Estimating Option Prices Using Black–Scholes and EGARCH Volatility

The common wisdom of financial theory is that returns are unpredictable as past performance is not indicative of future results. While returns on most assets tend to be uncorrelated at low frequencies – hours, days, etc. – the squared returns tend to display slowly decaying autocorrelation. In other words, while the direction of price change is unknown the magnitude of a change is somewhat predictable. Markets go through calm and turbulent periods – even if returns are uncorrelated, their variance changes. These empirical findings suggest that a good model could give a fair estimate of future volatility.

Ability to estimate volatility of a stock is essential in derivative pricing. A price of a European or American option naturally depends on volatility of a stock in the future. In this project, we will investigate and predict volatility dynamics of S&P500 using family of GARCH models. Then we will use Black-Scholes Equation to estimate price of an option given the predicted volatility. Finally, we will compare our estimates of prices with real prices of options.

Data Description

The data for the daily price of SMP 500 was downloaded from Yahoo Finance for the time period 22/08/2015 – 22/08/2025. The data for options was downloaded from Yahoo Finance with the options expiring at 29/08/2025.

We compute log-returns and squared log returns; below is brief preliminary statistical analysis.

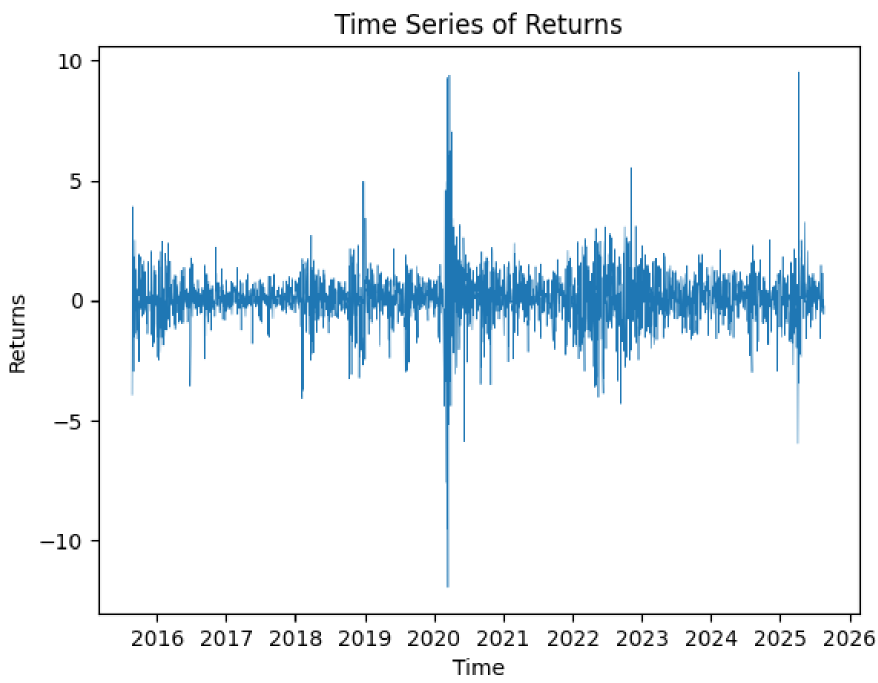
	MEAN	VARIANCE	SKEWNESS	KURTOSIS
PRICE	3599.466	1436526.157	0.509	-0.768
LOG-RETURNS	0.054	1.34	-0.358	14.853
SQUARED RETURNS	1.342	30.147	14.902	289.988

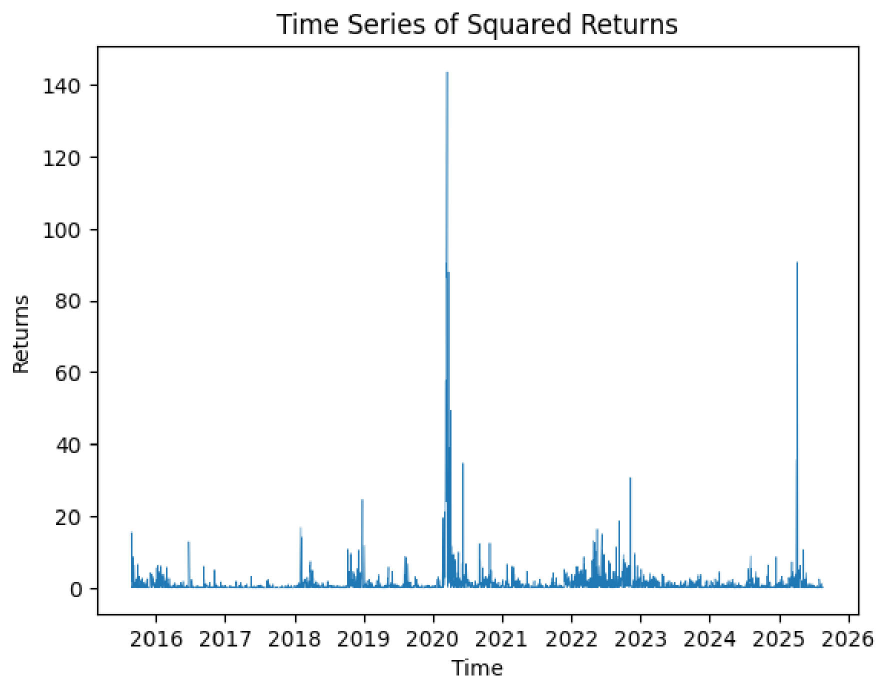
Table 1: Summary Statistics for normal and squared returns

Distribution of the rice of S&P500 index seems to be close to normal. High absolute variance is expected for the index as larger values of price lead to larger absolute swings. Skewness is slightly positive, meaning that prices move upwards more frequently than downwards. Kurtosis is low, slightly flatter than normal distribution but close to it. The prices are fairly symmetric and normally distributed.

Average daily return is slightly above 0, which is realistic for SMP500. Variance and skewness are also reasonable. Slight negative skewness implies more frequent small dips than gains. Kurtosis is 14.85, very leptokurtic, heavy tails large positive or negative returns are more likely than under normal distribution. These heavy tails justify the use of GARCH family models and fat-tailed distributions.

Squared returns are much more dispersed than raw returns – as expected. Extremely positive skewness is due to squaring amplifying the magnitude of shocks and making them positive. Kurtosis of 290 suggests extremely heavy tailed distribution. This confirms **volatility clustering** — most squared returns are small, but occasional very large values dominate the distribution. The use of the GARCH family models is appropriate.





Graphs 1 and 2: time series plots of normal returns and squared returns

Next step is to confirm that the assumptions of volatility-clustering models are satisfied:

- Returns are serially uncorrelated
- Squared returns are serially correlated
- Both normal and squared returns are stationary

First, let's test whether both series are stationary:

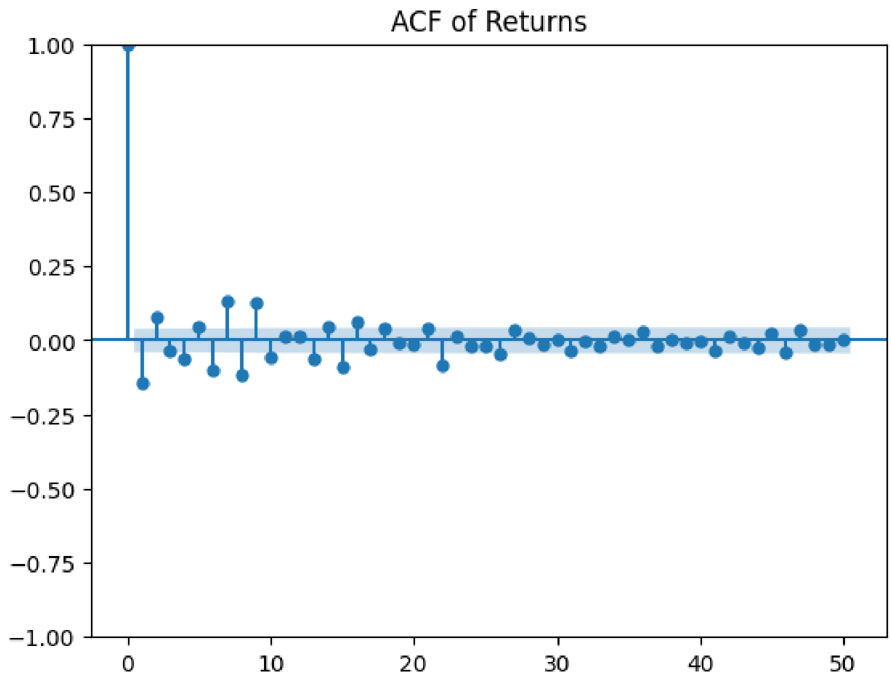
To test this, we use Augmented Dickey-Fuller test; the reason we use ADF test is that the test statistic does not converge in distribution to normal distribution.

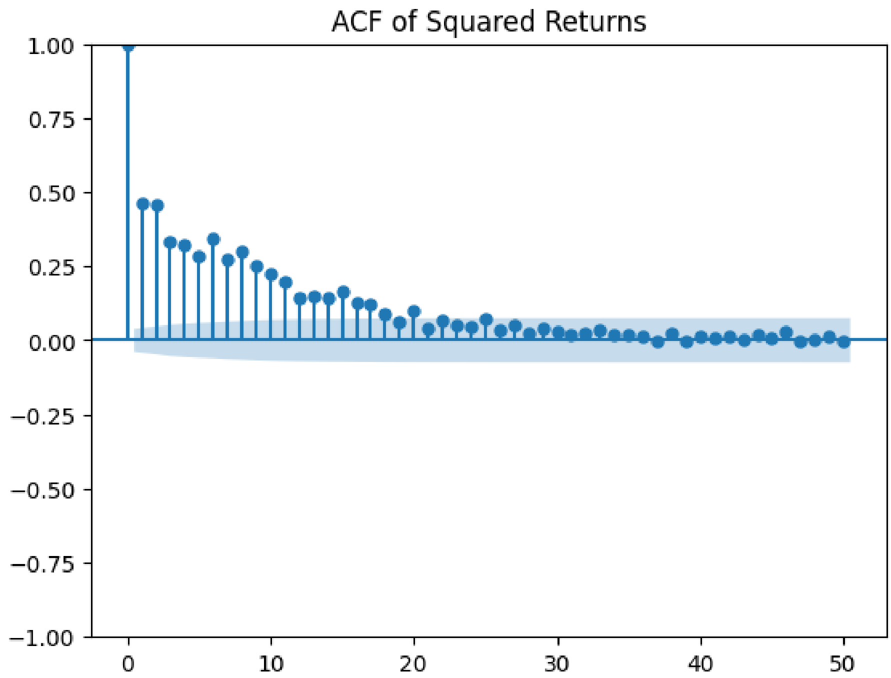
	ADF STATISTIC	SIGNIFICANT AT 5%?
RETURNS	-16.147	Yes
SQUARED RETURNS	-7.135	Yes

Table 2: Results of ADF test for returns and squared returns

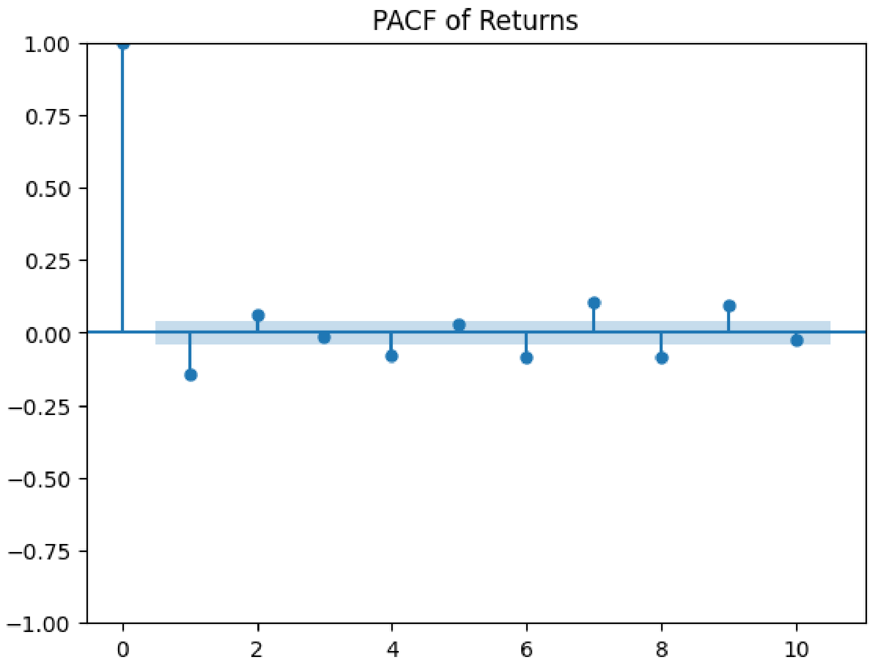
Both series are stationary at 5% significance level.

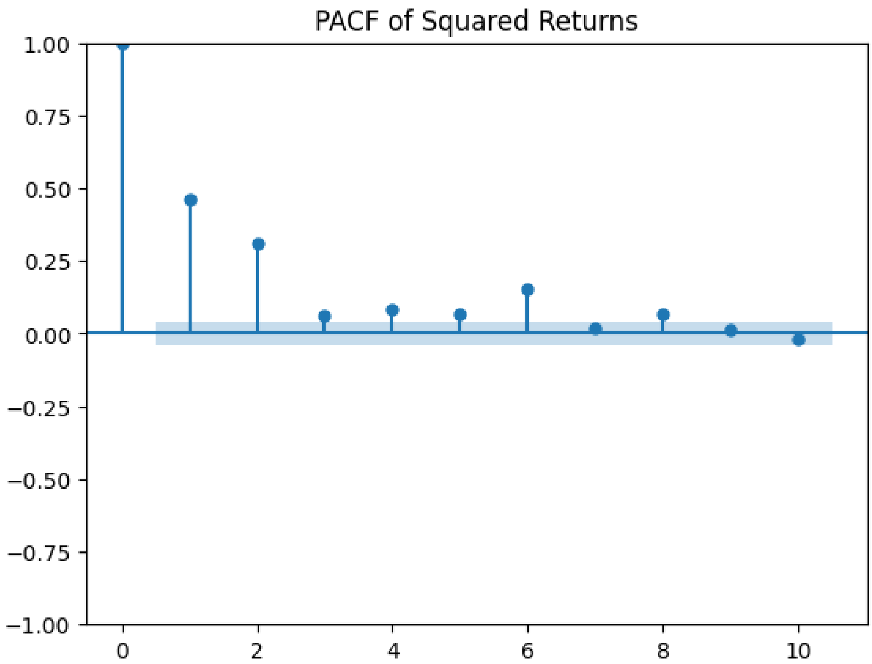
Second, we plot ACF and PACF to see the whether series are serially correlated.





Graphs 3 and 4: ACF of normal and squared returns





Graphs 5 and 6: PACF of normal and squared returns

From ACF plots, we see that normal returns are not serially correlated – ACF cuts of sharply after lag 0 - while for squared returns ACF decays slowly. Similarly, PACF of normal returns cuts off immediately while for squared returns PACF cuts off after second lag. This confirms that S&M 500 returns satisfy assumptions of volatility-clustering models; the direction of a daily change is unpredictable, but the magnitude of change is.

Model Specification

Given the significant autocorrelation in squared terms and evidence of volatility clustering, a model that captures conditional heteroskedasticity is appropriate. PACF of squared returns cuts-off sharply after second lag, we may deduce that the GARCH model will be of order (2,1). However, despite PACF suggesting order (2,1), the second lag is statistically insignificant; thus, the model GARCH(1,1) will be used. Additionally, given that the daily returns are uncorrelated: ACF and PACF of returns are about 0 for all lags >0, we do not need to estimate any models for the mean equation – it will simply be mean + shock. Finally, as the tails of distribution of squared returns are fat with excess kurtosis, we use student t-distribution for modelling innovations z_t.

We estimate the model using maximum likelihood method and get the following equations:

GARCH(1,1)

$$r_t = 0.0975 + \varepsilon_t$$

$$\sigma_t^2 = 0.0196 + 0.1706\varepsilon_{t-1}^2 + 0.8272\sigma_{t-1}^2$$

$$\varepsilon_t = \sigma_t z_t, \qquad z_t \sim t_{v=5.4586}(0,1)$$

While GARCH(1,1) captures volatility clustering and persistence, it has notable limitations. It treats positive and negative shocks symmetrically and can underestimate extreme events. Additionally, there is a problem with α + β ≈ 1, as in this case it seems to lead to very high unconditional variance . The unconditional variance of GARCH(1,1) is given by:

$$Var(r_t) = \frac{w}{1-\alpha_1-\beta_1} = \frac{0.0196}{1-0.1706-0.8272} \approx 8.864$$

This implies a daily standard deviation of:

$$\sigma_{daily} = \sqrt{8.864} \approx 2.977\%$$

Which in turn implies the yearly standard deviation of:

$$\sigma_{annual} \approx \sigma_{daily} \times \sqrt{252} = 47.6\%$$

(Market operates 252 days in a year)

This estimate of volatility is very high and is not consistent with empirical volatility. The high value is partly due to inputting returns as percentages rather than decimals.

These limitations motivate the use of more flexible models such as **EGARCH**, which allow for asymmetric and leverage effects.

EGARCH(1,1)

$$r_t = 0.066 + \varepsilon_t$$

$$\ln(\sigma_t^2) = -0.005 + 0.9709\ln(\sigma_{t-1}^2) + 0.1928\left(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - c\right) - 0.1785\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim t_{5.84}(0,1)$$

$$where\ c \approx 0.77\ is\ the\ mean\ of\ z_t$$

EGARCH captures leverage effect – people react more to negative news. This is captured by the $-0.1785\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ term. The Student-t distribution was used again to account for fat-tails. In EGARCH the unconditional mean does not have a closed-form solution; instead, we can forecast volatility for the next 1000 days and take the average. We get the unconditional variance estimate of 1.3, which converts to $\approx$ 18% yearly volatility that is realistic for S&M500.

Black-Scholes Model

Finally, given the estimated yearly volatility we can use Black-Scholes equation to estimate prices of call options. Black-Scholes formula:

$$C(S_0, K, T, r, \sigma) = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

Where

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \qquad and\ d_2 = d_1 - \sigma\sqrt{T}$$

S_0 - current price of an asset, as of 22/08/2025 price of S&M500 index is 6,466.91$

K – strike price of an option

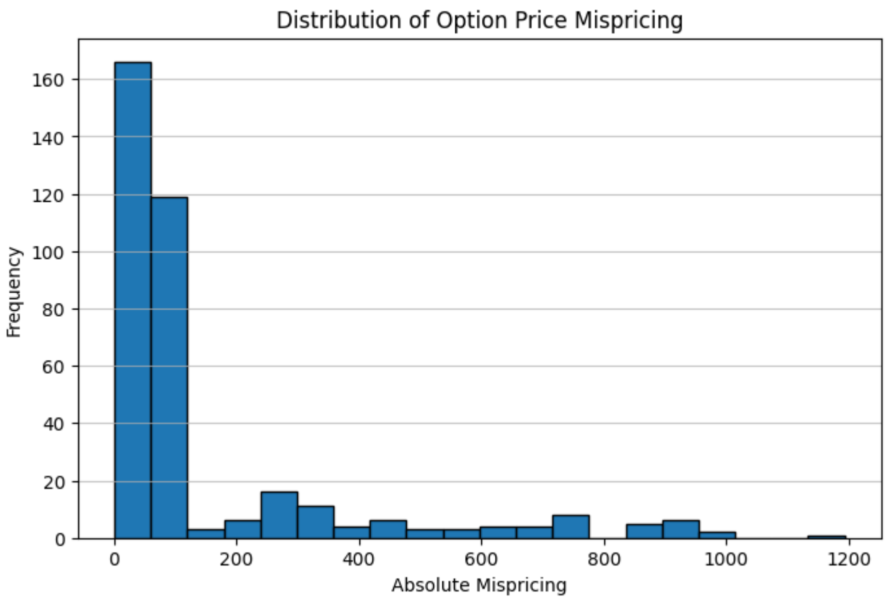
T – time to maturity in years, 5 trading days → $\frac{5}{252}$

r – risk-free interest rate, 10-year U.S. Treasury note yield = 4.3%

σ – annualized volatility, 18% estimated by the E-GARCH model

$N(d_i)$ - cumulative distribution function (CDF) of the standard normal distribution

We download the pricing data for American options expiring at 29/08/2025 and calculate the absolute difference between our predicted price and actual price.



Graph 7: histogram of absolute mispricing, calculated as the difference between estimated price and market price of an option

When comparing the predicted option prices obtained from the Black–Scholes model with the actual market prices, we observe that most of the price differences, or mispricings are relatively small—typically around $100. However, for a few options, the mismatches are much larger. This is not surprising, as these are real options traded by market participants, and actual prices can reflect factors beyond the Black–Scholes assumptions, such as liquidity constraints, transaction costs, market sentiment, or supply and demand imbalances.

Therefore, while the model provides a useful theoretical grounds for option pricing, occasional large deviations are expected in practice, and they indicate complex market dynamics and highlight the limitations of relying solely on theoretical models for pricing.

To obtain an estimate of the yearly volatility of S&M500 we applied EGARCH(1,1). The model, accounting for common features of financial markets - asymmetric impact of positive and negative shocks and clustering volatility – yielded a realistic estimate of 18% for yearly volatility.