Proceedings of ICMRED 2026 Symposium: Financial Innovation, Risk Governance, and the Dynamics of Global Capital Flows
DOI: 10.54254/2754-1169/2026.BL34132

Modeling the Ups and Downs: Using Asymmetric ARCH to Understand and Predict Commodity Price Volatility —A Case Study of WTI Crude Oil

Jiacheng Zhu

Jiangsu Huaiyin Middle School, Huai'an, China
2706488685@qq.com

Abstract. The dramatic fluctuations in commodity prices, such as crude oil, are a classic market phenomenon discussed in AP Economics. Traditional volatility models like GARCH assume that positive and negative shocks have symmetric effects on volatility, which contradicts the real-world intuition that "bad news often causes greater panic." This study focuses on WTI crude oil futures and attempts to combine Python programming with statistical knowledge to model and forecast price volatility using TARCH and EGARCH models, which can distinguish the different impacts of "good news" and "bad news." The core question is: Among models accessible to a high school student with available tools, which one better reflects the volatility characteristics of the crude oil market? Using daily WTI futures data, this study compares the standard GARCH(1,1) model with two asymmetric alternatives: TARCH and EGARCH. The empirical results reveal a statistically significant leverage effect, meaning that negative shocks ("bad news") increase future volatility more than positive shocks of the same magnitude. Among the three models, EGARCH provides the best in-sample fit and the most accurate out-of-sample forecasts, particularly in capturing extreme volatility clusters. The expected results verify the superiority of asymmetric models and provide a quantitative analysis case from a high school student's perspective for understanding market risk. This study concludes that ignoring asymmetry leads to underestimation of risk during market downturns, and that EGARCH is a practical choice for high school-level researchers to model commodity price volatility.

Keywords: Commodity price volatility, leverage effect, ARCH models, Python data analysis, WTI crude oil

1. Introduction

In AP Economics, the relationship between supply and demand is taught as the primary determinant of prices. However, in reality, commodities like crude oil do not follow smooth curves; they exhibit violent and unpredictable fluctuations. These fluctuations directly affect the global economy and raise the question: can more precise mathematical models be used to describe and forecast such volatility [1]?

DOI: 10.54254/2754-1169/2026.BL34132

Through self-study, it was discovered that the ARCH/GARCH family of models is commonly used in finance to capture "volatility clustering" [2, 3]. However, the standard models assume symmetry—that positive and negative shocks have the same effect on volatility—which contradicts the observed asymmetric market reactions to news (e.g., a war-related drop is often more violent than a peace-related rise). This leads to the research question: Does this asymmetry exist significantly in the crude oil market? In finance, this "bad news has a greater impact" phenomenon is called the "leverage effect" [4].

Therefore, this paper applies the hypothesis testing and regression concepts learned in AP Statistics, together with Python programming, to model WTI crude oil daily price data. The classic GARCH model is compared with two asymmetric models: TARCH (Threshold ARCH) and EGARCH (Exponential GARCH) [5, 6]. It is hoped that the analytical process and conclusions can provide a concrete, reproducible quantitative analysis example for peers to understand complex financial market volatility.

2. The beginning of the research: from AP classroom questions to real-world data exploration

2.1. The gap between the "ideal market" and the "real market"

In AP Economics, the classic supply-and-demand model is studied. Textbook diagrams show prices and quantities reaching equilibrium at the intersection of smooth curves. Yet when observing news of WTI crude oil prices plunging more than 10% in a single day, a question arises: why are real-world price movements so volatile? The smooth curves in textbooks cannot explain such dramatic swings. This sparked interest in market volatility and revealed that real markets are far more complex than idealized models.

2.2. The observed "asymmetric market reaction"

To test this intuition, key events were recorded. For example, after the Silicon Valley Bank collapse in March 2023, WTI crude oil prices fell sharply over the following trading days, with volatility markedly increasing. In contrast, when OPEC+ announced production cuts in June 2023, prices rose, but the increase was relatively mild and volatility did not spike as much. Through these comparisons, it was intuitively felt that "bad news" seemed to shake the market more than "good news." Later, in self-study, this phenomenon was identified as the "leverage effect" in finance.

2.3. First attempt: the inadequacy of static volatility measures

A natural starting point for analyzing volatility is the standard deviation of returns, which can be easily calculated in Excel. Applying this to WTI crude oil prices over rolling windows, however, revealed a fundamental limitation: the measure is static. It compresses information from a fixed time interval into a single value, obscuring any temporal structure. For example, during the Silicon Valley Bank crisis in March 2023, the market experienced several days of elevated turbulence. While Excel could compute a standard deviation for that period, it could not differentiate between the preevent calm and the postshock volatility, nor could it capture the tendency of large movements to cluster—a phenomenon known as "volatility clustering." This static approach thus fails to describe the dynamic evolution of volatility, motivating the use of timeseries models such as ARCH.

DOI: 10.54254/2754-1169/2026.BL34132

3. Self-teaching the ARCH family of models

3.1. How the ARCH model was encountered

While taking a self-paced financial econometrics course on Coursera, the concepts of "volatility clustering" and ARCH models were first encountered. The course explained that ARCH models can describe time-varying conditional variance, which perfectly matched the curiosity about volatility dynamics. Later, Python's arch library was discovered on GitHub, and it was decided to try solving this problem programmatically. For a comprehensive introduction to financial econometrics and time-series modeling, Brooks [7] provides an excellent reference.

3.2. Learning notes: key breakthroughs in understanding "asymmetric effects"

Understanding this asymmetry required a shift in perspective. A simple diagram contrasting the symmetric response of GARCH with the asymmetric responses of TARCH and EGARCH made the concept clear. GARCH assumes that positive and negative shocks have symmetric effects on volatility, whereas TARCH and EGARCH allow shocks of different signs to have different impacts. An analogy was used to help understand: market sentiment is like a spring—pushing it down (bad news) requires more energy to restore than pulling it up (good news), so bad news causes greater volatility.

3.3. From observation to model: understanding the mathematical expression of the leverage effect in TARCH and EGARCH

The TARCH model captures these asymmetric impacts by incorporating a dummy variable into the variance equation:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\epsilon_{t-1}^2 I_{t-1} + \beta\sigma_{t-1}^2 \tag{1}$$

where $It-1=1It-1=1$ if $\epsilon t-1<0\epsilon t-1<0$. This specification directly allows negative shocks ($\epsilon t-1<0\epsilon t-1<0$) to have a different $\alpha+\gamma\alpha+\gamma$

impact on volatility compared to positive shocks (α).

$$\ln\left(\sigma_t^2\right) = \omega + \alpha\left(\frac{|\epsilon_{t-1}|}{\sigma_t - 1} - \sqrt{\frac{2}{\pi}}\right) + \gamma\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2}$$

Detailed derivations and applications of these asymmetric models in Chinese financial research can be found in the paper *Financial Econometrics: From Basic to Advanced Modeling Methods* [8].

4. Research design: a real "data science project" practice

4.1. The twists and turns of data acquisition

Initially, an attempt was made to obtain data from the EIA (U.S. Energy Information Administration), but its API was found to be unfriendly for beginners, with complex documentation. Eventually, Yahoo Finance was chosen because it offers a simple Python interface (yfinance library) [9]. McKinney [10] provides a thorough treatment of data manipulation with Python, which was invaluable for cleaning and preparing the price series. During the first download, a time-zone misalignment caused data mismatches, and it took an entire day to resolve the issue by reading documentation and forum posts.

4.2. The thought process behind model selection

The simplest GARCH(1,1) model was started with, following the "from simple to complex" learning principle [2]. Then TARCH and EGARCH were introduced to test whether they could better capture the different effects of "good news" and "bad news" on volatility [5, 6]. Comparing these three models allowed both a baseline to be established and asymmetry to be explored.

4.3. Debugging in programming

When fitting the models, the error ValueError: The computed initial MA covariance matrix is not positive definite and repeatedly stopped the script. Consulting the Arch library documentation suggested issues with starting values, and a Stack Overflow post pointed to the optimizer. After experimenting, setting the tol parameter to 1e-5 and explicitly providing starting values based on a preliminary GARCH(1,1) fit resolved the issue. This allowed the code to loop through multiple parameter combinations without crashing. The code evolved from single fits to looping over multiple parameter combinations to find the optimal model.

5. Empirical results and analysis

5.1. Parameter estimates: first seeing statistical evidence

When the "volatility persistence" parameter($\alpha+\beta$)in the GARCH model was observed to be close to 1, it was realized that the market indeed has a "memory"—past volatility affects future volatility. Moreover, the "leverage effect" coefficients in TARCH and EGARCH were significantly positive, confirming the observation that bad news has a greater impact on the market. This validation of intuition by data was both exciting and surprising. These findings are consistent with earlier empirical studies on petroleum futures volatility, such as Sadorsky [11], who also documented strong asymmetric effects in oil markets.

5.2. Model diagnostics: testing models with AP statistics knowledge

The Ljung-Box test was applied to the model residuals, and p-values greater than 0.05 were found, indicating no significant autocorrelation—the models had captured the volatility patterns well [12]. The QQ plots of residuals showed some deviations from the theoretical line in the tails, suggesting that the models still have some bias in fitting extreme events, but overall they performed satisfactorily.

5.3. Intuitive comparison: which model looks more like real volatility?

Three comparison charts were plotted showing the true volatility (approximated by squared returns) against the fitted volatilities from GARCH, TARCH, and EGARCH. Among the three models, EGARCH stood out. Its fitted curve most closely tracked the true volatility trajectory, particularly during periods of extreme fluctuation, whereas the GARCH model produced an overly smooth estimate.

DOI: 10.54254/2754-1169/2026.BL34132

6. Forecasting experiment: facing the challenge of the unknown

6.1. Simulating real-world decisions: rolling window forecast design

A rolling window forecast experiment was designed: using the past 250 trading days to predict volatility for the next 20 days, then rolling forward step by step to simulate the continuous forecasting process in real investment decisions. The root mean square error (RMSE) and mean absolute error (MAE) were calculated to evaluate forecast accuracy. The results showed that short-term forecasts (e.g., next 5 days) had much smaller errors than long-term forecasts (e.g., next 20 days), indicating that volatility predictability decays over time.

6.2. A failed forecast and reflection

In September 2023, the model significantly underestimated actual volatility. Post-analysis revealed that it coincided with a Federal Reserve meeting, where market expectations shifted dramatically—an external event the model did not account for. This made it clear that a univariate time-series model cannot capture all external shocks; future work should incorporate more fundamental factors.

7. Conclusion

Through empirical analysis of WTI crude oil daily data, this study discovered a statistically significant leverage effect, confirming that negative shocks ("bad news") increase future volatility more than positive shocks of the same magnitude. Among the three models tested, the EGARCH model proved superior to both the symmetric GARCH and the TARCH model in describing volatility dynamics and forecasting future fluctuations. This finding directly answers the core research question: for a commodity market like crude oil, models that account for "bad news having a greater impact" are indeed closer to reality. In addition, this paper provided a practical bridge between textbook economic theories and real-world data analysis, demonstrating the value of programming (Python) as an essential tool for empirical research, the process itself was deeply educational.

Nevertheless, this study has three limitations: First, using daily data may miss important intraday volatility signals, such as flash crashes triggered by unexpected events. Second, fundamental factors like the U.S. dollar index or crude oil inventories were not included, even though they may have significant impacts on volatility. Third, the scope was limited to univariate GARCH-family models; more advanced multivariate specifications, such as DCC-GARCH, which captures volatility spillovers across different assets, were not considered due to my current level. If time permitted, future study could explore three extensions: applying the same framework to gold or copper to compare volatility characteristics across commodities; learning to compute Value at Risk (VaR) using my best model to understand risk management; and exploring whether machine learning methods like LSTM could improve volatility forecasting [9].

References

[1] Mankiw, N. G. (2018). Principles of Economics (8th ed.). Cengage Learning.
[2] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987-1007.
[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

DOI: 10.54254/2754-1169/2026.BL34132

[4] Black, F. (1976). Studies of stock price volatility changes. Proceedings of the 1976 Meetings of the American Statistical Association, pp.177-181.
[5] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59(2), 347-370.
[6] Zakoian, J.M. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and Control, 18(5), 931-955.
[7] Brooks, C. (2019). Introductory Econometrics for Finance (4th ed.). Cambridge University Press.
[8] Zhang, C. S. (2018). Financial Econometrics: From Basic to Advanced Modeling Methods. Dongbei University of Finance and Economics Press. (Original work published in Chinese). Retrievedfromhttps: //book.cnki.net/bookdetail/bookdetail?filename=ZLCJ201802001038& v=.
[9] Yahoo Finance. (2024). Crude Oil WTI Futures (CL=F) Historical Data. Retrieved January, 26. https: //finance.yahoo.com/quote/CL=F/history
[10] McKinney, W. (2017). Python for Data Analysis (2nd ed.). O'Reilly Media.
[11] Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. Energy Economics, 28(4), 467-488.
[12] Ljung, G. M., Box, G. E. P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.