\sqrt{V_t^s}dW_t^s \tag{16}$$

$$dW_t^S dW_t^f = \rho dt \tag{17}$$

where S_t is the asset price at time t, V_t^f and V_t^s are the fast and slow volatility components at time t, respectively, V_t is the combined volatility at time t, μ is the drift term, κ_f and κ_s are the mean reversion rates, θ_f and θ_s are the long-term means, σ_f and σ_s are the volatilities of the fast and slow volatility components,

W_t^S and W_t^f are correlated Wiener processes, ρ is the correlation coefficient, and J_t is a jump process with intensity λ and jump size distribution ν.

The MSVJ model extends the concepts introduced in the previous sections by incorporating multiple volatility components operating at different time scales. While the Geometric Brownian Motion (GBM) model assumes constant volatility and the Heston model introduces a single stochastic volatility component, the MSVJ model allows for a more flexible representation of volatility dynamics by considering both fast and slow volatility components. The fast volatility component, V_t^f, captures short-term fluctuations in volatility, while the slow volatility component, V_t^s, represents longer-term trends. The asset price dynamics in the MSVJ model are driven by the combined volatility, V_t, which is the sum of the fast (V_t^f) and slow (V_t^s) volatility components. This allows the model to capture both short-term and long-term volatility effects on the asset price.

Similar to the SVJD model, the MSVJ model also includes a jump component, which enhances its ability to represent the discontinuous and abrupt price movements in addition to the continuous price dynamics captured by the diffusion component.

3   Data and Methodology

3.1   Data Sources and Preprocessing

For this study, we used the historical daily price data of the ProShares UltraPro QQQ (TQQQ), a 3X-leveraged exchange-traded fund (ETF) that seeks to provide three times the daily investment results of the NASDAQ-100 Index over a 14-year period from September 1, 2010, to September 1, 2024. This timeframe captures various market regimes, including periods of low and high volatility, bull and bear markets, and significant economic events, providing a comprehensive dataset for model calibration and evaluation.

We split the dataset into two subsets: a calibration set and a testing set. The calibration set consists of data from September 1, 2010, to August 31, 2019, and is used to estimate the parameters of the various models under consideration. The testing set, spanning from September 1, 2019, to September 1, 2024, is used to assess the performance of the calibrated models in an out-of-sample setting. For the testing set, for each model, we generated 1,000 one-year simulations starting from September 1, 2019, and compared the simulated paths to the realized data. This process is repeated for each year from 2019 to 2024, providing a rigorous framework for evaluating the models' ability to capture the dynamics of the TQQQ ETF.

3.2   TQQQ Data Analysis

Figure 1 displays the auto-correlation function (ACF) and partial auto-correlation function (PACF) of TQQQ daily log returns, indicating minimal serial correlation

Table 1: Descriptive Statistics and Risk Analysis of TQQQ Daily Log Returns

Metric	Value
Mean	0.0014
Median	0.0035
Standard Deviation	0.0386
Skewness	-0.8913
Kurtosis	8.4526
Leverage Effect	-0.0204
ADF Test p-value	0.0000
KPSS Test p-value	0.1000
GPD Shape Parameter	0.1752
GPD Scale Parameter	0.0000
Value at Risk (95%)	-0.0635
Expected Shortfall (95%)	-0.0969
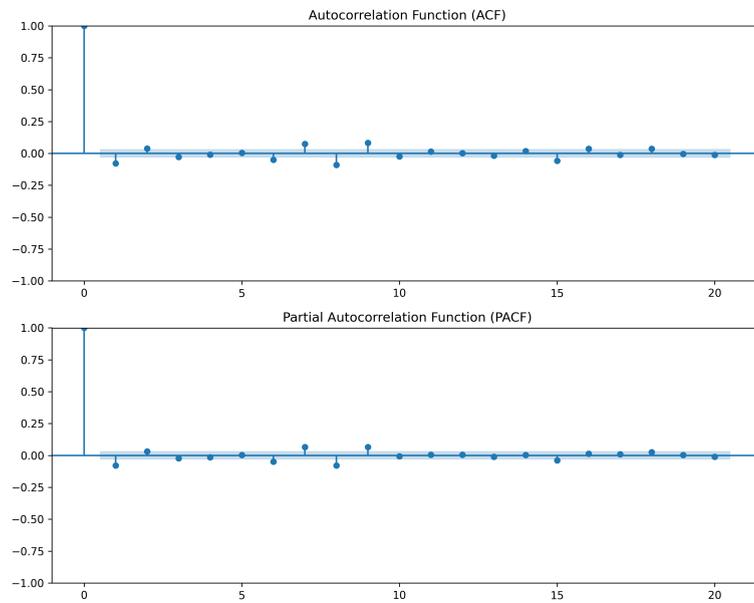
Figure 1: Auto-correlation and Partial Auto-correlation Functions of TQQQ Daily Log Returns

in the return series. Table 1 provides some of the descriptive statistics associated with TQQQ which suggest a left-skewed distribution with heavy tails. The resulting Pearson correlation coefficient (-0.0204) suggests a weak negative relationship between returns and volatility. To assess the stationarity of the returns, we employed the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. The ADF test rejects the null hypothesis of a unit root, while the KPSS test fails to reject the null hypothesis of stationarity at the 5% significance level indicating TQQQ daily log returns are indeed stationary.

4   Model Calibration

In this study, we calibrated the stochastic models using TQQQ's historical equity data. For volatility estimation, we relied solely on realized volatility, different from the common practice of using both realized and implied volatility for option pricing. Our approach may not fully capture market expectations of future volatility, which is typically reflected in implied volatility. Despite this potential limitation, our calibration method provided valuable insights into the models' capacity to replicate TQQQ's historical dynamics and predict future volatility movements.

4.1   Calibration Methodology

Rolling Window Calibration

To account for the time-varying nature of model parameters, we employed a rolling window calibration approach. The models were re-calibrated daily using a 252-day (approximately one trading year) lookback window. This approach allowed the models to adapt to evolving market conditions while maintaining a sufficient historical context for parameter estimation.

Objective Function

An objective function was designed to provide a comprehensive measure of the model's fit to the observed TQQQ data. It incorporated multiple metrics to capture various aspects of the price dynamics, including the accuracy of return and volatility estimates, the similarity of return distributions, and the matching of higher-order moments. The objective function was constructed as a weighted sum of the following components:

  (i)   Mean Squared Error (MSE) of returns

 (ii)   Mean Squared Error (MSE) of volatility

(iii)   Skewness difference

(iv)   Kurtosis difference

(v) Dynamic Time Warping (DTW) on price path

(vi) Dynamic Time Warping (DTW) distance on rolling volatility

The combined objective function was expressed as a weighted sum of these components, with weights determined through sensitivity analysis.

Optimization Algorithm

We employed a custom optimization algorithm inspired by the Differential Evolution (DE) algorithm [28] for parameter optimization. The pseudo-code for the optimization procedure is presented in Algorithm 1. The algorithm combines random exploration and exploitation of the best-performing parameters to calibrate the stochastic models using a rolling-window approach.

Algorithm 1 Calibrating Stochastic Models

 1: **Main Script:**
 2: Load data, split into historical and last year data
 3: Define parameter bounds for optimization
 4: Start with some initial set of parameters
 5: **Rolling Window Optimization:**
 6: **for** each window in data **do**
 7:     Initialize starting stock price S_0 for the simulations
 8:     **for** each iteration **do**
 9:         **if** first iteration or random exploration **then**
10:             Generate random parameters within bounds
11:         **else**
12:             Select top parameters and perturb them
13:         **end if**
14:         Generate simulated price paths using the model with current params
15:         Evaluate objective function, update best parameters if better score
16:     **end for**
17:     Store best parameters and score for the window
18: **end for**
19: Extract optimized parameters and performance scores from results
20: Select top 25 percentile results based on performance scores
21: Calculate median of top performing parameter sets to get best parameters

---

5   Model Evaluation

To evaluate how well each stochastic model captures TQQQ's price and volatility characteristics, we employ a comprehensive set of statistical measures. For price dynamics, we assess path trajectories using Dynamic Time Warping distance, price range capture using multi-band metrics, and return distributions using Kolmogorov-Smirnov tests. For volatility dynamics, we analyze realized volatility regression,

volatility persistence, jump capture, and volatility of volatility similarity through specific statistical measures.

5.1   Price Path Metrics

Dynamic Time Warping (DTW) Price

The Dynamic Time Warping (DTW) distance [8, 24] measures the similarity between the simulated price paths and the actual price path by finding the optimal alignment between them. The Dynamic Time Warping (DTW) distance between two time series X and Y after being normalized to X_{norm} and Y_{norm} is calculated as follows:

$$DTW(X_{norm}, Y_{norm}) = \min_{\mathbf{w} \in \mathcal{W}} \sqrt{\sum_{k=1}^{K} d(w_k)} \tag{18}$$

where $\mathbf{w} = (w_1, \ldots, w_K)$ is a warping path in the set of all possible warping paths $\mathcal{W}$, and $d(w_k)$ is the Euclidean distance between the corresponding elements of X_{norm} and Y_{norm} on the warping path. The warping path $\mathbf{w}$ represents a mapping between the indices of X_{norm} and Y_{norm} that minimizes the total distance while satisfying certain constraints, such as monotonicity and continuity. Lower DTW values indicate better performance, suggesting a higher degree of similarity in shape and timing between the simulated price paths and the actual price path.

Weighted Multi-band Capture Rate (WMCR) Price

The Weighted Multi-band Capture Rate (WMCR) Price metric assesses the model's ability to generate simulated price paths that capture the actual price path within specified percentage bands. It is calculated as follows:

$$WMCR(X, \hat{X}) = \sum_{b \in B} w_b \cdot \frac{1}{T} \sum_{t=1}^{T} I\left(M_t \cdot (1 - \frac{p_b}{100}) \leq X_t \leq M_t \cdot (1 + \frac{p_b}{100})\right) \tag{19}$$

where X is the actual price path, $\hat{X}$ represents the ensemble of simulated price paths, B is a set of percentage bands, w_b are band-specific weights, $M_t = \frac{\min(\hat{X}_t) + \max(\hat{X}_t)}{2}$ is the midpoint price derived from the simulated paths at time t, p_b is the percentage deviation corresponding to band b, $I(\cdot)$ is the indicator function, and T is the total number of time steps. The WMCR Price metric assigns higher weights to tighter bands, emphasizing the importance of the simulated price paths capturing the actual price path more closely. Higher WMCR Price values indicate better performance, suggesting that the actual price path consistently falls within the specified percentage bands derived from the simulated price paths. We used percentage bands from 20% to 60%, with weights decreasing as the band widens.

The 20% band had the highest weight of 5, while the 60% band had the lowest weight of 1.

Path Momentum Consistency (PMC)

The Path Momentum Consistency (PMC) metric measures the alignment of momentum characteristics between the simulated price paths and the actual price path. It is calculated using rolling windows to capture the momentum dynamics over time. The PMC metric is defined as follows:

$$PMC = \frac{1}{N(T - w + 1)} \sum_{i=1}^{N} \sum_{t=w}^{T} \not\Vdash(sign(m_{t-w+1:t}^{sim_i}) = sign(m_{t-w+1:t}^{actual})) \qquad (20)$$

where N is the number of simulated price paths, T is the total number of time steps, w is the size of the rolling window, $m_{t-w+1:t}^{sim_i}$ represents the rolling momentum of the i-th simulated price path from time $t - w + 1$ to t, $m_{t-w+1:t}^{actual}$ represents the rolling momentum of the actual price path from time $t - w + 1$ to t, and $\not\Vdash(\cdot)$ is the indicator function. The rolling momentum is calculated as the average return over the specified window. Higher PMC values indicate better alignment of momentum characteristics between the simulated and actual price paths.

Tail-Weighted Anderson-Darling (TWAD)

The Tail-Weighted Anderson-Darling (TWAD) metric is a modification of the original Anderson-Darling test [4] that places more emphasis on the tails of the distribution. The TWAD metric is defined as:

$$TWAD = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} \psi(x) dF(x) \qquad (21)$$

where $F_n(x)$ is the empirical cumulative distribution function (CDF) of the simulated returns, $F(x)$ is the CDF of the actual returns, n is the sample size, and $\psi(x)$ is a weight function that emphasizes the tails of the distribution. A common choice for the weight function is $\psi(x) = (F(x)(1 - F(x)))^{-1}$, which assigns more weight to the tails compared to the center of the distribution. Lower TWAD values indicate better agreement between the simulated and actual return distributions, particularly in the tails.

Kolmogorov-Smirnov (KS) Test

The Kolmogorov-Smirnov (KS) test is a non-parametric test that measures the maximum absolute difference between the empirical cumulative distribution functions (CDFs) of two samples [22]. It is used to determine whether two samples come from the same underlying distribution. The KS test statistic is defined as:

$$D_{n,m} = \sup_x |F_n(x) - G_m(x)| \tag{22}$$

where $F_n(x)$ and $G_m(x)$ are the empirical CDFs of the two samples, n and m are the respective sample sizes, and sup denotes the supremum (maximum) over all x values.

Lower values of the KS test statistic indicate better agreement between the two distributions being compared. Unlike the TWAD test which focuses on capturing tail behavior, the KS test provides a general measure of the difference between two distributions.

5.2   Volatility Metrics

Weighted Multi-band Capture Rate (WMCR) Volatility

Similar to WMCR Price, this metric measures the capture rate of the real volatility within the volatility bands derived from the simulated volatility. Higher values suggest better performance.

Dynamic Time Warping (DTW) Volatility

Similar to DTW Price, this metric measures the similarity between simulated and actual volatility paths using the DTW distance. Lower values indicate better performance.

Realized Volatility Regression (RVR)

The Realized Volatility Regression (RVR) metric assesses the model's ability to capture the relationship between the realized volatility and the model-implied volatility. It is estimated as follows:

$$RV_t = \beta_0 + \beta_1 \hat{\sigma}_t + \epsilon_t \tag{23}$$

where RV_t is the realized volatility at time t, and $\hat{\sigma}_t$ is the model-implied volatility [3]. We evaluate models based on the R^2 (values closer to 1 indicate better explanatory power) and the deviation of β_1 from 1 (values closer to 1 suggest better performance).

Volatility of Volatility Similarity (VVS)

The Volatility of Volatility Similarity (VVS) metric measures the similarity between the volatility of volatility (VoV) of the simulated paths and the VoV of the realized volatility. It is calculated as follows:

$$VVS = 1 - \frac{|\overline{VoV_{\hat{\sigma}}} - VoV_{RV}|}{\overline{VoV_{\hat{\sigma}}} + VoV_{RV}} \tag{24}$$

where $\overline{VoV_{\hat{\sigma}}}$ is the average VoV across all simulated paths, and VoV_{RV} is the VoV of the realized volatility [13].

For each simulated path i, the VoV is calculated as:

$$VoV_{\hat{\sigma}}^{(i)} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T-1}(\Delta\hat{\sigma}_t^{(i)} - \overline{\Delta\hat{\sigma}^{(i)}})^2}, \quad \Delta\hat{\sigma}_t^{(i)} = |\hat{\sigma}_t^{(i)} - \hat{\sigma}_{t-1}^{(i)}| \qquad (25)$$

where $\hat{\sigma}_t^{(i)}$ is the simulated volatility of path i at time t, $\Delta\hat{\sigma}_t^{(i)}$ is the absolute change in simulated volatility from time $t-1$ to t, and $\overline{\Delta\hat{\sigma}^{(i)}}$ is the mean of the absolute changes in simulated volatility for path i.

The VVS metric ranges from 0 to 1, with values closer to 1 indicating a higher similarity between the VoV of the simulated paths and the VoV of the realized volatility, suggesting that the model captures the variability of volatility changes more accurately.

Volatility Persistence Ratio (VPR)

The Volatility Persistence Ratio (VPR) measures the similarity between the persistence of volatility in the simulated paths and the persistence of volatility in the actual data. It is calculated as the ratio of the autocovariance of the simulated volatility to the autocovariance of the actual volatility:

$$VPR = \frac{\sum_{i=1}^{T-1}(\hat{\sigma}i+1^{sim} - \bar{\hat{\sigma}}^{sim})(\hat{\sigma}i^{sim} - \bar{\hat{\sigma}}^{sim})}{\sum i = 1^{T-1}(\sigma i+1^{actual} - \bar{\sigma}^{actual})(\sigma_i^{actual} - \bar{\sigma}^{actual})} \qquad (26)$$

where $\hat{\sigma}_i^{sim}$ is the simulated volatility at time i, $\bar{\hat{\sigma}}^{sim}$ is the average simulated volatility, σ_i^{actual} is the actual volatility at time i, $\bar{\sigma}^{actual}$ is the average actual volatility, and T is the total number of time steps [2]. The autocovariance measures the degree to which volatility at time i is related to volatility at time $i+1$. A higher autocovariance indicates stronger persistence in volatility, meaning that high (low) volatility periods tend to be followed by high (low) volatility periods.

The VPR compares the autocovariance of the simulated volatility to that of the actual volatility. A VPR close to 1 suggests that the model accurately captures the persistence of volatility observed in the actual data. Values greater than 1 indicate that the model overestimates volatility persistence, while values less than 1 indicate that the model underestimates volatility persistence.

Volatility Jump Capture (VJC)

The Volatility Jump Capture (VJC) metric measures the model's ability to capture sudden and significant changes (jumps) in volatility. It is calculated as the average proportion of actual volatility jumps that are successfully captured by the simulated volatility paths:

$$VJC = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \frac{\sum_{j=1}^{N_J} I(|\hat{\sigma}t_j^{sim_i} - \hat{\sigma}t_j - 1^{sim_i}| > \theta_i)}{\sum_{j=1}^{N_J} I(|\sigma_{t_j}^{actual} - \sigma_{t_j-1}^{actual}| > \theta_{actual})} \tag{27}$$

where N_{sim} is the number of simulated volatility paths, N_J is the number of actual volatility jumps, t_j are the times at which jumps occur, $\hat{\sigma}t_j^{sim_i}$ is the simulated volatility of path i at time t_j, σt_j^{actual} is the actual volatility at time t_j, θ_i is the jump threshold for the i-th simulated path, θ_{actual} is the jump threshold for the actual volatility, and $I(\cdot)$ is an indicator function that equals 1 if the condition inside the parentheses is true and 0 otherwise [6]. Higher values indicating better jump capture performance.

6 Results

In this section, we present the results of our comparative analysis of the Multi-Scale, GBM, SVJD, Heston, and GARCH models in capturing the dynamics of the TQQQ leveraged ETF. We evaluated the models' performance across three key aspects: path pricing, volatility modeling, and distributional properties of returns, using a comprehensive set of metrics and tests. Table 2 provides an overview of the models' performance across various metrics, highlighting their strengths and weaknesses in modelling the TQQQ leveraged ETF.

The performance metrics presented in Table 2 are averaged over the last five years of data, from September 2019 to September 2024. This period encompasses a wide range of market conditions, from low to high volatility, making it an ideal testing ground for evaluating the models' ability to capture the complex dynamics of the TQQQ leveraged ETF. In addition, we also conducted a more focused investigation of the models' performance using data from September 2023 to September 2024. During this one-year period, TQQQ experienced a significant surge of over 60%, reflecting a primarily bullish market sentiment. However, this period also witnessed substantial volatility, with TQQQ suffering a drawdown of 20% or more on four separate occasions and the VIX index, a measure of market volatility, reaching a value of 20 or higher four times, with it peaking at 65 during this timeframe.

6.1 Path Pricing Performance

To assess the models' ability to capture the overall price dynamics of TQQQ, we looked at the Dynamic Time Warping distance for price (DTW_{Price}) and the Weighted Multi-band Capture Rate for price ($WMCR_{Price}$).

The SVJD model achieved the lowest DTW_{Price} of 18.5719, closely followed by the GBM model (18.7055) and the Heston model (19.527), suggesting that these models best capture the overall price dynamics of TQQQ. The SVJD model's DTW_{Price} performance was nearly similar to that of GBM, GARCH performs the worst of all the models.

(a) GARCH Model

(b) GBM Model

(c) MSVJ model

(d) Heston model

(e) SVJD model
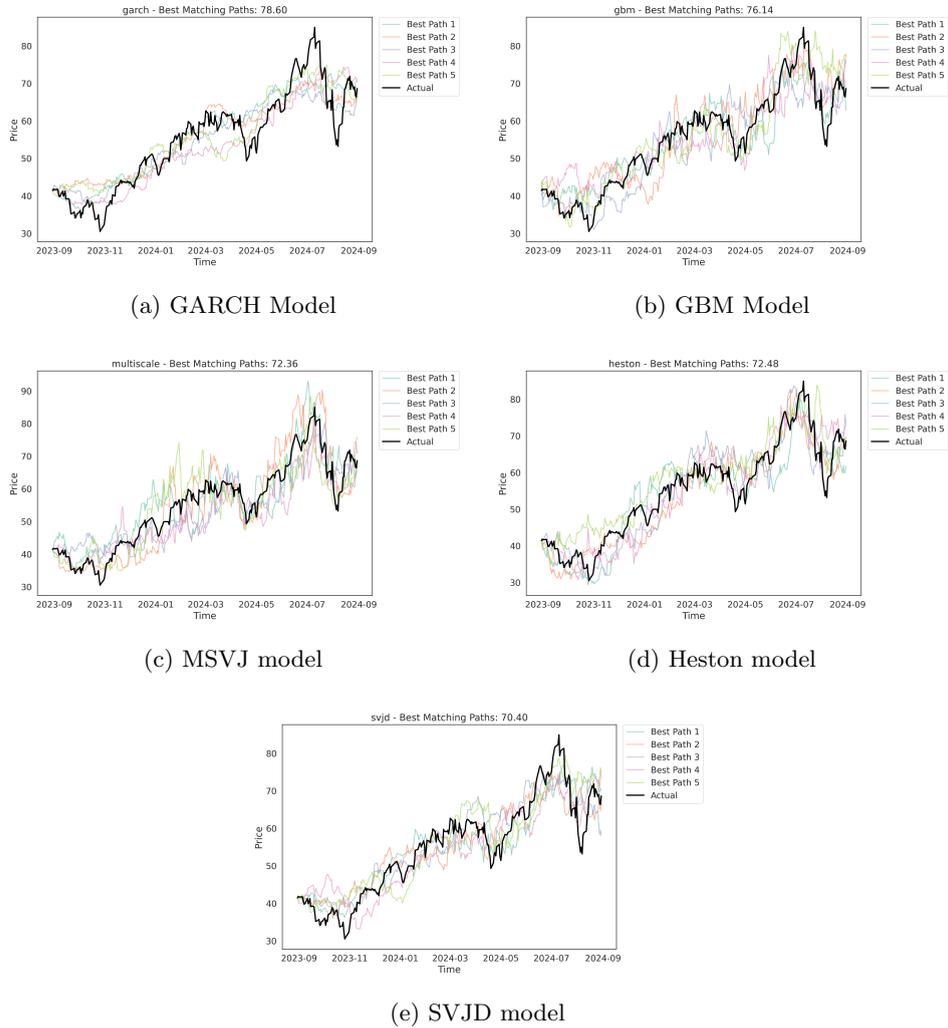
Figure 2: DTW Price distance for models from September 2023 to September 2024. A lower DTW distance indicates better performance in capturing the price path dynamics during this period.

(a) GARCH Model

(b) GBM Model

(c) MSVJ model

(d) Heston model

(e) SVJD model
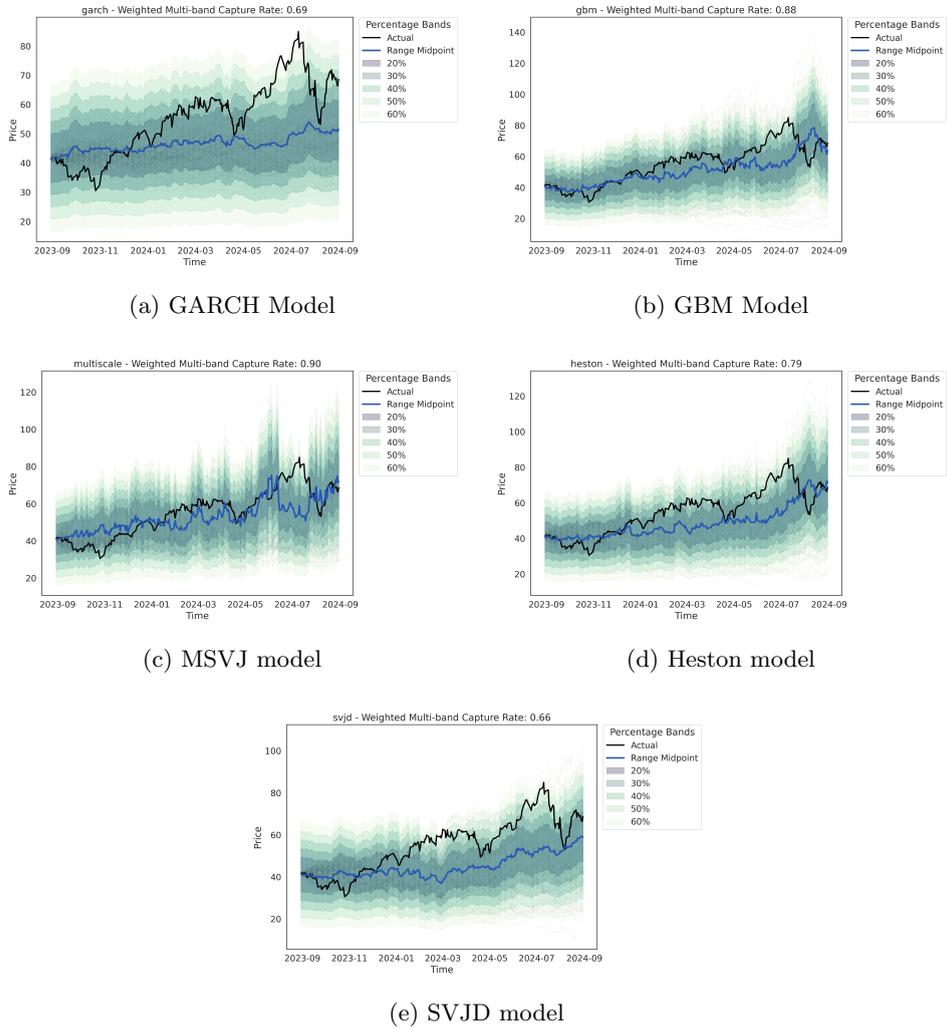
Figure 3: WMCR Price for all models from September 2023 to September 2024. A higher value indicates better performance in capturing the price paths dynamics within the specified range.

Metric	GARCH	Heston	GBM	SVJD	MSVJ
$\text{DTW}_{\text{Price}}$	20.0342	19.527	18.7055	**18.5719**	19.8595
$\text{DTW}_{\text{Volatility}}$	5.5814	2.9608	2.9401	**2.9353**	3.1328
$\text{WMCR}_{\text{Price}}$	0.7420	0.7510	0.7665	0.6916	**0.7699**
$\text{WMCR}_{\text{Volatility}}$	0.4143	0.4427	0.4313	0.4478	**0.4741**
RVR_{R2}	0.0729	0.0795	0.0713	0.1003	**0.1239**
RVR_{Beta}	-0.0326	0.0073	0.0059	**0.0388**	0.0016
VPR	-0.0003	0.0009	-0.0022	**0.0416**	0.0034
VJC	0.0691	0.0910	0.0908	**0.1079**	0.0951
VVS	0.6893	0.7357	0.7357	0.7396	**0.7403**
PMC	0.5045	0.5119	0.5132	0.5103	**0.5246**
TWAD	101.4809	8.8335	8.3610	17.0285	**6.8949**
KS_Statistic	0.2292	0.1273	0.1247	0.1259	**0.1196**
Median	0.0002	-0.0001	0.0007	0.0006	**0.0013**
Mean	0.0003	0.0005	**0.0014**	0.0001	0.0010
Standard Deviation	0.0146	0.0363	**0.0384**	0.0273	0.0352
$\text{VAR}_{5\%}$	-0.0222	-0.0569	**-0.0593**	-0.0449	-0.0567
$\text{ES}_{5\%}$	-0.0314	-0.0746	**-0.0777**	-0.0576	-0.0768

Table 2: Model performance comparison across various metrics, with the best values for each metric highlighted in bold

The MSVJ model also achieved the highest $\text{WMCR}_{\text{Price}}$ of 0.7699 followed by second best GBM model (0.7665), outperforming the other models, indicating its effectiveness in capturing price dynamics across multiple time scales.

To further evaluate the models' ability to capture the momentum characteristics of the price paths, we looked at the Path Momentum Consistency (PMC) metric. The MSVJ model achieved the highest PMC of 0.5246, suggesting better alignment of momentum characteristics compared to the other models.

Figure 2 illustrates the DTW distance for the simulated price paths of each model compared to the actual TQQQ price path from September 2023 to September 2024. The SVJD and Heston demonstrated the closest alignment with the actual price path, consistent with their strong average $\text{DTW}_{\text{Price}}$ metric performance in Table 2.

Figure 3 presents the WMCR Price distance for the models over the same period, with the MSVJ model exhibiting the highest capture rate followed by GBM model, consistent with its strong average $\text{WMCR}_{\text{Price}}$ metric performance in Table 2.

These results suggest that the SVJD and GBM models excel at capturing the overall price dynamics of TQQQ, as evidenced by their low DTWPrice values. The

MSVJ, on the other hand, was most effective in capturing the overall price range bands and aligning momentum characteristics, as indicated by its high WMCRPrice and PMC values. In the next section, we consider other aspects of the models' performance, such as volatility modeling and distributional properties, to gain a more comprehensive understanding of their suitability for modeling LETFs.

6.2   Volatility Modeling Performance

Accurate modelling of volatility dynamics is crucial for capturing the complex behaviour of LETFs like TQQQ. We assessed the models' volatility modeling performance using various metrics, including the DTW distance for volatility (DTWVolatility), Weighted Multi-band Capture Rate for volatility (WMCRVolatility), Realized Volatility Regression (RVR) metrics, Volatility of Volatility Similarity (VVS), Volatility Persistence Ratio (VPR), and Volatility Jump Capture (VJC).

The SVJD model achieved the lowest DTWVolatility of 2.9353, suggesting that it best captured the overall volatility dynamics of TQQQ. The GBM model followed with a DTWVolatility of 2.9401, and the Heston model with 2.9608. The MSVJ model achieved the highest $\text{WMCR}_{\text{Volatility}}$ of 0.4741, followed by the SVJD model with 0.4478, indicating their effectiveness in capturing volatility dynamics across different percentage bands.

The MSVJ model exhibited the highest RVRR2 value of 0.1239, indicating its superior explanatory power in predicting realized volatility. This suggests that the MSVJ model captured a larger proportion of the variation in the realized volatility compared to the other models. In contrast, the SVJD model had the RVRBeta coefficient closest to 1 (0.0388), implying the least bias in predicting realized volatility.

In terms of capturing the variance of volatility, the MSVJ model also achieved the highest VVS of 0.7403. The SVJD model, on the other hand, performed better in capturing the VPR and VJC, with values of 0.0416 and 0.1079, respectively, suggesting its effectiveness in modelling volatility persistence and jumps.

Figure 4 illustrates the DTW distance for the simulated volatility paths of each model compared to the actual TQQQ volatility path from September 2023 to September 2024. The Heston model and the SVJD model demonstrated the closest alignment with the actual volatility path, consistent with their strong average $\text{DTW}_{\text{Volatility}}$ metric performance in Table 2. Subsequently, Figure 5 presents the WMCR Volatility metric for the models over the same period, with the MSVJ model and SVJD model exhibiting the highest capture rate across multiple percentage bands, also consistent with their strong average $\text{WMCR}_{\text{Volatility}}$ metric performance in Table 2.

These results highlight the strengths of different models in capturing various aspects of volatility dynamics. The SVJD model demonstrated superior performance in capturing the overall volatility dynamics, as evidenced by its low DTWVolatility and high WMCRVolatility. The MSVJ model also excelled in capturing volatility dynamics across different percentage bands, as indicated by its high WMCRVolatility.

(a) GARCH model

(b) GBM model

(c) MSVJ model

(d) Heston model

(e) SVJD model
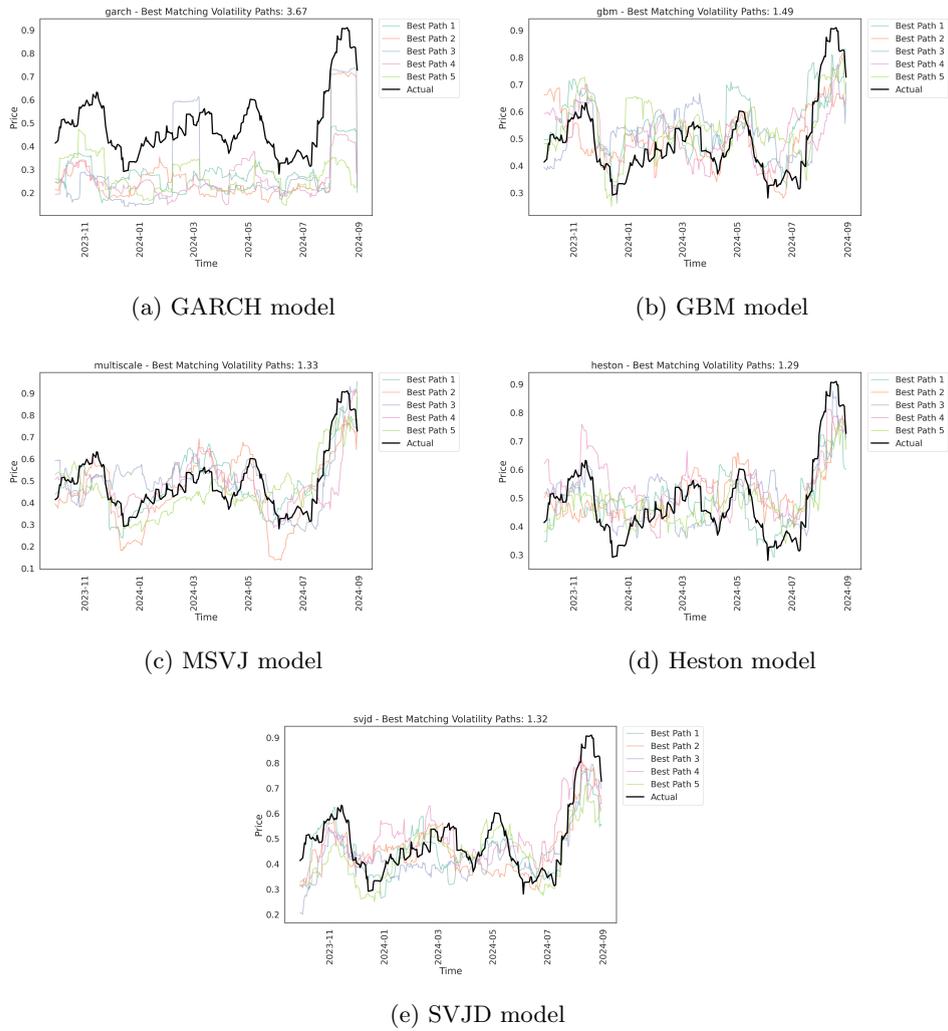
Figure 4: DTW rolling volatility distance for models from September 2023 to September 2024. A lower DTW distance indicates better performance in capturing the volatility path dynamics during this period.

Additionally, the MSVJ model showed the best explanatory power in predicting realized volatility and capturing the variability in volatility, as demonstrated by its high RVRR2 and VVS values. The SVJD model also showed effectiveness in modeling volatility persistence and jumps, as demonstrated by its strong VPR and VJC values. In the next section, we analyze the distributional properties of our simulated returns compared to the actual returns.

(a) GARCH model

(b) GBM model

(c) MSVJ model

(d) Heston model

(e) SVJD model
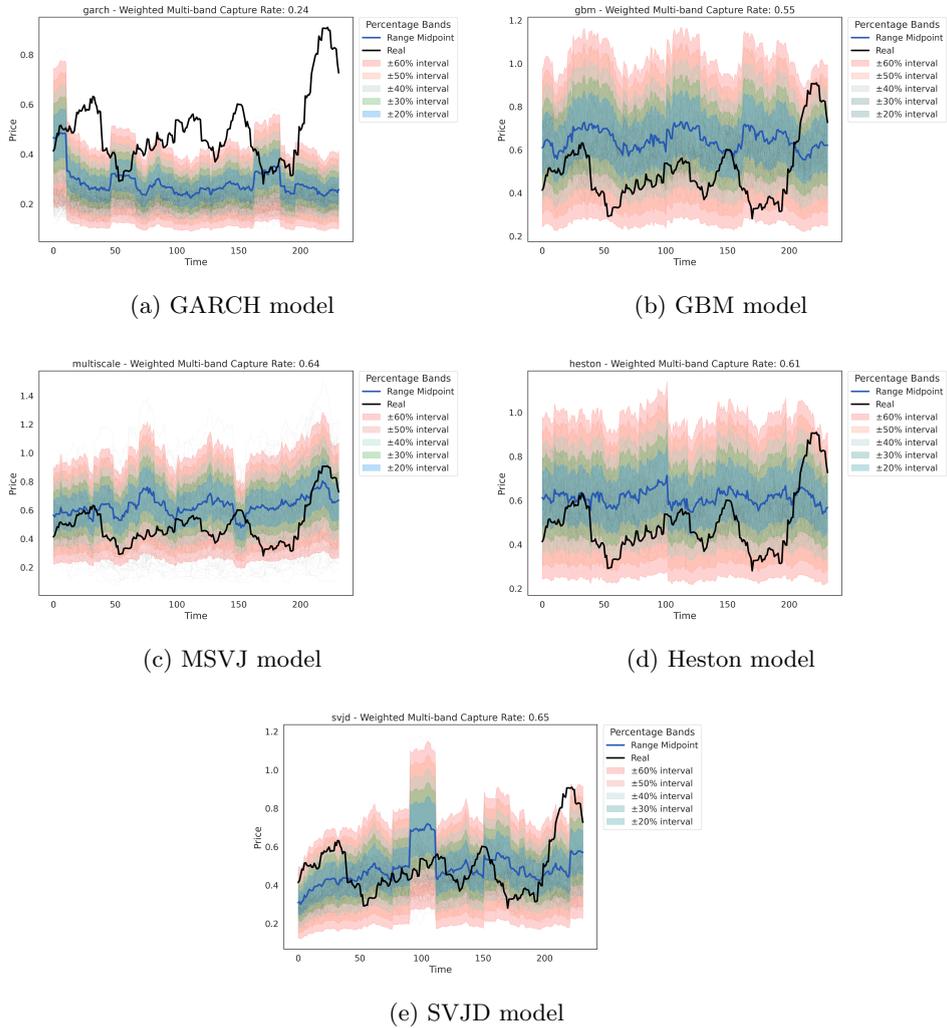
Figure 5: WMCR Volatility for all models from September 2023 to September 2024. A higher value indicates better performance in capturing the volatility dynamics within the specified range.

6.3   Distributional Properties of Returns

To assess the models' ability to reproduce the distributional properties of TQQQ returns, we compare the simulated return distributions with the actual TQQQ return distribution using various metrics and tests, as shown in Table 2.

The Tail-Weighted Anderson-Darling (TWAD) test [4] measures the goodness of fit between the simulated and actual return distributions, with lower values indicating better performance. The MSVJ model achieved the lowest TWAD of 6.8949, suggesting that it best captured the overall distribution of returns, particularly in the tails. The Kolmogorov-Smirnov (KS) test statistic [22], which measures the maximum distance between the empirical cumulative distribution functions of the simulated and actual returns, also favored the MSVJ, with the lowest KS statistic of 0.1196. While the MSVJ model achieved the closest median (0.0013), the GBM model achieved the closest mean (0.0014) and standard deviation (0.0384) to the actual returns. The GBM model also achieved the closest Value at Risk (VaR) and Expected Shortfall (ES) at the 5% significance level to the actual returns, with a VaR of -0.0593 and an ES of -0.0777.

Figure 6 presents the return distributions for the simulated returns of each model compared to the actual TQQQ return distribution from September 2023 to September 2024. The MSVJ model and SVJD model demonstrate the closest alignment with the actual return distribution, consistent with their strong performance in the KS statistic and other distributional metrics. These results suggest that the MSVJ model was the most effective in capturing the overall distribution of returns, particularly in the tails, while the GBM model excelled in reproducing the key distributional properties, such as mean, standard deviation, VaR, and ES.

The results underscore the importance of considering multiple evaluation metrics and choosing models based on the specific objectives and constraints of the application at hand, such as the relative importance of capturing price dynamics, volatility characteristics, or distributional properties.

6.4   Evaluating Historical Crashes

In the previous section, we assessed the models' general performance over the last five years and specifically analyzed their performance from September 2023 to September 2024, which was an overall bullish period. In this section, we assess the performance of different models in capturing the dynamics of bear market periods and historical crash events. We compared the simulated return distributions with the actual TQQQ return distribution during two significant market downturns: the COVID-19 crash in 2020 and the recent 2022 market drawdown driven by recession fears. The COVID-19 crash was characterized by a sharp and rapid decline in stock prices, with the S&P 500 falling by approximately 34% between February and March 2020 [32]. The 2022 market drawdown, on the other hand, was a more prolonged decline, with the S&P 500 dropping 27.5% from its peak and the Nasdaq-

(a) GARCH model

(b) GBM model

(c) MSVJ model

(d) Heston model

(e) SVJD model
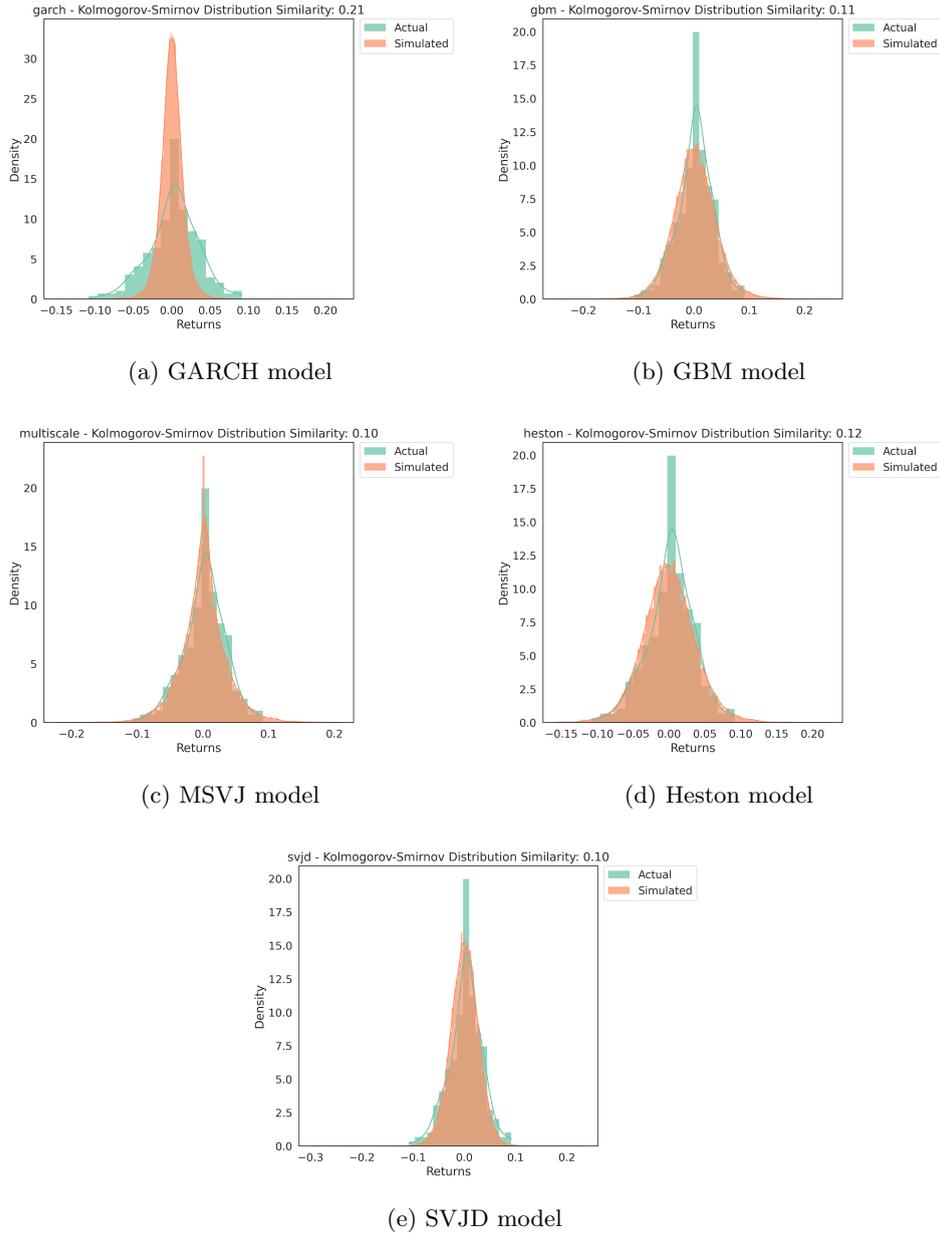
Figure 6: Kolmogorov-Smirnov (KS) test for comparing the return distributions of various models with the actual TQQQ return distribution from September 2023 to September 2024. A lower KS statistic indicates a better fit between the simulated and actual return distributions.

100 (QQQ) declining by 32.5% overall, with TQQQ suffering a drawdown of 80%, marking the worst annual performance since the 2008 financial crisis [26].

To evaluate the models' ability to capture these historical crashes, we focused on two key aspects: the similarity of the simulated return distributions to the actual TQQQ return distribution during the crash periods and the models' ability to generate realistic drawdown scenarios for both rapid and prolonged market declines.

First, we compared the simulated return distributions with the actual TQQQ return distribution during the COVID-19 crash and the 2022 market drawdown using the Kolmogorov-Smirnov (KS) test [22] and the Tail Weighted Anderson-Darling (TWAD) statistic [4]. The KS test measures the maximum absolute difference between the empirical cumulative distribution functions (CDFs) of two samples, while the TWAD statistic places more emphasis on the tails of the distribution, making it better for model evaluation in extreme market events.

Table 3 presents the KS and TWAD statistics for each model during the COVID-19 crash and the 2022 market drawdown, with lower values indicating better performance. The results show that the Heston model generated return distributions that closely matched the actual TQQQ return distribution during the COVID-19 crash, capturing the extreme negative returns and fat tails. The GBM model was the second-best performer in this regard. On the other hand, the GBM model was able to capture the more gradual and prolonged drawdown observed during the 2022 market crash better than the other models, with the Heston model being the second-best performer. These findings suggest that both the Heston and GBM model were well-suited for capturing the return distributions during different types of market crashes.

Model	COVID-19 KS	COVID-19 TWAD	2022 Decline KS	2022 Decline TWAD
MSVJ	0.4396	36.2375	0.2097	43.0122
GBM	0.2622	2.2647	**0.1722**	**31.6651**
SVJD	0.2962	6.9383	0.2392	101.0898
Heston	**0.2495**	**1.6266**	0.1834	39.8259
GARCH	0.4204	27.6852	0.3672	283.5072

Table 3: Kolmogorov-Smirnov (KS) and Tail Weighted Anderson-Darling (TWAD) statistics for each model during historical crash periods.

In addition to the distributional metrics, we assessed the models' ability to generate realistic price paths and price path ranges that simulated the crash periods using the DTW distance and the WMCR metric. The WMCR Price metric measured the proportion of simulated price paths that fall within certain percentage bands around the actual price path, with higher weights assigned to narrower bands. Table 4 presents the WMCR Price values for each model during the COVID-19 crash and the 2022 market drawdown. The results showed that the SVJD model achieved

(a) GARCH model

(b) GBM model

(c) MSVJ model

(d) Heston model
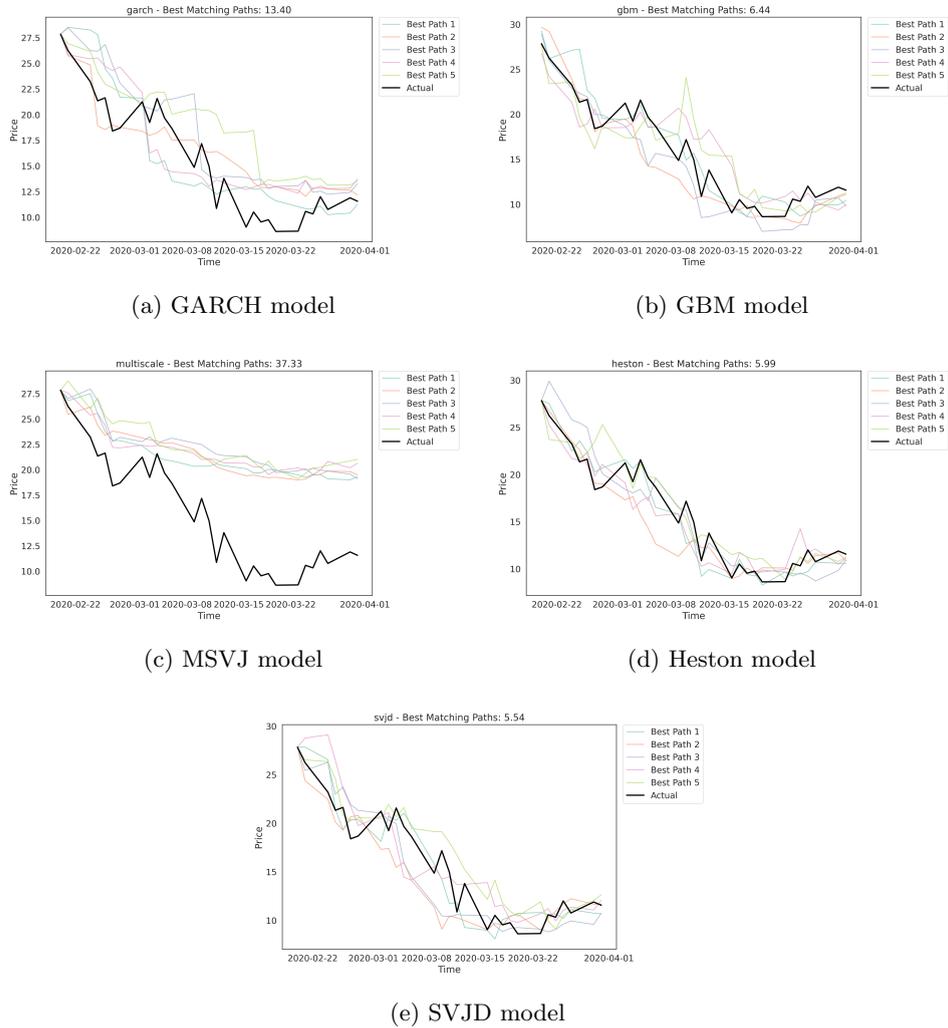
(e) SVJD model

Figure 7: DTW distance for models during the 2020 COVID-19 Market crash. The SVJD model has the lowest DTW distance, indicating better performance in capturing the price path during the crash.

(a) GARCH Model

(b) GBM Model

(c) MSVJ model

(d) Heston Model
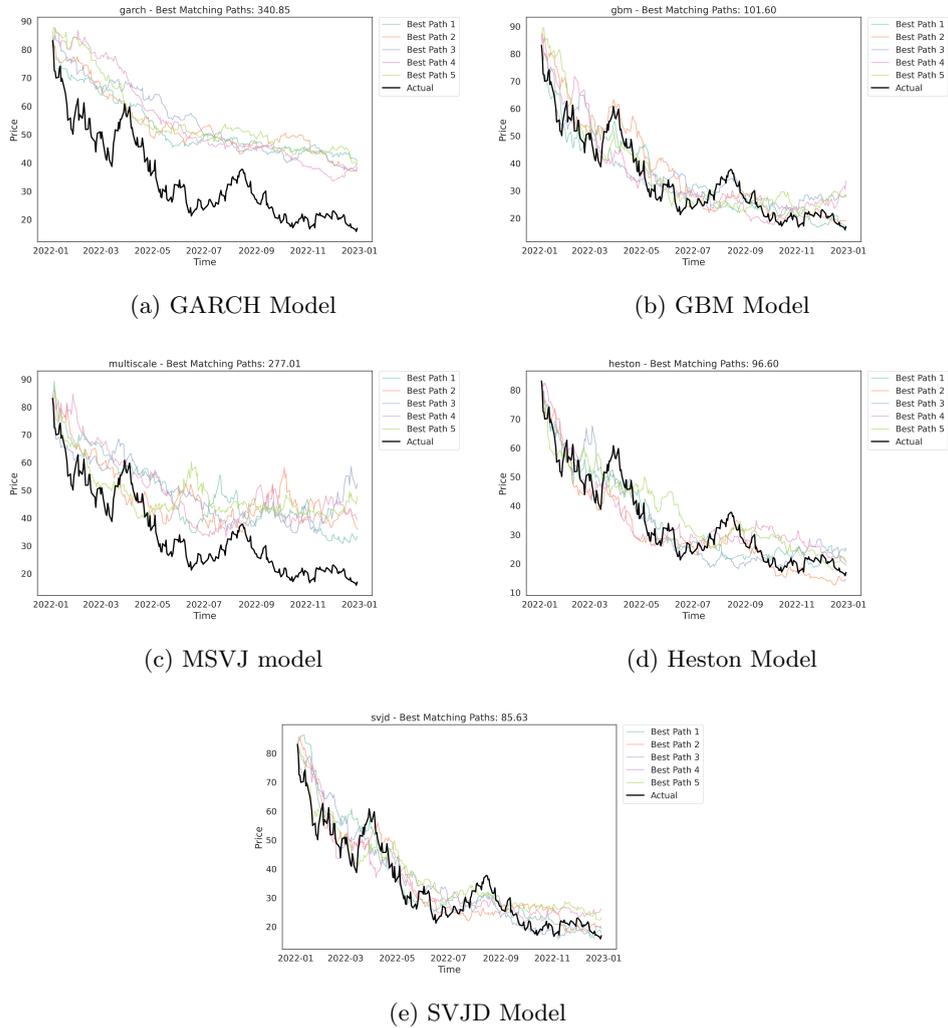
(e) SVJD Model

Figure 8: DTW distance for various models during the 2022 market drawdown. The SVJD model has the lowest DTW distance, indicating better performance in capturing the price path during the decline.

Model	COVID-19 WMCR Price	2022 Recession WMCR Price
MSVJ	0.3057	0.1043
GBM	0.2437	0.1480
SVJD	**0.3862**	**0.1955**
Heston	0.3287	0.1571
GARCH	0.1908	0.0923

Table 4: WMCR Price for models during the 2020 COVID-19 crash and 2022 market drawdown

the highest WMCR Price values for both crash periods, followed by the Heston and MSVJ.

To visualize the similarity between the simulated and actual price paths during the crash periods, we plotted the DTW distances for each model. Figure 8 presents the DTW distance plots for the 2022 market drawdown and Figure 7 shows the DTW distance plots for the COVID-19 crash. In these plots, a smaller DTW distance indicates a higher similarity between the simulated and actual price paths. The DTW distance plots reveal that the SVJD model generated most similar price paths to the actual TQQQ price path during both the 2022 market drawdown and the COVID-19 crash. The Heston model also performs well, ranking as the second-best model in terms of generating price paths that closely resemble the actual TQQQ price path during these crash events.

The evaluation of historical crashes using multiple metrics highlight the strengths of different models in capturing various aspects of market crashes. The Heston and GBM models exceled at reproducing the return distributions during the COVID-19 crash and the 2022 market drawdown, respectively, making them well-suited for risk management purposes.

The SVJD model demonstrated superior performance in generating realistic price paths and price path ranges during both crash periods, as evidenced by high WMCR Price values and low DTW distances. The Heston model offered a good balance between capturing return distributions and generating realistic price paths, with competitive performance in both the DTW distance and WMCR Price metrics.

In summary, the choice of model for simulating LETF dynamics during market crashes should be based on the specific characteristics of the crash event and the desired application. For risk management purposes, the Heston and GBM models may be preferred, while for applications that prioritize accurate price path simulation and realistic price path ranges, such as stress testing or scenario analysis, the SVJD model is the most appropriate choice.

7   Discussion

The findings of this study highlight the importance of selecting the appropriate stochastic model for simulating TQQQ price paths based on the specific objectives, market conditions and applications in quantitative finance.

Application	Best	2nd Best	Key Users & Applications
Modeling Sharp Crashes	Heston	GBM	• **Risk Mgrs:** estimating VaR/ES • **Port. Mgrs:** designing crash protection
Modeling Bear Markets	GBM	Heston	• **Risk Mgrs:** extended drawdown testing • **Port. Mgrs:** defensive rebalancing
Estimating Price Range	MSVJ	GBM	• **Options:** barrier/range option pricing • **Traders:** setting profit targets/stops
Simulating Price Paths	SVJD	GBM	• **Traders:** backtesting strategies • **Risk Mgrs:** stress testing portfolios
Simulating Volatility Paths	SVJD	GBM	• **Traders:** mean reversion strategies • **Risk Mgrs:** volatility regime testing

Table 5: Model performance summary

Table 5 provides a comprehensive summary of model performance across different applications based on our findings.

When the objective is to evaluate and simulate scenarios that reflect market crashes, both short-term events and long-term crises, models such as GBM and the Heston model have been shown to be more effective. These models are better equipped to capture the sudden and severe price movements associated with market crashes, as demonstrated by their performance in reproducing historical drawdowns and their ability to capture tail risk, as evidenced by the lowest Tail-Weighted Anderson-Darling (TWAD) and Kolmogorov-Smirnov (KS) values. Moreover, GBM model's ability to accurately estimate ES and VaR for TQQQ suggests that it is well-suited for risk management applications.

If the objective is to generate future scenario simulations for option pricing, the MSVJ model has proven to be the most suitable choice. The MSVJ model's superior performance in capturing the range of the actual TQQQ price, as evidenced by its highest WMCR for both price and volatility, makes it particularly valuable for option pricing. Options traders can leverage this model's accuracy in estimating potential price ranges to price exotic options more effectively, especially for barrier options and other path-dependent derivatives. The model's high VVS value indicates its ability to capture volatility dynamics, making it especially useful for trading volatility-based strategies such as straddles or strangles, where profitability

depends on accurate volatility forecasts.

When the primary goal is to simulate the most realistic price path and volatility paths for TQQQ, the SVJD model has demonstrated superior performance. By capturing both stochastic volatility and jump processes, the SVJD model can generate price and volatility trajectories that closely resemble the observed dynamics of TQQQ. Portfolio managers can utilize this model for more accurate backtesting of trading strategies and better assessment of portfolio risk under various market conditions. The SVJD model's strong performance in capturing volatility persistence (VPR) and jumps (VJC) makes it particularly valuable for designing volatility arbitrage strategies, where traders seek to profit from the difference between implied and realized volatility.

The choice of model should also take into account the computational complexity and data availability for calibration. More advanced models, such as the MSVJ model, may require more computational resources and longer calibration times compared to simpler models like GBM. Practitioners should weigh these computational costs against the potential benefits of improved accuracy for their specific application.

8   Conclusion

8.1   Key Contributions

This study provides a comprehensive comparative analysis of advanced stochastic models for simulating the price paths of TQQQ, a 3X- leveraged ETF tracking the NASDAQ-100 index. Our findings demonstrate that more sophisticated models, particularly the proposed MSVJ model and the SVJD model, outperform traditional approaches in capturing the complex dynamics of LETFs.

The key contributions of this work include:

1. The introduction of a custom MSVJ model that incorporates both slow and fast volatility components, as well as jump processes. This model effectively captures the complex volatility dynamics observed in LETFs, which are characterized by short-term fluctuations and long-term trends. The MSVJ model outperformed other models in estimating the range of the actual TQQQ price and volatility, as shown by its highest WMCR. It also best captures the distributional properties of returns, evident from its lowest TWAD and KS statistics. Furthermore, the MSVJ model excels in capturing changes in volatility over time, as indicated by its highest VVS value.

2. A comprehensive evaluation framework combining statistical, distributional, path-based, and financial performance metrics that provides insights into their strengths and weaknesses for different applications in quantitative finance.

In conclusion, this study underscores the importance of using sophisticated stochastic models when dealing with LETFs like TQQQ. As these financial in-

struments grow in popularity and complexity, the need for accurate simulation techniques becomes increasingly critical for investors, risk managers, and regulators alike. By providing insights into the best-performing models and their implications, this work contributes to the development of more reliable and effective tools for managing the risks and opportunities associated with LETFs.

8.2   Limitations and Future Work

Despite the breadth of stochastic models considered, several complexities of LETF dynamics may still remain unaddressed. Potential directions include exploring additional models such as the Variance Gamma (VG) [21]which can better capture jump behaviorand the Markov-Switching Multifractal (MSM) [11]which emphasizes volatility clustering.

A particularly promising avenue is the inclusion of option chain data (implied volatility) in conjunction with historical (realized) volatility for model calibration. Our current approach implicitly assumes that historical price data fully encodes the information needed to forecast future dynamics, but this overlooks the forward-looking information embedded in options markets. Integrating implied volatility can benefit both risk management and derivative pricing:

- **Better Parameter Estimation:** Calibration to both observed price data (real-world measure) and optionimplied data (riskneutral measure) enables reconciliation of realworld dynamics with market expectations of future variance. This dual-measure approach improves the capture of volatility term structures and skew while reducing model misspecification in future state simulations.

- **Improved Risk Management:** Integration of implied volatility surfaces enables more accurate tailrisk and stresstesting scenarios, as options markets often anticipate large moves or regime changes before their manifestation in realized volatility. This capability becomes particularly crucial for LETFs, which experience amplified volatility during market dislocations.

- **Extended Derivative Pricing:** Calibration to both historical returns and implied volatility enables direct application to pricing LETFlinked options and other derivatives. This approach yields riskneutral parameter estimates, essential for deriving fair values under standard pricing frameworks.

- **Enhanced Volatility Surface Analysis:** Beyond price path generation, calibration to implied volatility enables simulation of entire volatility surfaces across strikes and maturities, providing a comprehensive framework for scenario analysis. This capability proves particularly valuable for evaluating complex hedging strategies in pathdependent LETF options.

Lastly, the computational aspects of simulating LETFs could be further investigated. Future work could explore the use of more advanced optimization algorithms,

such as Bayesian optimization, particle swarm optimization, or deep learning, for model calibration [16, 27, 33]. These techniques could help identify optimal model parameters more efficiently, especially for complex models with numerous parameters.

By addressing these limitations and exploring the suggested avenues for future research, we can continue to enhance our understanding of LETF dynamics and develop more accurate and reliable simulation tools for these complex financial instruments, ultimately benefiting investors, risk managers, and regulators.

Bibliography

[1] Andrew Ahn, Martin Haugh, and Ashish Jain, *Consistent pricing of options on leveraged etfs*, SIAM Journal on Financial Mathematics **6** (2015), no. 1, 559–593.

[2] Torben G Andersen and Tim Bollerslev, *Intraday periodicity and volatility persistence in financial markets*, Journal of Empirical Finance **4** (1997), no. 2-3, 115–158.

[3] Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Paul Labys, *Modeling and forecasting realized volatility*, Econometrica **71** (2003), no. 2, 579–625.

[4] Theodore W Anderson and Donald A Darling, *A test of goodness of fit*, Journal of the American Statistical Association **49** (1954), no. 268, 765–769.

[5] Marco Avellaneda and Stanley Zhang, *Path-dependence of leveraged etf returns*, SIAM Journal on Financial Mathematics **1** (2010), no. 1, 586–603.

[6] Ole E Barndorff-Nielsen and Neil Shephard, *Power and bipower variation with stochastic volatility and jumps*, Journal of Financial Econometrics **2** (2004), no. 1, 1–37.

[7] David S Bates, *Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options*, The Review of Financial Studies **9** (1996), no. 1, 69–107.

[8] Donald J Berndt and James Clifford, *Using dynamic time warping to find patterns in time series*, KDD workshop, vol. 10, Seattle, WA, USA, 1994, pp. 359–370.

[9] Fischer Black and Myron Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), no. 3, 637–654.

[10] Tim Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics **31** (1986), no. 3, 307–327.

[11] Laurent E. Calvet and Adlai J. Fisher, *How to forecast long-run volatility: Regime switching and the estimation of multifractal processes*, Journal of Financial Econometrics **2** (2004), no. 1, 49–83.

[12] Minder Cheng and Ananth Madhavan, *The dynamics of leveraged and inverse exchange-traded funds*, Journal of Investment Management **7** (2009), no. 4, 43–62.

[13] Fulvio Corsi, Stefan Mittnik, Christian Pigorsch, and Uta Pigorsch, *The volatility of realized volatility*, Econometric Reviews **27** (2008), no. 1-3, 46–78.

[14] Doris Dobi and Marco Avellaneda, *Structural slippage of leveraged etfs*, Journal of Investment Strategies **2** (2012), no. 1, 3–32.

[15] Robert F. Engle, *Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation*, Econometrica **50** (1982), no. 4, 987–1007.

[16] Miloslav Fiura and Jií Witzany, *Historical calibration of svjd models with deep learning*, IES Working Papers 36/2023, Institute of Economic Studies, Faculty of Social Sciences, Charles University, 2023.

[17] Jean-Pierre Fouque, George Papanicolaou, and K. Ronnie Sircar, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, 2000.

[18] Steven L Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies **6** (1993), no. 2, 327–343.

[19] Tim Leung and Marco Santoli, *Leveraged exchange-traded funds: Price dynamics and options valuation*, Springer, Cham, 2016.

[20] Tim Leung and Ronnie Sircar, *Implied volatility of leveraged etf options*, Applied Mathematical Finance **22** (2015), no. 2, 162–188.

[21] Dilip B Madan and Eugene Seneta, *The variance gamma (vg) model for share market returns*, Journal of Business **63** (1990), no. 4, 511–524.

[22] Frank J Massey Jr, *The kolmogorov-smirnov test for goodness of fit*, Journal of the American Statistical Association **46** (1951), no. 253, 68–78.

[23] Robert C. Merton, *Theory of rational option pricing*, The Bell Journal of Economics and Management Science **4** (1973), no. 1, 141–183.

[24] Meinard Muller, *Dynamic time warping*, Information retrieval for music and motion (2007), 69–84.

[25] Svetlozar T Rachev, Christian Menn, and Frank J Fabozzi, *Fat-tailed and skewed asset return distributions: implications for risk management, portfolio selection, and option pricing*, John Wiley & Sons, Hoboken, NJ, 2005.

[26] Reuters, *Wall st ends 2022 with biggest annual drop since 2008*, Reuters (2022).

[27] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams, *Practical bayesian optimization of machine learning algorithms*, Advances in Neural Information Processing Systems, 2012, pp. 2951–2959.

[28] Rainer Storn and Kenneth Price, *Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces*, Journal of Global Optimization **11** (1997), no. 4, 341–359.

[29] Hongfei Tang and Xiaoqing Eleanor Xu, *Solving the return deviation conundrum of leveraged exchange-traded funds*, Journal of Financial and Quantitative Analysis **48** (2013), no. 1, 309–342.

[30] YCharts, *Proshares proshares ultrapro sp500 (upro) total assets under management*, https://ycharts.com/companies/UPRO/total_assets_under_management.

[31] , *Proshares ultrapro qqq (tqqq) total assets under management*, https://ycharts.com/companies/TQQQ/total_assets_under_management.

[32] Dayong Zhang, Min Hu, and Qiang Ji, *Financial markets under the global pandemic of covid-19*, Finance Research Letters **36** (2020), 101528.

[33] Yudong Zhang, Shuihua Wang, and Genlin Ji, *A comprehensive survey on particle swarm optimization algorithm and its applications*, Mathematical Problems in Engineering **2015** (2015), 1–38.