# Challenges in Leveraging GARCH and ARMA Models for Predicting Arbitrage Opportunities in Standard American Options

[Redacted] (@floor-licker)
Paper ID *****
May 1, 2023

Conforms to CVPR 2024 Submission Standard as of May 1, 2023

## Abstract

*This paper examines the limitations of GARCH and ARMA models in predicting arbitrage opportunities in the context of Standard American Options. My initial argument was that the characteristics of these models, primarily designed for modeling volatility and linear relationships in time series data, are not optimally aligned with the dynamic and non-linear nature of options markets, but that there may be some exceptions to this rule in some where we can transform and difference data to meet assumptions that are not otherwise met.*

## 1. Introduction

Standard American options grant their holders the right to exercise at any time before expiration. This flexibility introduces pricing dynamics that are influenced by market volatility, interest rates, and dividend yields. To effectively predict arbitrage opportunities in such a dynamic environment, I raise the question of whether GARCH and ARMA models may have specific use cases when price fluctuations, although inherently non-linear, may be observed to tend toward linear outcomes after data transformation or differencing.

The key distinction lies in the nature of continuous versus discrete systems – the former characterized by uninterrupted, seamless state changes, and the latter by distinct, separate intervals. However, an interesting phenomenon arises when discrete systems are observed with sufficient granularity: over an extended series of measurements, these systems begin to exhibit characteristics reminiscent of continuous dynamics. This convergence point between discrete measurements and continuous behavior provides a fertile ground for exploring the utility of GARCH and ARMA models in options pricing. I explore the exact bounds of this hypothesis with respect to the constraint of overfitting by case selection as well as by transformation and differencing of time-series data.

### 1.1. Understanding GARCH

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is predominantly used in financial modeling to predict and analyze the volatility of asset returns over time. It is mostly effective in capturing the 'volatility clustering' phenomenon, where high volatility periods tend to be followed by high volatility, and low volatility by low volatility. GARCH is widely employed in risk management for forecasting future volatility, which is crucial in pricing derivatives and managing portfolio risk. Its applications extend to various financial sectors, including stock, bond, and foreign exchange markets.

### 1.2. Understanding ARMA

The Autoregressive Moving Average (ARMA) model is an important tool in time series analysis, used to forecast and understand the patterns in stationary time series. It combines autoregression (AR), where future values are predicted based on past values, with moving average (MA), modeling the error term as a linear combination of past error terms. This blend enables ARMA to efficiently capture and predict time-dependent patterns with fewer parameters. Widely applied in economic and sales forecasting, ARMA is essential for analyzing time series where understanding historical trends is key to predicting future behavior. Its utility extends to fields requiring precise time series modeling, such as market analysis and weather forecasting.

### 1.3 Arbitrage with Standard American Options

For clarity's sake, arbitrage in the context of standard American options is most easily demonstrated through strategy known as a straddle, where one would purchase both a put option and call option for a given stock with the same strike price. In this sense, the investor is betting on the overall volatility to be seen by the underlying asset as opposed to a specific direction for the price to change in. This is because we end up only executing whichever contract yields the profit, so what is really being bet on is that one contract or the other will return a profit that is greater than the sum of the premiums for the two contracts. In this sense the strategy is called arbitrage as you are assuming a mismatch between the premium cost and the volatility of the underlying asset such that you can "*bet both ways*" just as in sports betting arbitrage, and still net a profit. Now it becomes clear that what is actually important is our ability to forecast volatility of the underlying asset effectively rather than guess the direction of price fluctuations. Implied volatility can be derived from the premium price using Black-Scholes, so to detect these opportunities, as well as the effectiveness of GARCH and ARMA models, we search for points in time where GARCH or ARMA predict a level of volatility that is higher than the implied volatility.

## 2.0 Challenges and Issues

We initially face at least a handful of issues in attempting to set up this experiment properly, namely due to the assumptions made by the models in question.

### 2.1 Catalysts of Non-Stationary Data in Financial Markets

First is changing volatility regimes. Financial markets experience periods of high and low volatility driven by macroeconomic events, policy changes, and market sentiment shifts. This is called volatility clustering, and it implies that the variance of returns changes over time, violating the assumption of constant variance required for stationarity.

Next is trends and random walks. Many financial time series, such as stock prices, exhibit trends or follow a random walk, where the mean drifts over time. This non-stationary behavior makes it difficult for models like ARMA, to accurately predict future values.

Impact of external shocks: External shocks, such as geopolitical events or earnings announcements can introduce sudden and unpredictable changes in the mean and variance of the time series.

Finally, non-linearity in volatility dynamics: The relationship between past and future volatility is often non-linear, while GARCH and ARMA models assume at least a semi-linear structure. This mismatch reduces their effectiveness in capturing the complexities of financial volatility.

### 2.2 Stationarity assumption of GARCH and ARMA Models

Financial time series data often exhibit non-stationarity, violating the stationarity assumption of GARCH and ARMA models. GARCH models are designed to predict volatility by assuming that periods of high or low volatility tend to cluster, but they need the overall structure of the data to remain stable over time for these predictions to hold. Similarly, ARMA models work by looking for repeating patterns in past data, assuming that these patterns will persist into the future. In this case we may apply differencing or  transformations to achieve stationarity and use formal tests to assess it.

## 3.0 Differencing and Transformation of Financial Time-Series Data

Differencing is a technique used to make data more stable by removing trends and shifts in the mean. This involves subtracting the value of a data point from the value of the previous data point, effectively focusing the analysis on changes or "returns" rather than absolute levels. As an example, instead of looking at the actual price of a stock, differencing might look at how much the price changes from one day to the next.

Transformations, for example taking logarithms, are a common method to prepare financial data for GARCH and ARMA models. These transformations help stabilize the variance, especially in cases where large price swings dominate the data. By converting absolute price values to percentage changes or log returns, we could reduce the impact of extreme fluctuations and make the data more consistent over time. These steps are necessary because GARCH and ARMA models depend on identifying patterns in volatility or mean reversion, and these patterns are much harder to detect in raw non-stationary data.

## 4.0 Why Differencing and Transformation of Financial Data is Futile

While differencing and transforming financial time series data can technically make it stationary enough to apply GARCH and ARMA models, this approach turns out to be logically redundant. By differencing or applying other transformations to stabilize the mean and variance, we are essentially creating a data set that assumes some level of mean reversion, which is the property that values will tend to revert to an average over time. If we are already making this assumption through the transformation process, then we could achieve similar results using much simpler mean-reversion strategies, without resorting to the complexity of GARCH and ARMA models.

This redundancy raises the question: why go through the elaborate process of modifying the data to meet the stationarity assumption when simpler methods are available. Mean reversion strategies, such as Bollinger Bands[1] or pairs trading are computationally efficient and don't require the discretionary data manipulation that would be demanded to implement GARCH and ARMA models for the same use case. These simpler methods rely on observable metrics thereby avoiding the black-box nature and complexity of transformed time series models.

Finally, the discretionary modification of financial data—such as deciding how and when to difference or transform—introduces another layer of potential bias, as well as an obviously large margin for human error. It's difficult to quantitatively assess and account for things like the measured financial impact of black swan geopolitical events. By altering the raw data, we are risking distorting the underlying market dynamics, which can lead to models that are overfitted to the transformed data but also might not even reflect real-world behavior.

## 5.0 Conclusion

Ultimately, if transformation and differencing effectively provide us with a data set which allows us to assume stationarity and mean-reversion, then simpler and more transparent mean-reversion strategies can be employed instead. These methods avoid the issues of over-engineering and allow for more direct application to real-time market conditions, making them both practical and conceptually cleaner. In this context, the complexity of GARCH and ARMA modeling becomes difficult to justify, especially given the availability of alternative strategies that achieve comparable or better outcomes with less effort and no human discretion in the manipulation of the input data.

[1] Bollinger On Bollinger Bands – The Seminar, I ISBN 978-0-9726111-0-7