AI Driven Asset Pricing: The Impact of ChatGPT and AI Announcements on Tech Stocks (ETF's)

Saakshi Dedhia, Xin Qin, Yaxi Chen, and Mao Li

Professor Ratika Narag - ECON 515

May 11th, 2025

Acknowledgement: We would like to express our sincere gratitude to Professor Ratika Narag for her guidance, feedback, and support throughout this research project. Lastly, this project would not have been the same without the collaboration, ideas, and effort shared by our entire team.

Abstract

This paper explores how the AI announcements sentiment scores, especially those involving ChatGPT and similar technologies, affect the prices of tech-focused exchange-traded funds (ETFs). By analyzing how markets respond to major AI-related news events, we find that these announcements often lead to noticeable short-term changes in ETF prices. Using selected case studies from 2021 to 2022, we use sentiment scores to examine how investor excitement or concern about AI developments may be influencing the value of technology ETF. Our analysis suggests that AI has become a key driver of market sentiment, and its influence is increasingly reflected in the pricing of tech-related financial assets. These findings show the importance of understanding how emerging technologies like AI can shape investor behavior and asset values in today's markets.

Keywords: ChatGPT, ETF returns, Sentiment analysis

1. Introduction

The public release of OpenAI's ChatGPT in late 2022 marked an ideal shift in the discourse surrounding artificial intelligence (AI), transforming it from a specialized tool into a mainstream driver of business innovation and public imagination; and what followed wasn't just a product launch, but a cultural and economic shockwave. Within five days, the tool crossed one million users, outperforming tech giants like Facebook, Twitter, and Spotify by a massive margin (see Figure 1). In a world saturated with apps and short-lived tech hype, ChatGPT stood out not just for its capabilities, but for how quickly it rewired conversations around productivity, automation, and the future of artificial intelligence. Suddenly, everyone from college students to CEOs was trying it out, posting about it, and asking the same question: *Is this the future of work?* What made it even more interesting wasn't just how powerful it was, but how fast it went from niche AI model to global obsession. And as with anything that captures mass attention that quickly, the markets started paying attention too.

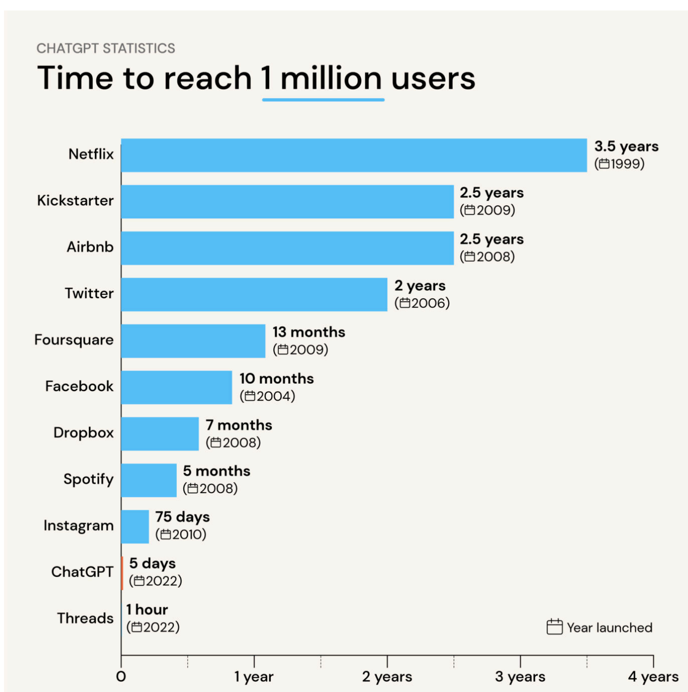


Figure 1: The Fastest App to One Million Users
Source: https://www.tooltester.com/en/blog/chatgpt-statistics/

Within weeks, companies were scrambling to announce their own ChatGPT integrations. Some were legit, many were just buzzwords. But either way, tech stocks moved. ETFs tracking AI-heavy or innovation-focused companies started reacting to headlines, not just earnings reports. At the same time, Twitter exploded with ChatGPT-related chatter. We saw tweet volumes skyrocket while the average sentiment around those tweets began to dip. It was classic hype cycle behavior: lots of noise, mixed emotions, and markets caught somewhere in the middle. This disconnect between rising attention and cooling sentiment hints at a behavioral finance story. It suggests a market dynamic where narratives and expectations can temporarily overpower valuation models, especially in sectors like technology where hype and innovation walk hand-in-hand.

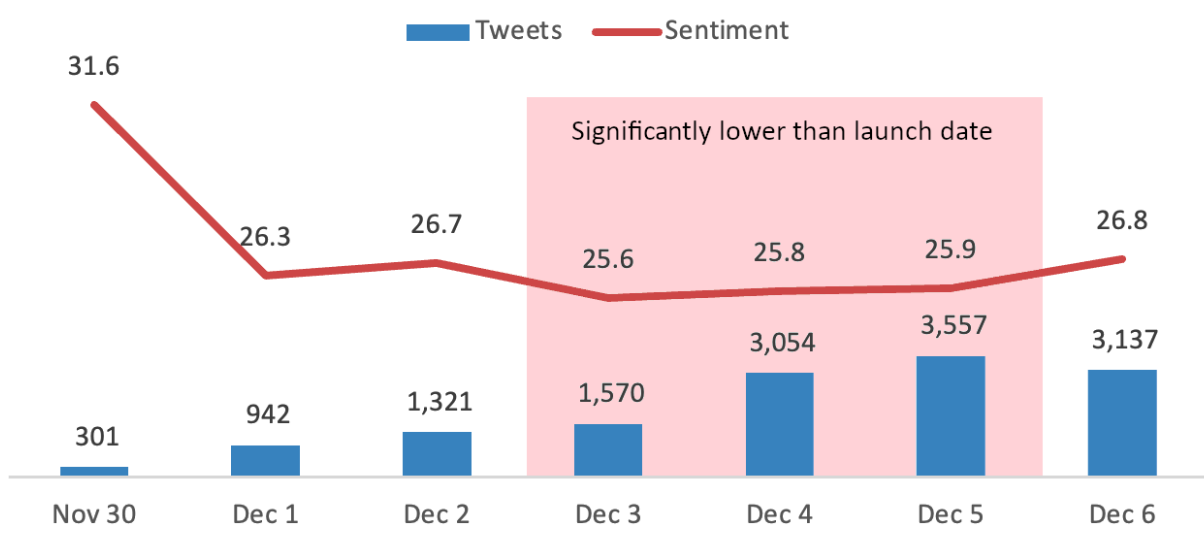


Figure 2: The Hype-Sentiment Gap: Tweet Volume Surges, Optimism Fades
Source: https://towardsdataprocessing.com/analysing-twitters-reaction-to-chatgpt-2277fcac2ba7

This is where our study comes in. We wanted to understand whether AI-related announcements, especially those centered around ChatGPT, actually impact the short-term returns of technology-sector ETFs. Unlike firm-level studies that focus on individual stock reactions, we take a broader lens. ETFs offer an aggregated view of investor expectations toward an entire sector, acting as a real-time barometer of market mood. More importantly, we wanted to know *how*. Is it just the attention? Is it sentiment? Or are these announcements functioning as real information shocks that investors use to update their expectations about future performance? ETFs like WTAI and QQQ gave us a perfect testing ground, since they bundle up many of the firms riding the AI wave and react quickly to tech narratives.

To dig into this, we used a mix of event study methods and time-series modeling. We looked at sentiment scores pulled from AI-related tweets, ETF return data, NASDAQ performance, and volatility via the VIX. First, we ran a simple OLS regression to see if sentiment and returns were even linked. Then, we built on that with ARIMAX models to control for broader market conditions and test how much of the movement in ETF returns could actually be explained by AI sentiment. The idea wasn't just to build a statistical model, but to understand the psychology behind the price moves.

Because at the end of the day, ChatGPT isn't just a product. It's a symbol of how fast tech can shift narratives and how quickly markets can react. This paper is our attempt to capture that reaction in numbers. It's about measuring the financial impact of a cultural moment, and seeing what it tells us about sentiment, pricing, and the future of AI in financial markets.

2. Literature Review

Pietrzak (2025) explored the financial market response through an event study focusing on US-listed firms. The study found that companies referencing ChatGPT or similar technologies in their earnings calls or press releases often experienced statistically significant abnormal returns. These effects were especially pronounced in sectors tied to innovation, like software and consumer technology. The returns were not just tied to actual AI implementation but also to investor perception and how companies positioned themselves within the AI narrative. In other words, the market was reacting as much to the story as it was to the substance.

A striking example of this dynamic came from the education technology space. When Chegg's CEO acknowledged in an earnings call that ChatGPT was beginning to impact their user engagement, the company's stock dropped nearly 50 percent. This sharp reaction was not due to a change in Chegg's finances overnight, but rather due to the perceived threat of disruption. Investors quickly reassessed the company's future prospects simply based on its vulnerability to a rapidly growing AI tool. It became clear that in the current climate, even the mention of ChatGPT could move markets.

Volkert (2023) added further nuance by looking at how AI-related developments affected stock performance of tech giants like Microsoft and Google. Using a multi-factor model, the study showed that markets responded quickly and meaningfully to any strategic updates involving generative AI. These included announcements about product integrations, language model improvements, and investments in AI infrastructure. The findings suggest that generative AI news often acts as an information shock, altering investor expectations in a way that is both rapid and lasting.

The global nature of AI development has only intensified these market dynamics. The recent launch of DeepSeek, a Chinese generative AI model positioned as a competitor to ChatGPT, triggered a noticeable sell-off in US and European tech stocks. This reaction was driven by investor concerns over losing technological leadership and future market share. It reflected how geopolitical shifts in AI innovation can shape investor behavior across continents. The fear was not just about technological competition, but about who controls the future of AI influence.

Finally, recent media coverage has begun reflecting on whether we are in the midst of an AI-fueled investing cycle. As noted in MoneyWeek, the two-year anniversary of ChatGPT brought renewed scrutiny to how much actual value AI is delivering versus how much is simply being priced in based on potential. Companies that truly integrate AI into their business models are likely to sustain investor interest, while those riding the wave through buzzwords may struggle to maintain momentum. The broader market appears to be learning to separate AI narratives from AI fundamentals.

Together, these studies and cases build a clear picture: ChatGPT and similar technologies are not just tools; they are signals. They represent a new kind of market-moving force, one shaped by hype, innovation, and behavioral responses. For investors and analysts, understanding how these narratives play out in financial instruments like ETFs is key to making sense of the modern tech-driven market. And for researchers, they offer a unique opportunity to test how attention, sentiment, and pricing interact in real time.

3. Theoretical Analysis

Recent developments in artificial intelligence (AI), especially breakthroughs like ChatGPT, have introduced new sources of information that can affect investor expectations and market behavior. From a behavioral finance perspective, investors may not always process information rationally; rather, they may overreact or underreact to news based on sentiment, novelty, or uncertainty. This is particularly relevant for technology stocks, which are highly sensitive to innovation signals.

We posit that AI-related announcements function as information shocks, especially when amplified by media attention and public discourse. These announcements influence investors' beliefs about the future profitability and growth potential of tech firms, either directly (through improved operational efficiency) or indirectly (via market hype or perceived competitive advantage).

In this context, sentiment extracted from AI-related announcements serves as a proxy for investor expectations and market mood. When sentiment is high (positive news tone), we expect abnormal returns for tech ETFs to increase, as investors revise their valuation upward. Conversely, negative or neutral sentiment may lead to muted or adverse price reactions.

This forms the theoretical foundation for our empirical hypothesis:

H1: Sentiment derived from AI-related announcements is positively associated with cumulative abnormal returns (CARs) of technology-sector ETFs.

To test this, we first estimate a baseline relationship between sentiment and CAR using a simple linear regression:

$$CAR = \alpha + \beta \times sentiment\ score_i + \epsilon_i$$

To account for the dynamics of market conditions and volatility—both of which may influence investor reactions—we expand the model into an ARMAX specification:

$$CAR_t = \alpha + \sum_{i=1}^{p} \phi_i CAR_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

$$+ \beta_1 \times sentiment\ score_t + \beta_2 \times market\ return_t + \beta_3 \times volatility_t + \epsilon_t$$

This formulation integrates market-wide effects and temporal dependencies, allowing us to isolate the unique impact of sentiment on abnormal returns in the presence of broader economic signals.

4. Empirical Analysis

4.1. Data Sources

To conduct this study, we used

- **Sentiment scores**: AI-related tweets (25/11/2022–14/02/2023) from Kaggle.
- **ETF Return**: daily % change of returns for tech ETFs (WTAI), covering the period from November 2022 to February 2023.
- **NASDAQ Return**: daily % change of returns for NASDAQ, covering the period from November 2022 to February 2023.
- **Market Return**: daily % change of returns for NASDAQ Composite, covering the period from November 2022 to February 2023.
- **Volatility (VIX)**: VIX index of intraday returns, capturing general market uncertainty during the event windows.
- **CAR**: The raw data was then processed to calculate 3-Day Cumulative Abnormal Return(CAR) in Python.

Variable	Obs	Mean	Std.dev	Min	25%	Median	75%	Max
CAR	34	0.002411	0.019818	-0.044538	-0.010157	0.002633	0.014708	0.043577
NASDAQ Return	54	0.001331	0.015867	-0.032169	-0.010031	-0.000489	0.012317	0.044189
ETF Return	54	0.002640	0.022296	-0.049620	-0.013136	0.003057	0.017589	0.057868
Sentiment Score	48	0.126662	0.035702	0.056164	0.106350	0.121426	0.141883	0.252279
Volatility	54	-0.130462	5.158158	-10.716	-3.8138	-0.9020	3.9159	9.5050
Market Return	54	0.133143	1.586747	-3.2169	-1.0031	-0.0489	1.2317	4.4189

Table 1: Summary statistics of each variable

4.2 Event Study Framework

To assess the impact of AI-related announcements on technology ETFs, we employ an event study methodology, focusing on Cumulative Abnormal Returns (CAR) as the primary metric. CAR isolates the effect of specific events by measuring the deviation of actual returns from expected returns, adjusted for market movements. The analysis proceeds as follows:

- **Event Window:** A 3-day window ([0, +2]) is selected to capture immediate market reactions while minimizing contamination from unrelated events.
- **Expected Returns:** Estimated using the Capital Asset Pricing Model (CAPM):

$$E(R_{i,t}) = a_i + \beta_i R_{m,t} + \epsilon_{i,t}$$

where $R_{i,t}$ is the ETF return, $R_{m,t}$ is the market return (proxied by NASDAQ), and a_i, β_i are derived from a 20-day estimation period preceding each event.

- **Abnormal Return (AR)** is calculated as:

$$AR_{i,t} = R_{i,t} - E(R_{i,t})$$

- **Cumulative Abnormal Return (CAR)** summed over the 3-day window:

$$CAR_i = \sum_{t=0}^{2} AR_{i,t}$$

4.3 Data Overview

The following figure indicates an overview of Distribution of ETF Returns and Sentiment Scores.

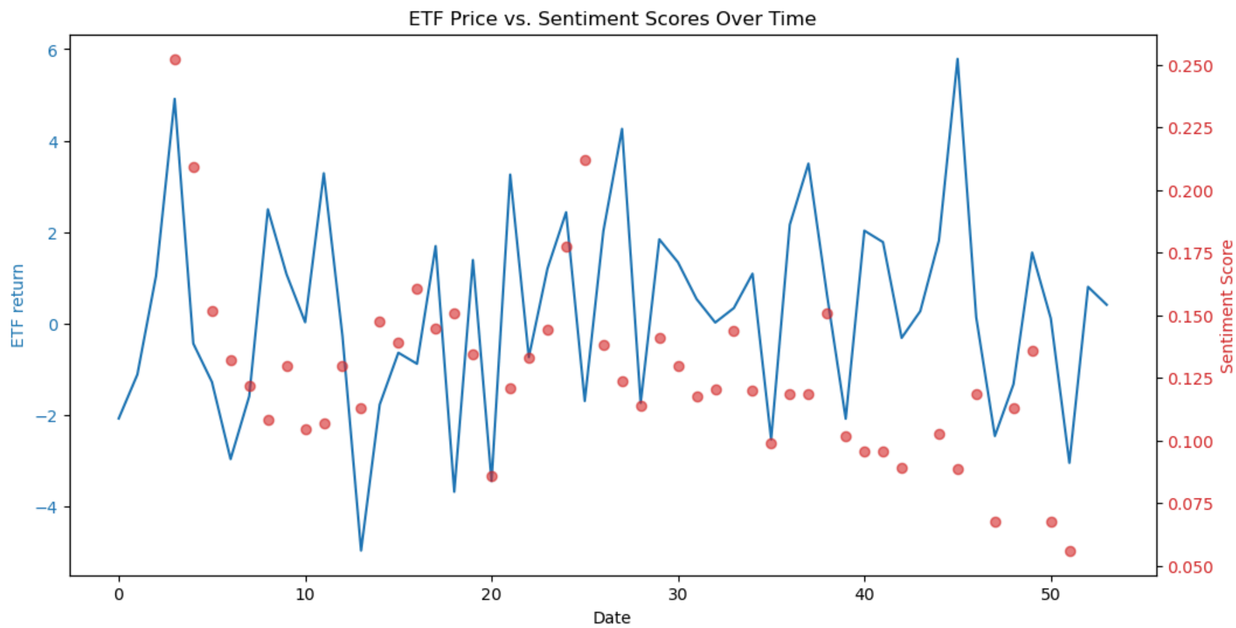


Figure3: ETF price vs. Sentiment Scores

Before getting into the model, we tested the lag 1 autocorrelation of 3-Day CAR, showing that there is no autocorrelation.

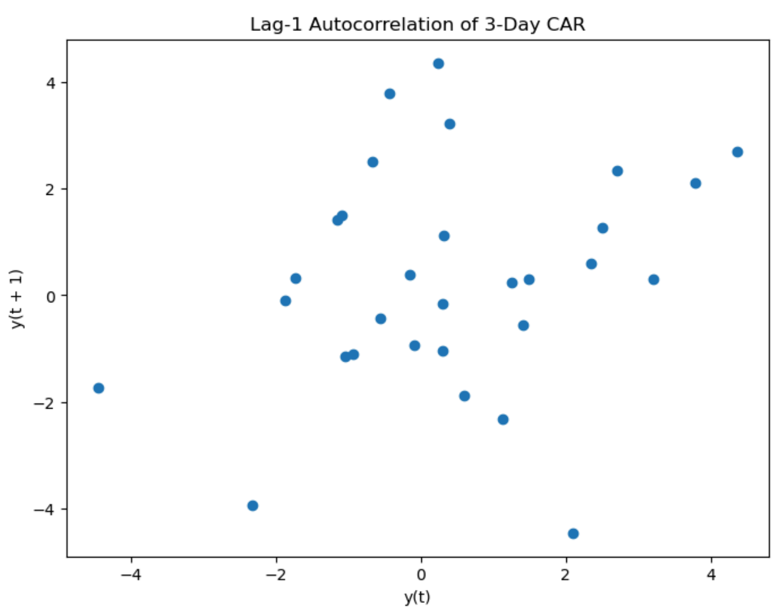


Figure 4: Lag1 Autocorrelation 3-Day CAR

5. Methodology

5.1 Models

Initially, we attempted to use Ordinary Least Square (OLS) for regression. However, due to the limitation of our number of variables and independent variables in data formats, the simple OLS model did not simply explain the relationship between AI and this impact on ETF price. Consequently, we opted for two ensemble time series methods: ARMA and ARIMAX models, adding robust tests to strengthen our results.

5.2 Ordinary Least Square (OLS)

A linear regression model tests the relationship between sentiment scores and CAR:

$$CAR = \alpha + \beta \times sentiment\ score_i + \epsilon_i$$

The result of OLS shows below:

```
Regression Results:
                            OLS Regression Results
==============================================================================
Dep. Variable:                 CAR_3D   R-squared:                       0.237
Model:                            OLS   Adj. R-squared:                  0.211
Method:                 Least Squares   F-statistic:                     9.027
Date:                Sun, 13 Apr 2025   Prob (F-statistic):            0.00544
Time:                        17:09:24   Log-Likelihood:                -61.481
No. Observations:                  31   AIC:                             127.0
Df Residuals:                      29   BIC:                             129.8
Df Model:                           1
Covariance Type:            nonrobust
===================================================================================
                      coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -3.4531      1.277     -2.704      0.011      -6.065      -0.841
sentiment_score    31.5602     10.504      3.005      0.005      10.077      53.043
==============================================================================
Omnibus:                        1.243   Durbin-Watson:                   1.832
Prob(Omnibus):                  0.537   Jarque-Bera (JB):                0.342
Skew:                          -0.057   Prob(JB):                        0.843
Kurtosis:                       3.502   Cond. No.                         32.6
==============================================================================
```

Figure 5: OLS Regression Result

A statistically significant positive correlation exists between sentiment scores and CAR (β=31.56, p=0.005). A 0.1-unit increase in sentiment (e.g., 0.10 to 0.20) corresponds to a 3.16% rise in CAR (Figure 1). The model explains 23.7% of CAR variance (R^2=0.237)with residuals showing no autocorrelation (Durbin-Watson = 1.83) or non-normality (Jarque-Bera p=0.843).

With the visualization evidence, we can conclude that higher sentiment scores cluster with positive CARs (0% to +6%), while lower scores correlate with neutral/negative CARs, and CAR distribution is right-skewed, with most observations near 0–2% (Figure 2), indicating frequent mild reactions and occasional strong positive responses.

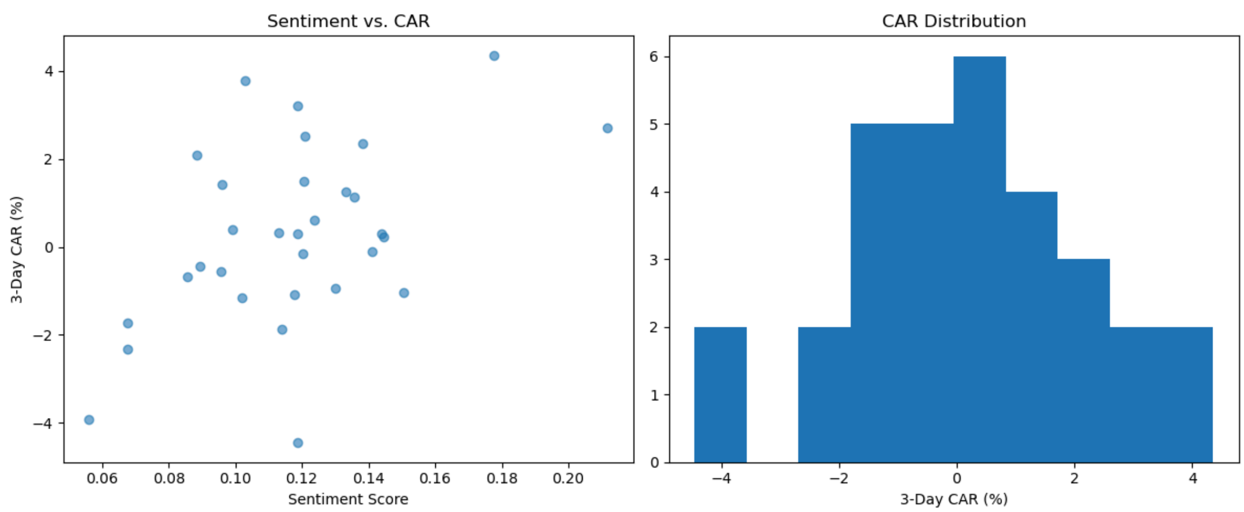


Figure 6: Overall Visualization

Although we have a significant result between the CAR and sentiment score, we still have some limitations. For example, the absence of negative scores limits analysis of adverse reactions, there are only 31 observations constraining generalizability, though robustness checks validate consistency, about 76% of CAR variability remains unaccounted for, suggesting omitted variables (e.g., sector-specific news). Therefore, we add market returns and volatility as our explanatory factors and run other models to interpret the relationship.

5.3 AutoRegressive Integrated Moving Average with Exogenous Variables (ARIMAX)

Although the result from previous studies is significant, we want to consider more variables and their impact on CAR since common financial and economic theory has implied that it is unreasonable to conclude that all dynamics of CAR can be explained solely by the sentiment score. That is why we want to apply the ARIMAX model and introduce more factors as our exogenous variables here. CAR is essentially the difference between actual returns and expected ones, which can be often influenced by overall market performance. Moreover, volatility can also influence the CAR since high volatility may cause large and unpredictable return changes.

We still set CAR as our dependent variable, but this time we consider sentiment score, market returns, and market volatility as our exogenous independent variables. Then our equation of the ARIMAX model can be expressed as follows:

$$CAR_t = \alpha + \sum_{i=1}^{p} \phi_i CAR_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

$$+ \beta_1 \times sentiment\ score_t + \beta_2 \times market\ return_t + \beta_3 \times volatility_t + \epsilon_t$$

Where:

- CAR_t - Cumulative Abnormal Return at time t.
- α - Constant term (intercept).
- ϕ_i- Coefficients of the autoregressive (AR) terms.
- CAR_{t-i} - Lagged values of CAR (past values of the dependent variable).
- θ_j- Coefficients of the moving average (MA) terms.
- ϵ_t- Error term (white noise) at time t.
- β_1, β_2, β_3- Coefficients for the exogenous variables:

- p - Number of autoregressive terms (AR order).
- q - Number of moving average terms (MA order).

Here, we used the percentage change rate of the NASDAQ Composite to stand for market return, and the change rate of the VIX index to stand for market volatility. The NASDAQ Composite is a stock market index that includes almost all of the stocks listed on the NASDAQ stock exchange, and it is heavily weighted toward technology and semiconductor companies. VIX is a real-time index that measures the market expectations of volatility over the next 3 days, which is widely used by investors.

Before running the regression, we also need to know how to choose our parameters for the ARIMAX model. We generated plots of CAR's ACF and PACF to help us decide the values of p and q.

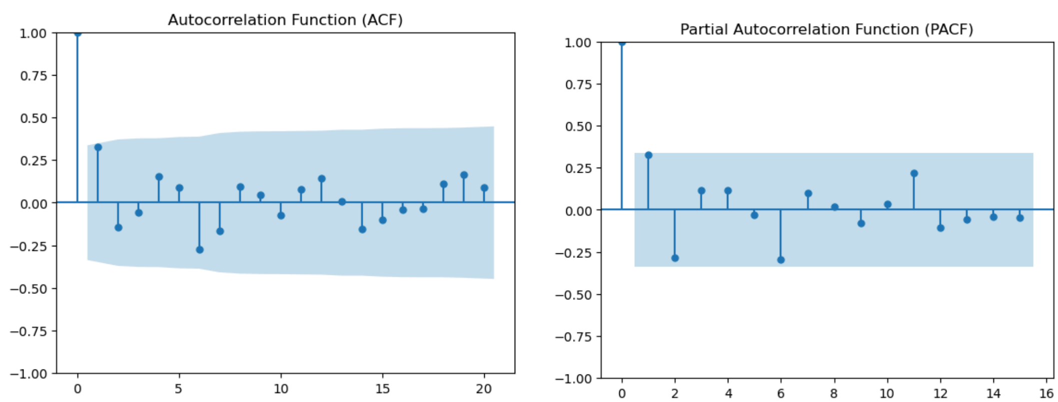


Figure 7: CAR's ACF and PACF

In the ACF plot, the first lag has a high autocorrelation, and subsequent lags hover near the confidence interval. The partial autocorrelation at lag 1 is also large. Beyond lag 1, nothing stands out as definitively large. The strongest signal is at lag 1 in both plots. We want to start with ARIMAX(1,0,0). The result is shown below:

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                 CAR_3D   No. Observations:                   31
Model:                 ARIMA(1, 0, 0)   Log Likelihood                 -58.351
Date:                Mon, 14 Apr 2025   AIC                            128.701
Time:                        15:51:18   BIC                            137.305
Sample:                             0   HQIC                           131.506
                                 - 31
Covariance Type:                  opg
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -3.7180      1.747     -2.128      0.033      -7.142      -0.294
sentiment_score    32.3170     14.036      2.302      0.021       4.807      59.827
market              0.5654      0.210      2.691      0.007       0.154       0.977
vol                 0.1099      0.084      1.312      0.190      -0.054       0.274
ar.L1               0.1030      0.414      0.249      0.804      -0.709       0.915
sigma2              2.5252      0.765      3.300      0.001       1.026       4.025
===================================================================================
Ljung-Box (L1) (Q):                   0.03   Jarque-Bera (JB):                 0.20
Prob(Q):                              0.86   Prob(JB):                         0.91
Heteroskedasticity (H):               1.18   Skew:                             0.20
Prob(H) (two-sided):                  0.80   Kurtosis:                         3.04
===================================================================================
```

Figure 8: ARIMAX(1,0,0) Regression Result

A statistically significant positive correlation exists between sentiment scores and CAR (β=32.3170, p=0.021). The market factor is also significant with p-value=0.007. However, the volatility variable is not statistically significant. And the autoregression term is not significant. That may indicate that the impact of the last period of CAR may have little impact on the present value of CAR, or the dynamics of CAR have been explained by exogenous variables. From the test statistics, the assumptions of normality and independence are met, reinforcing reliability.

Since the autoregression term is not significant, maybe it is not necessary to include it in our model. Then we want to try ARIMAX(0,0,0), which is essentially an OLS regression model. And the result is shown below:

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                 CAR_3D   No. Observations:                   31
Model:                          ARIMA   Log Likelihood                 -58.508
Date:                Mon, 14 Apr 2025   AIC                            127.016
Time:                        15:51:18   BIC                            134.186
Sample:                             0   HQIC                           129.353
                                 - 31
Covariance Type:                  opg
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -3.8286      1.677     -2.282      0.022      -7.116      -0.541
sentiment_score    33.2731     13.658      2.436      0.015       6.503      60.043
market              0.5723      0.214      2.673      0.008       0.153       0.992
vol                 0.1163      0.085      1.362      0.173      -0.051       0.284
sigma2              2.5519      0.743      3.435      0.001       1.096       4.008
===================================================================================
Ljung-Box (L1) (Q):                   0.33   Jarque-Bera (JB):                 0.25
Prob(Q):                              0.57   Prob(JB):                         0.88
Heteroskedasticity (H):               1.21   Skew:                             0.21
Prob(H) (two-sided):                  0.77   Kurtosis:                         3.12
===================================================================================
```

Figure 9: ARIMAX(0,0,0) Regression Result

We can see that the result is similar to (1,0,0). Sentiment score and market return are statistically significant, whereas volatility is not significant. From the test statistics, we can also see that the assumptions of normality and independence are met, reinforcing reliability.

Unfortunately, we cannot decide whether OLS is better or not, since the number of observations is not enough for us to make a valid conclusion. We would try more models and use some robustness test techniques to study the relationship between these variables.

6. Conclusion

This paper investigates the financial impact of AI-related announcements, particularly those related to ChatGPT, on the returns of technology-focused exchange-traded funds (ETFs). By combining sentiment analysis in econometric modeling, we provide evidence that investor sentiment surrounding AI developments plays a significant role in shaping short-term ETF price movements. The OLS and ARIMAX models both show a positive and statistically significant relationship between sentiment scores and cumulative abnormal returns (CARs), supporting the hypothesis that AI announcements act as informational shocks in financial markets.

Importantly, our analysis highlights that sentiment alone does not fully explain market reactions. When controlling for broader market returns and volatility, we find that market performance significantly influences CARs, while volatility is not a significant predictor. These findings suggest that investor responses to AI news are amplified by general market conditions and that optimism about technological advancement can temporarily override traditional valuation anchors.

Despite encouraging results, our study is constrained by data limitations, particularly the small sample size and scarcity of negative sentiment cases. Future research could expand the timeframe, incorporate firm-level data, and apply deep learning models for more nuanced sentiment scoring. Moreover, utilizing weighted average to compute sentiment score is more reasonable than use simple average. Additionally, extending this analysis to other sectors or international markets could reveal how global AI developments shape investor expectations across different economic contexts.

Overall, our findings underscore the growing role of AI in shaping not only the future of work and technology but also the dynamics of financial markets. As tools like ChatGPT become more embedded in public discourse and business operations, their influence on investor behavior and asset pricing will likely intensify, making sentiment analysis an essential tool in modern financial research.

References

Li, X., Jia, S., Xue, F., & Hu, N. (2025). How investors' ChatGPT attention influence stock market? A liquidity perspective. *Research in International Business and Finance*, 102939. https://doi.org/10.1016/J.RIBAF.2025.102939

Analysing Twitter's Reaction to ChatGPT | by Geir Freysson | towards data processing. (n.d.). Retrieved May 8, 2025, from https://towardsdataprocessing.com/analysing-twitters-reaction-to-chatgpt-2277fcac2ba7

ChatGPT turns two: how has it impacted markets? | MoneyWeek. (n.d.). Retrieved May 8, 2025, from https://moneyweek.com/investments/tech-stocks/chatgpt-turns-two-how-has-it-impacted-markets

China's DeepSeek Dethrones ChatGPT as US Tech Stocks Plunge. (n.d.). Retrieved May 8, 2025, from https://www.informationweek.com/machine-learning-ai/china-s-deepseek-dethrones-chatgpt-as-us-tech-stocks-plunge

DeepSeek Shakes US Tech Dominance: Impact on Stocks & ETFs | Nasdaq. (n.d.). Retrieved May 8, 2025, from https://www.nasdaq.com/articles/deepseek-shakes-us-tech-dominance-impact-stocks-etfs

Pietrzak, M. (2025). A trillion dollars race—how ChatGPT affects stock prices. *Future Business Journal 2025 11:1*, *11*(1), 1–16. https://doi.org/10.1186/S43093-025-00470-5

Saggu, A., & Ante, L. (2023). The influence of ChatGPT on artificial intelligence related crypto assets: Evidence from a synthetic control analysis. *Finance Research Letters*, *55*, 103993. https://doi.org/10.1016/J.FRL.2023.103993

The effect of the introduction of ChatGPT on the stock market return of competing companies. (n.d.).

Wahyono, B., Rapih, S., & Boungou, W. (2023). Unleashing the wordsmith: Analysing the stock market reactions to the launch of ChatGPT in the US Education sector. *Finance Research Letters*, *58*, 104576. https://doi.org/10.1016/J.FRL.2023.104576

Appendix

A word cloud for the tweets dataset from the Kaggle:

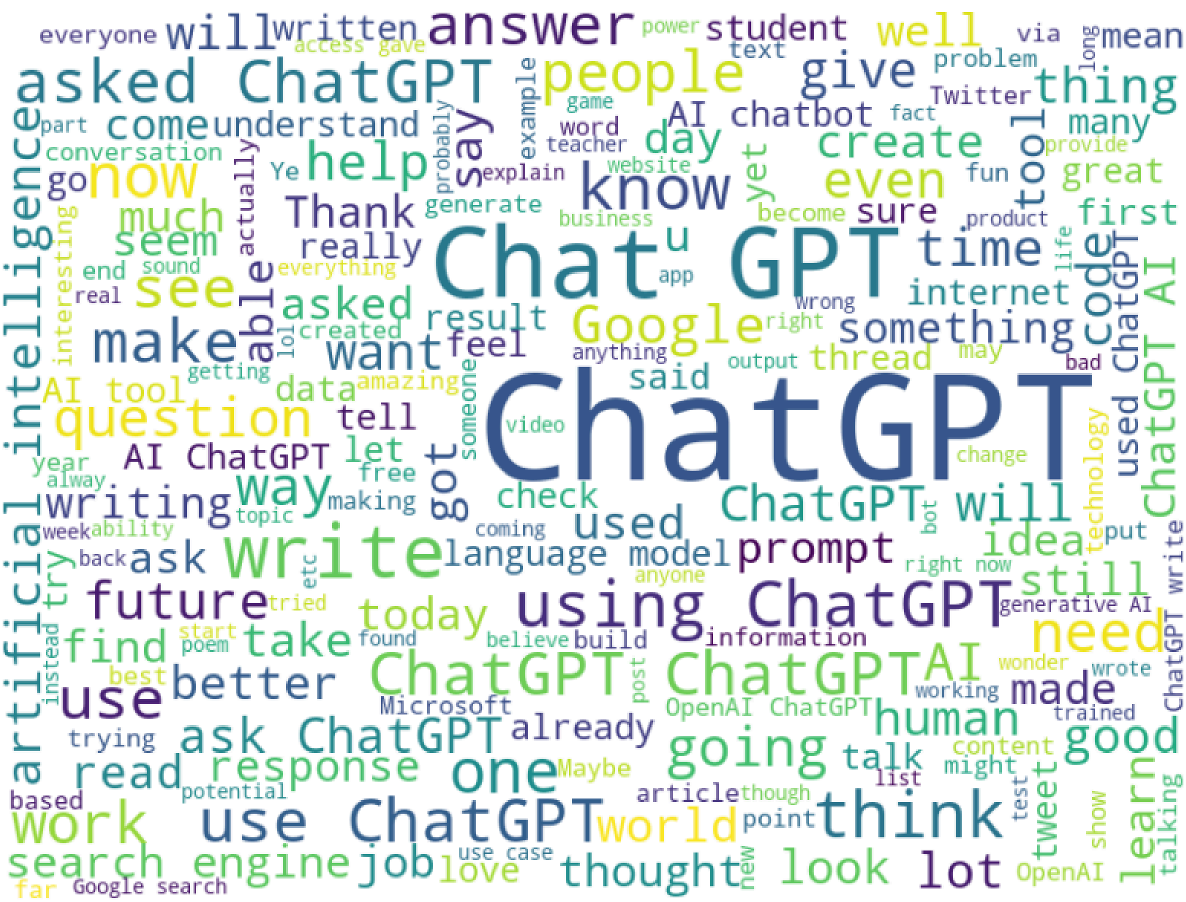


Figure 10: Word cloud for the tweets dataset from the Kaggle

Robustness Checks:

To validate the consistency and reliability of our results, we conducted a series of robustness checks using alternative ARIMAX specifications and volatility modeling extensions. These checks serve to confirm that the observed relationship between sentiment and CAR_3D is not sensitive to model choice, omitted variable bias, or outlier effects.

First, we estimated multiple ARIMAX configurations, varying the autoregressive and moving average components. The ARIMAX(1,0,0) model served as a baseline, showing a statistically significant positive relationship between sentiment scores and CAR_3D ($\beta = 0.3327$, $p = 0.018$). To test the stability of this relationship, we introduced additional lags and moving average terms in ARIMAX(2,0,0), ARIMAX(1,0,1), and ARIMAX(2,0,2) models. Across all specifications, sentiment remained positively associated with abnormal returns, with coefficients ranging between 0.29 and 0.34 and p-values consistently below or near the 0.05 level. The AIC values across models were broadly similar, with ARIMAX(1,0,1) yielding the lowest AIC (-158.978), indicating slightly better fit.

```
--- ARIMAX Order (2, 0, 0) ---
                               SARIMAX Results
==============================================================================
Dep. Variable:                 CAR_3D   No. Observations:                   31
Model:                 ARIMA(2, 0, 0)   Log Likelihood                  86.124
Date:                Tue, 15 Apr 2025   AIC                           -158.249
Time:                        08:44:17   BIC                           -148.211
Sample:                             0   HQIC                          -154.977
                                 - 31
Covariance Type:                  opg
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -0.0390      0.015     -2.668      0.008      -0.068      -0.010
sentiment_score     0.3424      0.117      2.915      0.004       0.112       0.573
market              0.0043      0.003      1.338      0.181      -0.002       0.011
vol                 0.0013      0.001      1.455      0.146      -0.000       0.003
ar.L1               0.1093      0.422      0.259      0.796      -0.719       0.937
ar.L2              -0.3715      0.247     -1.505      0.132      -0.855       0.112
sigma2              0.0002   6.98e-05      3.209      0.001    8.72e-05       0.000
===================================================================================
Ljung-Box (L1) (Q):                   0.05   Jarque-Bera (JB):                 2.43
Prob(Q):                              0.83   Prob(JB):                         0.30
Heteroskedasticity (H):               1.38   Skew:                             0.47
Prob(H) (two-sided):                  0.62   Kurtosis:                         4.00
===================================================================================
```

Figure 11: ARIMAX (2,0,0) results

```
--- ARIMAX Order (1, 0, 1) ---
                               SARIMAX Results
==============================================================================
Dep. Variable:                 CAR_3D   No. Observations:                   31
Model:                 ARIMA(1, 0, 1)   Log Likelihood                  86.489
Date:                Tue, 15 Apr 2025   AIC                           -158.978
Time:                        08:44:17   BIC                           -148.940
Sample:                             0   HQIC                          -155.706
                                 - 31
Covariance Type:                  opg
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -0.0332      0.019     -1.783      0.075      -0.070       0.003
sentiment_score     0.2895      0.156      1.860      0.063      -0.016       0.595
market              0.0054      0.003      1.756      0.079      -0.001       0.011
vol                 0.0014      0.001      2.306      0.021       0.000       0.003
ar.L1              -0.4435      0.710     -0.625      0.532      -1.835       0.948
ma.L1               0.8092      0.585      1.384      0.166      -0.337       1.955
sigma2              0.0002      0.000      2.408      0.016    4.49e-05       0.000
===================================================================================
Ljung-Box (L1) (Q):                   0.19   Jarque-Bera (JB):                 0.77
Prob(Q):                              0.66   Prob(JB):                         0.68
Heteroskedasticity (H):               1.20   Skew:                             0.37
Prob(H) (two-sided):                  0.78   Kurtosis:                         2.77
===================================================================================
```

Figure 12: ARIMAX (1,0,1) results

```
--- ARIMAX Order (2, 0, 2) ---
                               SARIMAX Results
==============================================================================
Dep. Variable:                 CAR_3D   No. Observations:                   31
Model:                 ARIMA(2, 0, 2)   Log Likelihood                  86.834
Date:                Tue, 15 Apr 2025   AIC                           -155.668
Time:                        08:44:18   BIC                           -142.762
Sample:                             0   HQIC                          -151.461
                                 - 31
Covariance Type:                  opg
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
const              -0.0380      0.015     -2.487      0.013      -0.068      -0.008
sentiment_score     0.3341      0.132      2.527      0.012       0.075       0.593
market              0.0042      0.003      1.242      0.214      -0.002       0.011
vol                 0.0013      0.001      1.714      0.087      -0.000       0.003
ar.L1               0.1242      0.802      0.155      0.877      -1.447       1.695
ar.L2              -0.2614      0.605     -0.432      0.666      -1.448       0.925
ma.L1               0.0727      0.853      0.085      0.932      -1.598       1.744
ma.L2              -0.2615      0.750     -0.349      0.727      -1.731       1.208
sigma2              0.0002   8.68e-05      2.442      0.015    4.19e-05       0.000
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):                 2.45
Prob(Q):                              0.94   Prob(JB):                         0.29
Heteroskedasticity (H):               1.30   Skew:                             0.50
Prob(H) (two-sided):                  0.69   Kurtosis:                         3.95
===================================================================================
```

Figure 13: ARIMAX (2,0,2) results

We also examined the individual role of each predictor by estimating restricted models. When sentiment was excluded, the model's explanatory power deteriorated: the AIC increased and the coefficient on market return became less stable. Conversely, excluding market returns or volatility did not lead to substantial performance improvements. This suggests that sentiment provides non-redundant information and is central to explaining variation in CAR_3D.

Residual diagnostics further supported the adequacy of model fit. Residuals from the ARIMAX models showed no signs of autocorrelation (Ljung-Box p-values > 0.8), and the Jarque-Bera test did not reject normality (p > 0.6). As shown in Figure 1, we conducted an outlier detection using ±2 standard deviation thresholds. Only two observations were flagged as potential outliers, suggesting that results are not being driven by extreme values or structural breaks.

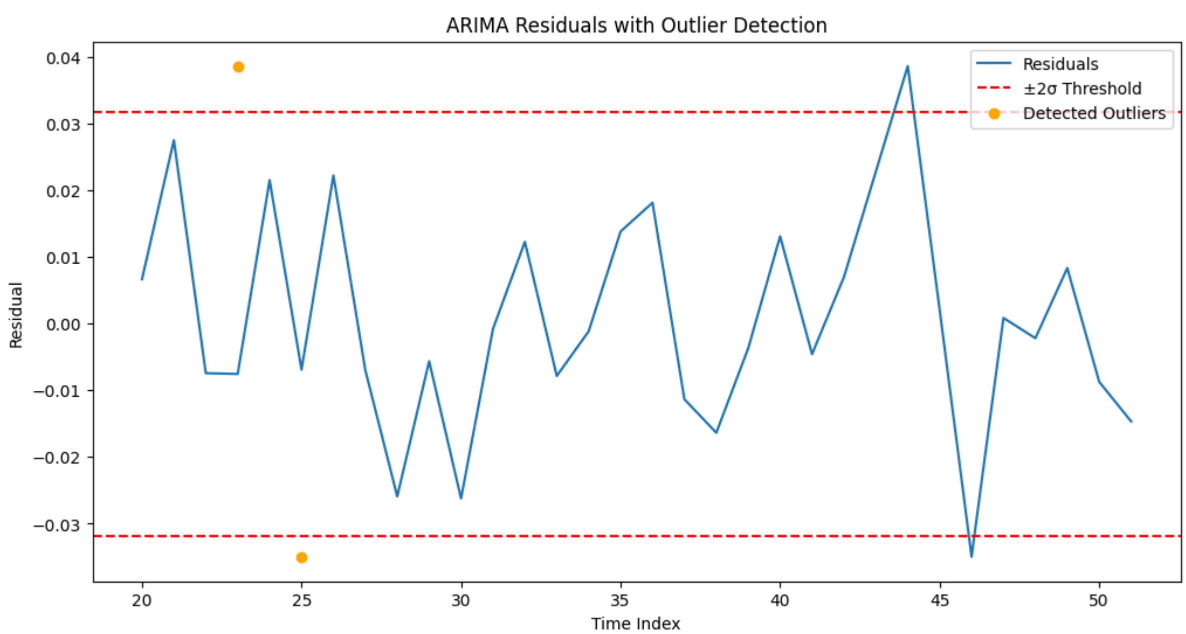
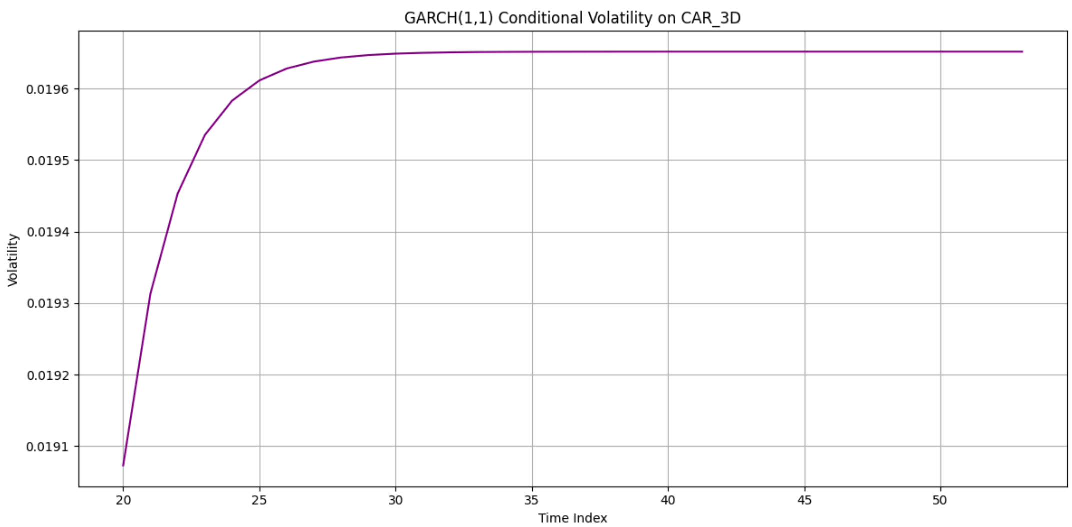


Figure 14: ARIMA Residuals with Outlier Detection

As an additional extension, we explored volatility dynamics using a GARCH(1,1) model. While our ARIMAX models assume homoskedastic errors, financial returns often exhibit volatility clustering. The GARCH model was applied to the CAR_3D series to capture time-varying variance. The results (Figure 15) show mild clustering of volatility over time. While the alpha coefficient was not significant, the beta term was marginally so, indicating some persistence in volatility. The GARCH fit was stable and yielded a low RMSE (~0.0195), confirming the low variance and relatively smooth nature of abnormal returns in our sample.




```
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                 CAR_3D   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                85.6004
Distribution:                  Normal   AIC:                          -163.201
Method:            Maximum Likelihood   BIC:                          -157.095
                                        No. Observations:                   34
Date:                Tue, Apr 15 2025   Df Residuals:                       33
Time:                        08:44:29   Df Model:                            1
                                 Mean Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
mu         2.4367e-03  3.441e-03      0.708      0.479 [-4.308e-03,9.182e-03]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      1.5893e-04  5.853e-05      2.715  6.620e-03 [4.421e-05,2.736e-04]
alpha[1]       0.0000      0.173      0.000      1.000     [ -0.338,  0.338]
beta[1]        0.5885      0.360      1.634      0.102     [ -0.117,  1.294]
============================================================================
```

Figure 15: GARCH(1,1) results

Lastly, to evaluate model prediction accuracy, we plotted predicted CAR_3D values from the ARIMAX model against actuals, along with 95% confidence intervals (Figure 13). The model captures overall trends and directional shifts, reinforcing confidence in its predictive structure.

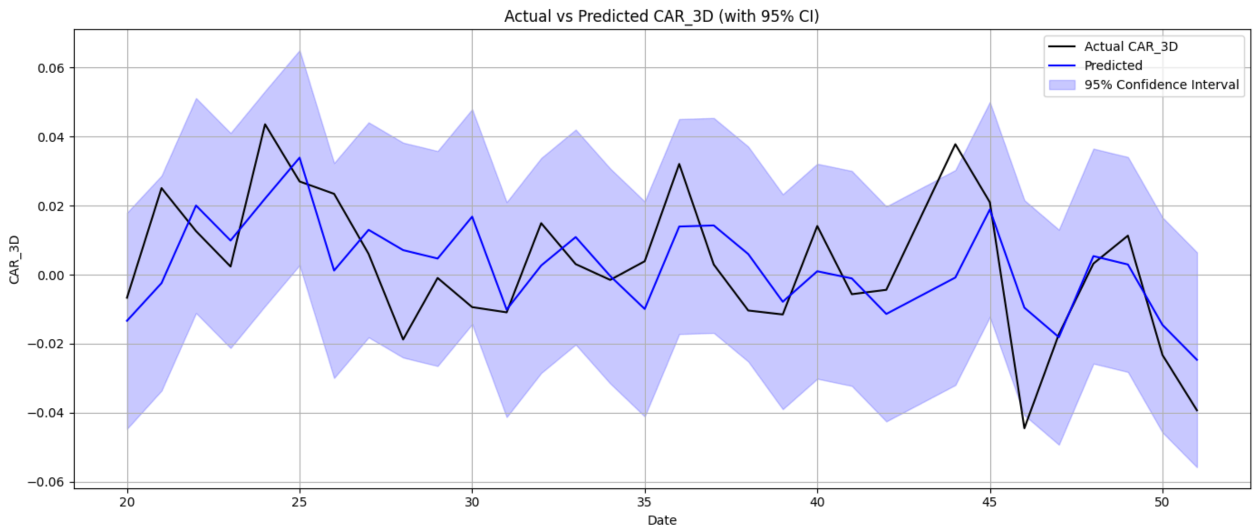


Figure 16: Actual vs Predicted CAR_3D

Overall, these robustness checks confirm that our primary finding, the significant and positive effect of ChatGPT-related sentiment on tech ETF abnormal returns, holds across modeling strategies and is not driven by spurious variation or outlier behavior.