The Power of Media:
Sentiment Shifts and Stock Price Dynamics*

Sentiment predicts price movements for Netflix, but not Amazon or Meta.

Cristina Su Lam   Parisa Pham   Sum Yee Chan

ECO475H1S: Applied Econometrics II

Department of Economics, University of Toronto

Prof. Antoine Djogbenou

April 14, 2025

Abstract. This study investigates whether financial news sentiment affects daily stock price changes and price direction for three large-cap technology firms: Netflix, Amazon, and Meta. Using *New York Times* headlines from 2014–2025, we construct sentiment scores with a finance-specific dictionary and link them to stock prices, trading volume, macroeconomic indicators, and market trends. Employing OLS, fixed effects, IV, and logistic models, we find that sentiment significantly predicts price movements for Netflix but has no meaningful impact on Amazon or Meta. For Netflix, IV estimates remain directionally consistent but imprecise, while logistic models show weak predictive power across all firms. These results highlight that sentiment effects are firm-specific and most pronounced in volatile, media-sensitive environments, suggesting that forecasting models should account for structural differences rather than apply a one-size-fits-all approach.

Contents

1 Introduction

Financial economists have long treated asset prices as functions of hard information such as earnings, cash-flow expectations, and macroeconomic conditions (Fama, 1990). Recent research in behavioural finance and text-as-data methods, however, shows that qualitative signals embedded in news coverage can also move markets in the short run. Pioneering work by Tetlock (2007) and Loughran and McDonald (2011) documents that the tone of newspaper language predicts subsequent returns and trading volume, suggesting that investors react not only to fundamentals but also to the narratives that frame them.

This study investigates whether the sentiment conveyed by daily financial headlines measurably influences the stock price movements of large technology firms. Answering this question has both theoretical and practical implications: if sentiment effects differ across firms, pooled models may misrepresent their importance, and trading strategies that ignore firm characteristics may systematically misprice risk.

We focus on three members of the FAANG cohort—Amazon, Meta, and Netflix. Technology stocks attract intense coverage and retail participation, making them a natural laboratory for sentiment-driven trading (Cristescu et al., 2023). Apple and Google (Alphabet) are excluded because their common names create substantial lexical ambiguity for headline scraping, whereas "Amazon," "Meta," and "Netflix" map cleanly onto single tickers. Restricting attention to these firms reduces classification error while still capturing a wide range of business models and investor bases within the sector.

Our empirical design combines daily *New York Times* headlines from 2014–2025 with adjusted-close prices from Yahoo Finance. Sentiment scores are computed using the domain-specific dictionary of Loughran and McDonald (2011). We estimate a sequence of models: baseline OLS, specifications with macroeconomic and market controls, firm and calendar fixed effects, an instrumental-variables model that uses one-day-lagged sentiment to mitigate simultaneity bias (Tetlock, 2007), and a complementary logistic regression that predicts the sign of price movements.

The results reveal pronounced firm-level heterogeneity. For Netflix, a one-unit increase in

the sentiment score is associated with an \$0.85 to \$1.43 rise in next-day price depending on the specification, and the effect remains positive (though not significant) in the IV model. For Amazon and Meta, sentiment coefficients are small and statistically indistinguishable from zero once market-wide factors are controlled. These findings imply that sentiment matters most for firms with high retail ownership and narrative-driven volatility, while diversified giants anchored by institutional investors respond mainly to systematic market factors.

By isolating firm-specific sentiment effects using causal econometric tools, this study contributes to the existing literature in three ways. It shows that media tone is not a uniform predictor even within a single sector, illustrates how lagged sentiment can help address simultaneity concerns in high-frequency settings, and offers a replicable firm-level panel linking news sentiment to stock performance. While the analysis is limited by headline-based sentiment measures and assumptions about instrument validity, the findings provide new insight into how structural characteristics shape the role of sentiment in asset pricing.

The remainder of the paper is structured as follows. Section 1.1 reviews the relevant literature. Section 2 outlines the data sources and variable construction. Section 3 presents the empirical framework, Section 4 reports the findings, Section 5 offers interpretation and limitations, and Section 6 concludes.

1.1   Literature Review

Over the past two decades, the study of asset–pricing dynamics has moved beyond traditional balance-sheet fundamentals to consider how qualitative signals shape investor behaviour. Growing evidence shows that the tone, framing, and timing of financial news can influence trading activity and short-term returns, prompting a large empirical literature that blends insights from behavioural finance, natural-language processing, and applied econometrics.

Early contributions establish the basic link between media tone and market movements. Tetlock (2007) construct a content-based index from the *Wall Street Journal* and find that days dominated by negative language are followed by lower market returns and higher trading volume, patterns that

4

reverse within a week. The study interprets this short-lived drift as evidence that investors initially overreact to pessimistic coverage. Extending the analysis, Dougal et al. (2012) exploit variation in individual columnist style at the same newspaper. After controlling for firm fundamentals, they show that a more pessimistic writer induces additional same-day price declines, providing quasi-experimental support for a causal effect of sentiment.

Subsequent work broadens the empirical setting and refines measurement. Firm-level studies in the technology sector, where media exposure and retail participation are pronounced, reveal particularly strong effects (Cristescu et al., 2023). Using wavelet correlation, Cristescu et al. (2023) report that news titles influence Tesla's opening prices, whereas article descriptions matter more for Microsoft, highlighting that sentiment effects vary across both companies and information windows. At the index level, Fazlija and Harder (2022) combine a BERT sentiment model with a random-forest classifier to forecast the daily direction of the S&P500, achieving substantial gains in predictive accuracy. Parallel research shows that measurement choices are critical: dictionaries tailored to financial language outperform generic word lists (Xiao and Ihnaini, 2023; Loughran and McDonald, 2011), and sentiment derived from *New York Times* headlines can affect global equity indices (García-Medina et al., 2018).

Two broad methodological approaches have emerged in response. Machine-learning studies focus on forecasting and often employ algorithms such as random forests, long short-term memory networks, and BERT variants (Patil et al., 2024; Xiao and Ihnaini, 2023). While these models deliver strong out-of-sample performance, they seldom incorporate design features that address endogeneity or firm-specific heterogeneity. Recent work has begun to explore hybrid approaches that blend predictive accuracy with causal inference, but such methods remain relatively underutilized in financial sentiment research. By comparison, econometric studies rely on ordinary least squares (OLS), fixed effects, instrumental variables (IV), and binary response models such as logistic regression to estimate causal effects or directional return probabilities. Notable approaches include the use of newspaper strikes as an instrument for media coverage (Engelberg and Parsons, 2011), lagged sentiment to capture short-term reversals (Tetlock, 2007), and logistic regression

to model the likelihood of positive stock returns based on sentiment extracted from social media (Chen et al., 2014). Nevertheless, even this stream of research rarely compares sentiment effects across firms with differing investor bases, volatility profiles, or media visibility.

While the literature has grown increasingly sophisticated, three key limitations persist in how sentiment effects are modeled and compared. First, much of the predictive literature treats sentiment as an exogenous input, leaving open the possibility that news tone reacts to price changes rather than causes them. Second, few studies test whether the magnitude or persistence of sentiment effects varies systematically across firms. Third, IV and panel-fixed-effects strategies—which can mitigate simultaneity bias and unobserved heterogeneity—are still underused.

To address these gaps, the present study links *New York Times* headline sentiment to stock returns for Amazon, Meta, and Netflix, three technology firms that differ in investor composition and media visibility (Cristescu et al., 2023). Sentiment scores are computed with the Loughran and McDonald (2011) dictionary, chosen for its domain relevance. The empirical strategy proceeds in three stages. Baseline OLS models establish the unconditional association between sentiment and next-day returns. Specifications with firm and calendar-date fixed effects then control for unobserved heterogeneity and common shocks. Finally, a two-stage least squares model instruments current sentiment with its one-day lag, following Tetlock (2007), to reduce simultaneity bias. A complementary logistic regression tests whether sentiment predicts the sign of price changes. By applying causal econometric tools to comparable firms within a single sector, the study provides new evidence on how structural characteristics condition the impact of media tone on asset prices.

2   Data

This section describes the datasets employed in the empirical analysis and summarises the construction and justification of the key variables. Section 2.1 details the data sources, coverage, and panel construction. Section 2.2 explains the theoretical rationale behind variable selection. Section 2.3 presents summary statistics and exploratory visualizations.

2.1 Overview

This study constructs a firm-level daily panel dataset to examine how financial news sentiment influences short-term stock price movements, controlling for macroeconomic and market-wide conditions. The panel spans from 2 January 2014 to 1 April 2025 and covers three major technology firms: Amazon, Netflix, and Meta. The unit of observation is a firm–day, where each row corresponds to one company on a given trading day. The final dataset comprises approximately 2,943 firm–day observations.

All data processing and integration steps were conducted in R (R Core Team, 2022). The sentiment component of the dataset is derived from the *New York Times Articles Metadata* collection available on Kaggle (Singh, 2025). To isolate company-relevant articles, regular expression keyword filters were applied to the `headline`, `lead_paragraph`, and `snippet` fields. The resulting text was tokenized using the `tidytext` package (Silge and Robinson, 2016), and sentiment was computed using the financial sentiment dictionary developed by Loughran and McDonald (2011). For each firm i on day t, the sentiment score is defined as:

$$S_{it} = \frac{\text{pos}_{it} - \text{neg}_{it}}{\text{pos}_{it} + \text{neg}_{it}},$$

where pos_{it} and neg_{it} denote the number of positive and negative tokens, respectively. If multiple articles mention a firm on the same day, all tokens from those articles are pooled, and the sentiment score reflects the total count of positive and negative words for that day. In the absence of any articles for a given firm–day, the sentiment score is set to $S_{it} = 0$, treating the lack of news as neutral tone. A one-day lag of S_{it} is also constructed for use in instrumental variable models.

Stock market data, including adjusted closing prices and trading volumes, were obtained from Yahoo Finance via the `quantmod` package (Ryan, 2020). The primary dependent variable is the daily return, calculated as the first difference of adjusted close prices:

$$\Delta p_{it} = p_{it} - p_{it-1}.$$

A binary variable is additionally constructed, taking the value 1 if the return is positive on day t, and 0 otherwise. To capture broader market movements, the daily change in the NASDAQ Composite Index is included, retrieved from Yahoo Finance using the same method as firm-level stock data.

To control for macroeconomic conditions, three indicators from the Federal Reserve Economic Data (FRED) are incorporated: the Effective Federal Funds Rate (EFFR) (Federal Reserve Bank of New York, 2025), the Consumer Price Index (CPI) (U.S. Bureau of Labor Statistics, 2025a), and the Unemployment Rate (UNRATE) (U.S. Bureau of Labor Statistics, 2025b). As CPI and UNRATE are reported monthly, their values are forward-filled to a daily frequency using the zoo package (Zeileis and Grothendieck, 2005) to align with the higher-frequency stock and sentiment data.

Prior to merging, all text fields were standardized by lowercasing and trimming whitespace. Datasets were then joined by calendar date. Observations with missing stock prices—primarily non-trading days—were excluded. Categorical variables for calendar month and year were generated to accommodate time fixed effects in the regressions. The resulting integrated panel includes sentiment scores, stock prices, trading volume, NASDAQ movements, and macroeconomic variables for each firm–day. This structure enables a rigorous econometric assessment of how textual sentiment influences short-run price dynamics, conditional on firm-level and systemic factors. A sample of the cleaned dataset structure is shown in Table 1.

2.2 Variable Selection

The variables included in this study are selected based on established empirical findings and theoretical relevance to short-term asset pricing. The primary independent variable is a firm-specific daily sentiment score (S_{it}), computed from *New York Times* headlines using the financial-domain dictionary by Loughran and McDonald (2011), which has been shown to outperform general-purpose lexicons in financial contexts (Xiao and Ihnaini, 2023). This follows a growing literature linking media tone to stock price fluctuations (Tetlock, 2007; Cristescu et al., 2023).

8

The dependent variable is daily stock price change (Δp_{it}), which captures short-run price sensitivity to sentiment shocks. To assess directionality, a binary variable indicating positive or negative return is also constructed.

Macroeconomic controls—including interest rates (proxied by the Effective Federal Funds Rate), inflation (measured by the Consumer Price Index), and the unemployment rate—are included to capture broader economic conditions that influence investor expectations and equity valuation. These are standard controls in studies of return variation (e.g., Fama 1990).

NASDAQ daily return serves as a market-wide control for systematic risk, particularly relevant for large-cap technology firms whose returns are tightly correlated with overall index trends. Firm-level trading volume (V_{it}) is included to represent liquidity and investor attention, both of which may amplify or moderate the influence of sentiment.

Finally, lagged sentiment ($S_{i,t-1}$) is used as an instrument in 2SLS models to address potential simultaneity bias—an approach grounded in the methodology of Tetlock (2007) and consistent with recent applications of instrumental variables in financial sentiment analysis.

This variable set enables a flexible yet interpretable framework for estimating sentiment effects on stock price dynamics while mitigating omitted variable bias and endogeneity concerns.

2.3 Descriptive Statistics

To motivate the regression analysis, we first summarize the distributional characteristics of the key variables. Tables 2, 3, and 4 present summary statistics for the main variables across the three stocks of interest. The dependent variable, daily stock price change, exhibits significant variability across firms. Meta and Amazon display wide distributions, though with somewhat less extreme values. Netflix, in particular, shows the highest volatility, with daily changes ranging from –122.42 to 84.57 and a notably wide spread between quartiles, indicating that its returns are more sensitive to shocks or abrupt investor reactions than those of its peers.

Among the key independent variables, the sentiment score ranges from –1 to 1. All three firms exhibit distributions skewed slightly negative, with median values below zero. On average, Netflix

has a sentiment score of –0.17, Amazon –0.25, and Meta –0.39. This indicates a persistent tilt toward negatively toned coverage, particularly for Meta, potentially reflecting reputational controversies or regulatory scrutiny during the sample period. Such skewness aligns with patterns in financial journalism, which tends to focus on risk and volatility more than positive developments—especially for high-profile firms under public or governmental pressure.

Macroeconomic indicators such as interest rate, inflation, and unemployment rate serve as general controls for economic conditions. These variables exhibit relatively modest variation over the sample period and are identical across firms on a given day, as they do not vary at the firm level. While interest rate fluctuates daily, inflation and unemployment rates are released monthly and are forward-filled to align with the daily panel structure. The interest rate ranges from 0.04% to 5.33%, inflation from \$234.7 to \$317.6, and unemployment rate from 3.4% to 14.8%. Including these variables helps account for changes in the overall economy, such as rising interest rates or inflation, that could influence stock prices across all firms. By controlling for these factors, we can better isolate the specific effect of news sentiment on stock returns without mistakenly attributing economy-wide movements to firm-specific news.

The NASDAQ daily change is included as a market-wide control to account for systemic factors that influence all stocks on a given day. It exhibits substantial variability, ranging from below –700 to above 670, reflecting broad market fluctuations such as the COVID-19 crash and recovery. At the firm level, trading volume is incorporated to capture shifts in investor attention and liquidity. Amazon averages the highest daily volume (approximately 86 million shares), followed by Meta (26 million) and Netflix (9 million), with all three displaying right-skewed distributions. These spikes in trading activity often coincide with firm-specific events such as earnings releases or major announcements and may be associated with larger price movements. Together, these controls help disentangle the effect of sentiment from both market-wide dynamics and firm-specific trading behavior. The figures that follow provide a visual overview of price trends and the relationship between sentiment and returns, further illustrating firm-level differences.

Figure 1 displays adjusted closing prices over time. Netflix exhibits the steepest growth and

most pronounced volatility, particularly between 2020 and 2022. Meta shows a more stable price path, though it experiences a visible correction post-2022. Amazon's trajectory is smoother and more gradual across the entire period. These trends reflect differing exposures to sector-specific disruptions, such as streaming demand shocks, regulatory shifts, and global tech cycles.

Figure 2 presents a scatter plot of sentiment scores against daily stock price change. While the relationship is noisy, a few patterns emerge. First, the density of observations near zero on both axes suggests that most news headlines carry weak sentiment and coincide with minimal price movement, which is consistent with market efficiency or sentiment neutrality. However, the plot also shows greater variance in returns when sentiment scores approach ±1, particularly for Netflix, where extreme sentiment scores often coincide with sharp price changes. For Meta and Amazon, this pattern is less pronounced, which may indicate that sentiment is a weaker signal for these firms or that their stock returns are more driven by fundamentals or broader market factors. This uneven pattern suggests that investors respond differently to news depending on the firm, which may affect how we model the impact of sentiment on stock returns.

Overall, these descriptive insights highlight meaningful differences across firms in both sentiment exposure and stock return behavior. The evidence suggests that sentiment may play a role in influencing prices, particularly for firms like Netflix that exhibit high return dispersion. These findings motivate a more formal econometric investigation to assess whether sentiment effects persist after accounting for macroeconomic and market-wide influences. Accordingly, the next section develops regression models to examine these relationships in greater detail.

3 Econometric Model

Building on the patterns observed in the previous section, the following analysis outlines the functional forms, estimation techniques, and underlying assumptions used to assess the impact of news sentiment on stock price movements. The empirical strategy incrementally builds from a simple baseline model to more comprehensive specifications that incorporate macroeconomic and

financial controls, time fixed effects, and instrumental variables to address potential endogeneity.

3.1 Functional Forms

Let Y_{it} represent the daily change in adjusted closing price for firm i on day t. The key explanatory variable is the sentiment score S_{it} derived from financial news.

The analysis begins with the following baseline linear model:

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \epsilon_{it} \tag{1}$$

To control for macroeconomic conditions, the model is extended as follows:

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 R_t + \beta_3 U_t + \beta_4 \pi_t + \epsilon_{it} \tag{2}$$

where R_t denotes the interest rate, U_t the unemployment rate, and π_t the inflation level.

Market-wide and firm-level financial factors are then introduced:

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 R_t + \beta_3 U_t + \beta_4 \pi_t + \beta_5 N_t + \beta_6 V_{it} + \epsilon_{it} \tag{3}$$

where N_t represents the NASDAQ daily change and V_{it} the trading volume.

The following equation represents the full model, including all controls and time fixed effects:

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 R_t + \beta_3 U_t + \beta_4 \pi_t + \beta_5 N_t + \beta_6 V_{it} + \gamma_m + \delta_y + \epsilon_{it} \tag{4}$$

with γ_m and δ_y denoting month and year fixed effects, respectively.

Finally, to address potential endogeneity of S_{it}, a two-stage least squares (2SLS) specification

is introduced:

$$\text{1st stage:} \quad S_{it} = \alpha_0 + \alpha_1 S_{i,t-1} + \alpha_2 X_{it} + u_{it} \tag{5}$$

$$\text{2nd stage:} \quad Y_{it} = \beta_0 + \beta_1 \hat{S}_{it} + \beta_2 X_{it} + \epsilon_{it} \tag{6}$$

where X_{it} includes all controls used in the full specification.

3.2 Estimation Methods

The baseline and extended models are estimated using Ordinary Least Squares (OLS). Given the known presence of heteroskedasticity in financial time series, all standard errors are corrected using the robust variance estimator:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'Y$$

Given the possibility of endogeneity in the sentiment variable, the instrumental variables (IV) approach uses lagged sentiment as an instrument. Inflation is excluded from Equation 6 to ensure that the model remains identified and computationally stable. In this setup, current sentiment is the only endogenous variable, and lagged sentiment is the only available instrument that satisfies the identification condition of 2SLS. Although other control variables—such as interest rate, unemployment, NASDAQ return, and trading volume—are included in the model, they are treated as exogenous and do not pose a problem for estimation.

Inflation, on the other hand, causes the IV regression to fail despite not being endogenous. This failure is not due to multicollinearity, as confirmed by variance inflation factor (VIF) values below 5 in the first-stage regression (Table 5), but is likely caused by limited variation in the inflation variable and overlap with time fixed effects. Including it may introduce near-linear dependence among covariates, making the model matrix singular and the estimation infeasible.

To examine the direction rather than the magnitude of price changes, a logistic regression is

13

estimated using the same covariates as in Equation 4.

$$\Pr(D_{it} = 1) = \frac{\exp(\beta_0 + \beta_1 S_{it} + \ldots)}{1 + \exp(\beta_0 + \beta_1 S_{it} + \ldots)} \tag{7}$$

where D_{it} equals 1 if the daily price change is positive on day t and 0 otherwise.

3.3 Assumptions and Justification

The linear regression models rest on several key assumptions. First, the relationship between predictors and the outcome is assumed to be linear and additive, implying that the effect of each variable on returns is constant and independent of the others. This assumption is reasonable given the short daily frequency of the data and the focus on capturing marginal effects rather than non-linear dynamics.

Second, OLS estimation assumes strict exogeneity, meaning that the regressors must be uncorrelated with the error term. This assumption may be violated if sentiment is influenced by the same-day stock return it is intended to explain. For example, a sharp drop in a stock's price early in the trading day could trigger the publication of more negative headlines by market close. In this case, sentiment and returns are jointly determined, leading to simultaneity bias and inconsistency in the OLS estimates.

To address this concern, a 2SLS approach is used with lagged sentiment as an instrument for current sentiment. This instrument satisfies the relevance condition because sentiment tends to be autocorrelated — the tone of media coverage often carries over from one day to the next. It also plausibly satisfies the exclusion restriction, as sentiment from the previous day is unlikely to be influenced by unobservable return shocks occurring today, especially after conditioning on market and macroeconomic controls. This setup helps isolate the component of sentiment that is not simultaneously reacting to current price movements, allowing for a more credible estimate of its causal effect on returns.

Third, heteroskedasticity is a common feature of financial return data, as volatility varies across

14

firms and over time. Robust standard errors are used in all linear models to ensure valid statistical inference in the presence of non-constant error variance.

Fourth, the independence assumption requires that errors are not correlated across time. This may not fully hold in high-frequency financial data due to autocorrelation or unobserved firm-level shocks. However, including lagged instruments, time fixed effects, and market controls reduces the risk of omitted variable bias and partially addresses serial correlation.

Lastly, the logistic regression model assumes a linear relationship between the covariates and the log-odds of a binary outcome. This allows sentiment and other predictors to influence the probability of a price increase in a flexible but interpretable way. While it does not model return magnitudes, this binary approach provides complementary insights into the directional impact of sentiment, particularly when the magnitude of returns is difficult to interpret or subject to high variance.

3.4   Firm Level Implementation

The baseline, full, and logistic regression models are estimated for all three firms. The instrumental variables model is implemented only for Netflix, as regression results indicate that the sentiment score is statistically significant for this firm. This raises greater concern about potential endogeneity, making the use of an IV approach appropriate in this case. For Amazon and Meta, where sentiment does not appear to meaningfully explain return variation, the IV model is excluded to maintain parsimony and avoid unnecessary complexity.

Overall, this modeling framework allows for a detailed assessment of how sentiment influences stock returns, accounting for both magnitude and direction, while addressing firm-level variation and potential confounding factors.

4 Results

This section presents the empirical findings on the relationship between financial news sentiment and daily stock price changes and directional movements. Results are reported separately for Netflix, Amazon, and Meta to capture firm-level heterogeneity in response to sentiment.

4.1 Netflix: Regression Result

Table 6 presents results from a series of regression models analyzing the relationship between financial news sentiment and daily stock price changes for Netflix. Across all specifications, the coefficient on the Netflix Sentiment Score remains positive, indicating that favorable news coverage is generally associated with upward price movements. This finding aligns with earlier descriptive analysis that identified Netflix as the most sentiment-sensitive and volatile firm in the sample.

4.1.1 Baseline and Extended Models

In the baseline model, which includes only the sentiment score, a one-unit increase in sentiment score is associated with a \$1.43 increase in Netflix's adjusted closing price ($p < 0.001$). This relationship remains strong and statistically significant (coefficient = 1.24, $p < 0.01$) even after adding macroeconomic controls (interest rate, unemployment, and inflation), suggesting that sentiment captures information beyond general economic conditions.

However, when market-level variables (NASDAQ daily change) and firm-specific trading volume are added in the third model, the sentiment coefficient falls to 0.61 and becomes statistically insignificant. This decline suggests that earlier estimates may have partially captured underlying market trends or trading activity unrelated to sentiment itself.

4.1.2 Full Fixed Effects Model

In the full model with month and year fixed effects, the sentiment coefficient increases to 0.85 ($p < 0.05$), regaining significance. This indicated that sentiment continues to exert a meaningful

influence on price movement even after controlling for unobserved seasonal and time-specific factors. This model shows the strongest overall performance, with an adjusted R² of 0.352 and the lowest RMSE (8.91), highlighting improved predictive power.

NASDAQ daily change remains a consistently strong predictor (coefficients $\approx$ 0.038, $p < 0.001$), emphasizing the importance of market-wide forces in explaining daily price movements. Additionally, the interest becomes a significant predictor (coefficient = 1.76, $p < 0.05$), reflecting sensitivity to monetary policy announcements. In contrast, unemployment and inflation remain largely insignificant.

4.1.3 Instrumental Variables (IV) Model

The IV estimate remains positive (0.8) but becomes statistically insignificant, with a considerably larger standard error, likely due to weak instrument bias. The adjusted R² also declines slightly, from 0.352 in the full OLS model to 0.335. This reflects a common tradeoff in IV models between internal validity and estimation precision.

4.2 Amazon: Regression Result

Table 7 reports regression results for Amazon. Unlike Netflix, the sentiment score does not exhibit a statistically significant relationship with daily stock price changes in any specification.

4.2.1 Baseline and Extended Models

In the baseline model, the sentiment coefficient is negative and statistically insignificant (–0.034), indicating no link between sentiment score and price movement. Adding macroeconomic controls does not meaningfully alter the outcome, with the sentiment coefficient remaining insignificant (–0.014).

With the inclusion of NASDAQ returns and trading volume, the sentiment coefficient turns slightly positive (0.049) but remains statistically insignificant. However, NASDAQ daily change emerges as a consistently significant driver of price movements (coefficient = 0.013, $p < 0.001$),

indicating that Amazon's stock responds more to aggregate market movements than firm-specific sentiment.

4.2.2 Full Fixed Effects Model

The full fixed effects model yields a similar outcome, with the sentiment coefficient remaining insignificant (0.032). NASDAQ daily change continues to exert a strong, statistically significant influence, while all other variables, including macroeconomic indicators and trading volume, are insignificant. Despite the lack of sentiment effects, the model fit is strong, with an adjusted R^2 of 0.565 and a low RMSE of 1.62. These results suggest that Amazon's short-term price behavior is primarily shaped by market-wide and seasonal factors, rather than firm-specific sentiment. This may reflect the firm's large institutional ownership base, diversified operations, and lower sensitivity to speculative news.

4.3 Meta: Regression Result

Table 8 presents regression results for Meta. Similar to Amazon, the sentiment coefficient is statistically insignificant across all specifications, indicating that financial news sentiment does not meaningfully explain Meta's short-term price changes.

4.3.1 Baseline and Extended Models

In the baseline model, the sentiment coefficient is –0.053 and insignificant. Including macroeconomic controls does not significantly alter the outcome (–0.061), although interest rate (coefficient = 0.42, $p < 0.1$) and unemployment rate (coefficient = 0.28, $p < 0.05$) become weakly significant, indicating some sensitivity to economic conditions.

Once market and firm-specific controls are added, sentiment turns slightly positive (0.0462) but remains insignificant. The NASDAQ daily change is highly significant (coefficient = 0.024, $p < 0.001$), enforcing the predominance of systematic market trends. Interest rate and unemployment remain marginally significant.

18

4.3.2 Full Fixed Effects Model

In the final model, the sentiment coefficient remains small and statistically insignificant (0.024), while NASDAQ daily change remains highly significant (0.024, $p < 0.001$). All other explanatory variables lose significance, indicating that once fixed effects are accounted for, little additional explanatory power is provided by sentiment, macroeconomic controls, or trading volume.

Despite this, the model fits moderately well, with an adjusted R² of 0.433 and an RMSE of 3.59. These findings suggest that Meta's stock price behavior is driven largely by market trends and seasonal effects rather than sentiment. This is consistent with its lower volatility and media exposure relative to Netflix.

4.4 Model Comparisons Across Firms

Table 9 summarizes model performance across all three firms. The full model consistently outperforms the baseline across R² and RMSE metrics, confirming that adding macroeconomic, financial, and time-fixed controls improves explanatory power.

Netflix demonstrates the clearest role for sentiment, with the R² improving from 0.011 to 0.375 and RMSE falling from 10.97 to 8.74. Amazon shows the largest gain in fit, from a 0.0001 R² to a 0.582 R², driven almost entirely by market-level variables. Meta already exhibits strong baseline model performance (R² = 0.427), with only marginal improvement under the full model (R² = 0.443). These differences reinforce the notion that sentiment's influence is firm-specific, contingent on characteristics such as volatility, investor base, and news sensitivity.

4.5 Logistic Regression: Directional Effects

Table 10 discusses the logistic regression results across all three companies. For Netflix, the sentiment coefficient is 0.16, positive but statistically insignificant, indicating a potential directional relationship between sentiment and the likelihood of a price increase, though the effect is not strong enough to be distinguished from random variation. For Amazon (0.0056) and Meta

19

(0.095), sentiment remains insignificant, reinforcing the absence of meaningful influence in either magnitude or direction.

NASDAQ daily change is highly significant ($p < 0.001$) across all three firms, again highlighting the importance of systematic market trends. Other control variables, including interest rate, unemployment, inflation, and trading volume, do not show consistent effects.

5 Discussion

This section interprets the empirical results in light of existing literature and firm-specific characteristics. We explore why sentiment effects are more pronounced for Netflix, compare structural differences across firms, highlight the contribution of our econometric approach, and discuss how time-series patterns support the regression findings.

5.1 Summary of Key Findings

This study examines whether financial news sentiment influences daily stock returns and return direction, revealing notable heterogeneity across three major technology firms: Netflix, Amazon, and Meta. For Netflix, sentiment emerges as a statistically significant and positive predictor of price changes in most specifications, even after controlling for macroeconomic conditions, market trends, and fixed effects. Although the IV model introduces greater estimation uncertainty, the sentiment coefficient remains directionally consistent. Logistic regression results also show a weak but positive association between sentiment and the likelihood of a price increase. In contrast, Amazon and Meta show no significant relationship between sentiment and price changes, suggesting that the impact of sentiment is highly firm-specific.

5.2 Why Sentiment Matters for Netflix

Netflix's stronger reaction to sentiment likely stems from its structural characteristics. As a highly volatile, consumer-facing company operating in a media-intensive industry, Netflix is reg-

ularly in the spotlight for events such as show releases, celebrity partnerships, and subscriber updates. These events create attention-driven shocks that can temporarily move prices away from fundamentals. Such coverage is more likely to attract retail investors, who are especially responsive to salient news and prone to sentiment-driven trading behavior (Barber and Odean, 2008). The resulting increase in trading activity and short-term volatility makes Netflix's stock more susceptible to sentiment effects. These findings align with prior literature by Tetlock (2007), Chen et al. (2014), and Cristescu et al. (2023), who report stronger sentiment effects in firms with greater media visibility, higher volatility, and larger retail investor participation.

5.3   Structural Differences in Amazon and Meta

Unlike Netflix, Amazon and Meta exhibit return behavior largely explained by market-wide forces and macroeconomic trends. In both cases, sentiment coefficients remain statistically insignificant across all models. Amazon, in particular, shows a strong dependence on NASDAQ movements, with a substantial increase in model fit after adding market-level controls—suggesting a high degree of market integration. This could reflect Amazon's size, institutional ownership, and operational scale, which buffer the firm against sentiment-induced noise. Meta, meanwhile, displays relatively strong baseline predictability and only modest improvements from additional controls. The muted role of sentiment for both firms may be attributed to their stability, lower volatility, and reduced media sensitivity.

5.4   Methodological Contributions and Literature Comparison

This study extends the literature by adopting a causal econometric framework rather than a purely predictive one. While many recent studies leverage machine learning models—such as random forests and BERT variants—for sentiment forecasting (Fazlija and Harder, 2022; Patil et al., 2024), these models rarely account for endogeneity or firm-level variation. In contrast, this paper applies OLS, fixed effects, IV, and logistic regression to estimate the marginal effect of sentiment, while explicitly addressing simultaneity bias through the use of lagged sentiment as an

instrument ([Tetlock](#), [2007](#)). This approach allows us to control for unobserved heterogeneity and time shocks, offering a more rigorous causal interpretation. Moreover, our cross-firm comparison under a consistent modeling strategy addresses a major gap in existing sentiment literature, which often analyzes firms in isolation.

5.5   Time Series Patterns and Empirical Support

Visualizations of key variables (see Appendix 7) provide further empirical support for these findings. Netflix shows pronounced variability in both sentiment scores and daily price changes, particularly during earnings cycles and major news events between 2020 and 2022. By contrast, Amazon and Meta display tighter clustering around zero in both sentiment and price volatility, consistent with more stable return dynamics. These visual patterns reinforce the regression findings and suggest that firm-specific sensitivity to sentiment is not merely statistical but reflected in the underlying time-series behavior of the data.

5.6   Limitations

While this study provides meaningful insights into how sentiment influences stock price movements, the analysis faces a few important limitations that should be taken into account.

The sentiment variable is constructed using a dictionary-based method applied to headlines, which captures word polarity but misses the broader nuance often embedded in full-text articles. This limitation means that tone, irony, and context are likely underrepresented, introducing measurement error. If this error correlates with unobserved factors that also affect price movements, then the estimates may still suffer from omitted-variable bias, despite the inclusion of market-wide and macroeconomic controls.

Another challenge lies in addressing endogeneity. The two-stage least squares approach relies on lagged sentiment as an instrument for current sentiment, which is common in prior research. However, this method relies on an exclusion restriction that cannot be directly tested. While lagged sentiment is likely exogenous to contemporaneous price changes after controlling for observed

factors, there is no guarantee that it influences prices only through current sentiment. Media cycles or external events may influence both lagged sentiment and price movements, raising concerns about instrument validity. Additionally, this strategy could only be implemented for Netflix, as the instrument lacked sufficient strength for Amazon and Meta. This restricts the ability to generalize causal inferences across firms.

The scope of the sample also limits generalizability. All three firms are large, U.S.-based technology companies with high public visibility and extensive English-language coverage. These characteristics may not reflect the media dynamics, investor profiles, or market behavior of firms in other sectors, countries, or sizes.

The use of daily data helps isolate short-term sentiment effects but does not capture within-day volatility or delayed reactions that unfold over longer timeframes. As a result, some dynamics may be missed. Future studies could enhance this work by employing higher-frequency financial and sentiment data, experimenting with more sophisticated text analysis tools, and expanding the sample to include a wider range of firms and market contexts.

6   Conclusion

This study examines whether daily financial news sentiment has a measurable impact on stock price changes and price direction for three large-cap technology firms: Netflix, Amazon, and Meta. Leveraging a multi-stage empirical framework—including OLS, fixed effects, instrumental variables, and logistic regression—we estimate both the marginal and causal effects of sentiment, while accounting for macroeconomic conditions and market-wide influences.

The analysis reveals that the role of sentiment is not uniform across firms. For Netflix, sentiment exhibits a statistically significant association with daily price change, supporting the view that media tone can influence price dynamics, particularly in firms with higher volatility and retail investor presence. In contrast, sentiment effects are negligible for Amazon and Meta, suggesting that firm-level characteristics such as investor base, media visibility, and sensitivity to narrative-driven

movements shape the extent to which sentiment matters.

These findings have both empirical and practical implications. From a modeling standpoint, they underscore the need for firm-specific approaches rather than pooled sentiment-based strategies. From an investor perspective, sentiment metrics may offer forecasting value primarily for firms operating in volatile, consumer-facing environments where attention and news cycles play a prominent role. Overall, this study contributes to the growing literature on financial text analysis by demonstrating how sentiment interacts with structural firm attributes to affect short-term asset pricing.

References

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies 21*(2), 785–818.

Chen, H., P. De, Y. J. Hu, and B.-H. Hwang (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies 27*(5), 1367–1403.

Cristescu, M. P., D. A. Mara, R. A. Nerișanu, L. C. Culda, and I. Maniu (2023). Analyzing the impact of financial news sentiments on stock prices—a wavelet correlation. *Mathematics 11*(23), 4830.

Dougal, C., J. Engelberg, D. Garcia, and C. A. Parsons (2012). Journalists and the stock market. *The Review of Financial Studies 25*(3), 639–679.

Engelberg, J. and C. A. Parsons (2011). The causal impact of media in financial markets. *The Journal of Finance 66*(1), 67–97.

Fama, E. F. (1990). Stock returns, expected returns, and real activity. *Journal of Finance 45*(4), 1089–1108.

Fazlija, B. and P. Harder (2022). Using financial news sentiment for stock price direction prediction. *Mathematics 10*(13), 2156.

Federal Reserve Bank of New York (2025). Effective Federal Funds Rate [EFFR]. Retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/EFFR (accessed March 23, 2025).

García-Medina, A., L. Sandoval Junior, E. Urrutia Báñuelos, and A. Martínez-Argüello (2018). Correlations and flow of information between the new york times and stock markets. *Physica A: Statistical Mechanics and its Applications 502*, 403–415.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance 66*(1), 35–65.

Patil, A., H. Sharma, and A. Sinha (2024). Sentiment analysis of financial news and its impact on the stock market. In *2024 2nd World Conference on Communication Computing (WCONF)*, pp. 1–5.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ryan, J. A. (2020). Quantmod: Quantitative financial modelling framework. R package version 0.4-17.

Silge, J. and D. Robinson (2016). Tidytext: Text mining and analysis using tidy data principles in r. *The Journal of Open Source Software 1*(3), 37.

Singh, A. (2025). Nyt articles: 2.1m+ (2000-present) daily updated. Kaggle. https://www.kaggle.com/datasets/aryansingh0909/nyt-articles-21m-2000-present.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance 62*(3), 1139–1168.

U.S. Bureau of Labor Statistics (2025a). Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL]. Retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/CPIAUCSL (accessed March 23, 2025).

U.S. Bureau of Labor Statistics (2025b). Unemployment Rate [UNRATE]. Retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/UNRATE (accessed March 23, 2025).

Xiao, Q. and B. Ihnaini (2023). Stock trend prediction using sentiment analysis. *PeerJ. Computer Science 9*, e1293–e1293.

Zeileis, A. and G. Grothendieck (2005). zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software 14*(6), 1–27.

7 Appendix
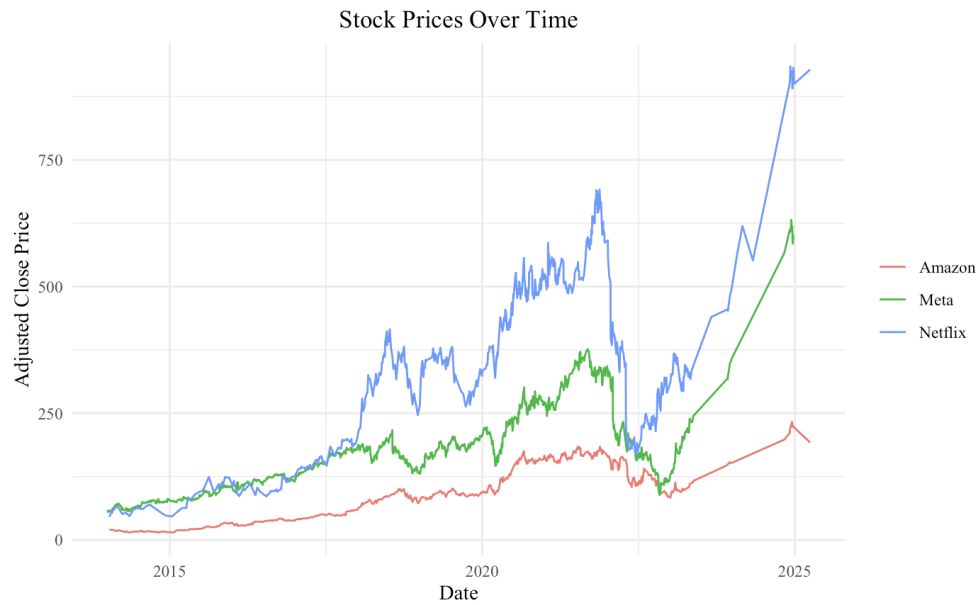
Figure 1: Time Series of Adjusted Closing Prices for Netflix, Meta, and Amazon (2014–2025)
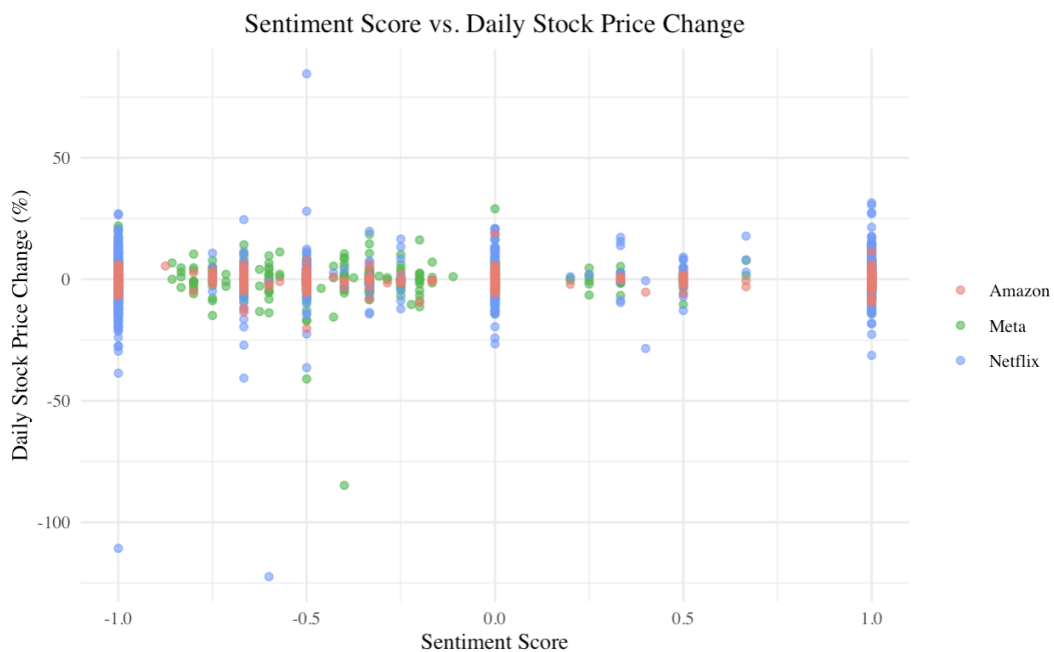
Figure 2: Sentiment Score vs. Daily Stock Price Change for Netflix, Meta, and Amazon

Note. Each point represents one trading day matched with its corresponding sentiment score derived from financial news articles. Sentiment scores are based on the Loughran-McDonald dictionary.

Date	Close Price	Daily Change	Volume	Sentiment Score	NASDAQ Change	Interest Rate	Inflation	Unemp. Rate	Price Direction
2014-01-02	54.453	–	43,195,500	-0.667	–	0.080	235.288	6.6	–
2014-01-03	54.304	-0.149	38,246,200	-1.000	-11.160	0.080	235.288	6.6	0
2014-01-06	56.932	2.628	68,852,600	-1.000	-18.230	0.080	235.288	6.6	1
2014-01-07	57.648	0.717	77,207,400	1.000	39.500	0.070	235.288	6.6	1
2014-01-10	57.668	0.717	42,449,500	-0.500	18.480	0.070	235.288	6.6	1
2014-01-14	57.469	1.821	37,503,600	-1.000	69.720	0.070	235.288	6.6	1

Table 1: Sample dataset for Meta. Amazon and Netflix datasets follow the same format

Variable	n	Mean	SD	Min	Max
Netflix Daily Change	768	-0.113	11.036	-122.420	84.570
Netflix Sentiment Score	768	-0.173	0.816	-1.000	1.000
NASDAQ Daily Change	768	9.347	153.432	-716.371	760.980
Interest Rate	764	1.267	1.442	0.040	5.330
Inflation	767	264.916	21.037	234.747	317.603
Unemployment Rate	767	5.022	2.124	3.400	14.800
Netflix Volume	768	9,044,852.953	9,692,733.123	1,144,000	133,387,500

Table 2: Summary Statistics for Netflix Variables

Variable	n	Mean	SD	Min	Max
Amazon Daily Change	710	0.004	2.457	-20.315	18.794
Amazon Sentiment Score	710	-0.246	0.764	-1.000	1.000
NASDAQ Daily Change	710	6.883	137.278	-647.170	673.080
Interest Rate	704	1.087	1.245	0.040	5.330
Inflation	709	259.082	19.723	234.747	317.603
Unemployment Rate	709	5.062	2.083	3.400	14.800
Amazon Volume	710	86,068,604.085	45,375,356.176	28,321,200	396,022,000

Table 3: Summary Statistics for Amazon Variables

Variable	n	Mean	SD	Min	Max
Meta Daily Change	1464	0.032	4.754	-84.840	29.023
Meta Sentiment Score	1465	-0.393	0.639	-1.000	1.000
NASDAQ Daily Change	1464	4.396	126.515	-624.940	673.080
Interest Rate	1455	1.058	1.228	0.040	5.330
Inflation	1465	256.064	19.305	234.747	317.603
Unemployment Rate	1465	4.995	1.799	3.400	14.800
Meta Volume	1465	26,494,957.270	17,823,100.000	6,702,000	188,119,900

Table 4: Summary Statistics for Meta Variables

Variable	VIF
Lagged Sentiment	1.151
NASDAQ Daily Change	1.006
Interest Rate	1.811
Netflix Volume	1.056
Unemployment Rate	1.338
Inflation	1.710

Table 5: VIF from First-Stage Regression

Table 6: Regression Results for Netflix Stock Returns

Variable	Baseline	Economic Controls	Financial Controls	Full Model (FE)	IV Model
(Intercept)	0.1345	7.2864	15.9144	-51.0483	0.5341
	(0.3776)	(6.1827)	(10.7053)	(64.1600)	(1.6654)
Netflix Sentiment Score	1.4336***	1.2410**	0.6085	0.8541*	0.8020
	(0.4319)	(0.4290)	(0.4182)	(0.3771)	(1.2337)
Interest Rate		0.8620+	0.7568+	1.7577*	0.2853
		(0.4946)	(0.4250)	(0.7817)	(0.2447)
Unemployment Rate		0.3936+	0.1477	0.2777	0.1567
		(0.2204)	(0.1575)	(0.2734)	(0.1626)
Inflation		-0.0387	-0.0596+	0.2513	
		(0.0275)	(0.0376)	(0.2748)	
NASDAQ Daily Change			0.0378***	0.0375***	0.0381***
			(0.0028)	(0.0028)	(0.0029)
Netflix Volume			-2.404e-07	-3.177e-07	-2.208e-07
			(1.796e-07)	(2.122e-07)	(1.742e-07)
Month Fixed Effects	No	No	No	Yes	No
Year Fixed Effects	No	No	No	Yes	No
Observations	768	764	764	764	763
R^2	0.011	0.020	0.347	0.375	0.339
Adjusted R^2	0.010	0.015	0.342	0.352	0.335
RMSE	10.98	10.98	8.97	8.91	9.03

Note. Robust standard errors in parentheses. $^+p < 0.1$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7: Regression Results for Amazon Stock Returns

Variable	Baseline	Economic Controls	Financial Controls	Full Model (FE)
(Intercept)	-0.0042	-0.1208	0.6733	1.4651
	(0.0985)	(1.7681)	(1.6521)	(9.6734)
Amazon Sentiment Score	-0.0343	-0.0143	0.0490	0.0322
	(0.1118)	(0.1115)	(0.0797)	(0.0788)
Interest Rate		0.2324	0.1283	-0.0144
		(0.1417)	(0.0998)	(0.2228)
Unemployment Rate		0.0821	0.0437	0.0471
		(0.0666)	(0.0498)	(0.0702)
Inflation		-0.0021	-0.0032	-0.0060
		(0.0079)	(0.0061)	(0.0411)
NASDAQ Daily Change			0.0132***	0.0131***
			(0.0008)	(0.0009)
Amazon Volume			-3.378e-09	-2.639e-09
			(3.637e-09)	(3.895e-09)
Month Fixed Effects	No	No	No	Yes
Year Fixed Effects	No	No	No	Yes
Observations	710	704	704	704
R^2	0.0001	0.010	0.566	0.582
Adjusted R^2	-0.0013	0.004	0.562	0.565
RMSE	2.46	2.45	1.62	1.62

Note. Robust standard errors in parentheses. $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 8: Regression Results for Meta Stock Returns

Variable	Baseline	Economic Controls	Financial Controls	Full Model (FE)
(Intercept)	0.0113	1.9534	3.2218	4.4286
	(0.1362)	(2.8214)	(2.6826)	(21.577)
Meta Sentiment Score	-0.0531	-0.0610	0.0462	0.0240
	(0.1596)	(0.1597)	(0.1174)	(0.1206)
Interest Rate		0.4219^{+}	0.2909^{+}	0.4516
		(0.2346)	(0.1630)	(0.3241)
Unemployment Rate		0.2807^{*}	0.1609^{+}	0.1109
		(0.1139)	(0.0858)	(0.1247)
Inflation		-0.0149	-0.0132	-0.0085
		(0.0129)	(0.0100)	(0.0959)
NASDAQ Daily Change			0.0235^{***}	0.0235^{***}
			(0.0016)	(0.0015)
Meta Volume			-3.817e-08	-5.455e-08
			(2.615e-08)	(3.463e-08)
Month Fixed Effects	No	No	No	Yes
Year Fixed Effects	No	No	No	Yes
Observations	1464	1454	1454	1454
R^2	0.0001	0.010	0.429	0.443
Adjusted R^2	-0.0006	0.007	0.427	0.433
RMSE	4.76	4.74	3.60	3.59

Note. Robust standard errors in parentheses. $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 9: Model Performance Metrics Across Companies and Specifications

Company	Model	R^2	F-statistic	RMSE
Netflix	Baseline	0.011	8.70	10.97
Netflix	Full	0.375	16.33	8.74
Amazon	Baseline	0.0001	0.08	2.45
Amazon	Full	0.582	36.25	1.58
Meta	Baseline	0.427	215.90	3.60
Meta	Full	0.443	43.66	3.55

Note. RMSE = Root Mean Squared Error. Models labeled "Full" include economic and financial controls, as well as fixed effects.

Table 10: Logistic Regression Results for Stock Price Direction

Variable	Netflix	Amazon	Meta
Sentiment Score	0.1605	0.0056	0.0947
	(0.1153)	(0.1340)	(0.0958)
NASDAQ Daily Change	0.0109***	0.0197***	0.0152***
	(0.0012)	(0.0024)	(0.0013)
Interest Rate	0.3385	−0.1519	0.0370
	(0.2085)	(0.3102)	(0.2189)
Volume	0.0000	−0.0000	−0.0000
	(0.0000)	(0.0000)	(0.0000)
Unemployment Rate	0.0365	0.0348	−0.0519
	(0.0820)	(0.1209)	(0.0774)
Inflation	0.0485		
	(0.0625)		
Constant	−11.7432	0.6562	0.9634
	(14.8519)	(0.8501)	(0.5880)
Month Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	764	704	1,454

Note. Robust standard errors in parentheses. $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.