Do LLMs Make Markets More Efficient?

Runjing Lu
University of Toronto

Luka Vulicevic
UC San Diego

Yongxin Xu
Monash University[*]

March 14, 2026

Abstract

This paper evaluates whether large language models (LLMs) improve or impair market efficiency. Exploiting plausibly exogenous outages at major LLM providers, we find that LLM availability eliminates at least 46% to 61% of post-news drift into the next trading day. The absence of reversal alongside the magnitude of this reduction suggests that LLMs play a *first-order* role in short-horizon price discovery. Because LLM supply is highly concentrated, outages expose a single point of failure in market-wide information processing. Consistent with this interpretation, investors insulated from major LLM outages can earn more than twice the day $t+1$ returns of the long-short news-sentiment strategy following outage events.

Keywords: LLMs, Market Efficiency, News, Sentiment
JEL Classification Numbers: G14, G12

[*]Contact: runjing.lu@utoronto.ca, lvulicevic@ucsd.edu, yongxin.xu@monash.edu. We thank Joseph Engelberg and William Mullins for useful comments. We also thank Refine.ink for checking the paper. The Rady Research Fund and the Brandes Center provided critical funding for this project.

1. Introduction

Asset prices reflect information, and the speed at which prices incorporate information is central to market efficiency. Large language models (LLMs) have profoundly altered information processing, leading to rapid adoption by institutional and retail investors (Alves, 2025, Wang, 2025). By lowering the cost of processing unstructured data, LLMs may accelerate price discovery, enabling investors to extract and act on value-relevant signals more rapidly. However, the centralized provision of LLMs creates a countervailing risk: when the same models process the same news, they can generate similar interpretations and correlated trading, amplifying price movements beyond fundamentals and creating short-term mispricing (e.g., Bikhchandani et al., 1992, Allen et al., 2006, Dou et al., 2025). This tension raises a central question: do LLMs improve or impair market efficiency?

The ideal experiment would compare two otherwise identical worlds—one with LLMs and one without—and measure how quickly each incorporates new information into prices. While such a comparison is infeasible, outages at major LLM providers offer an empirical analog. LLM provision is highly centralized: a few developers, such as OpenAI, Anthropic, and Google, serve the vast majority of retail and enterprise users (Menlo Ventures, 2025, FirstPageSage, 2026). Outages at these providers generate an abrupt shock to investors' ability to process unstructured information, but *not* to the information itself. In our sample, outages occur every four days on average and have a median duration of seven hours, resulting in frequent and economically meaningful disruptions to LLM-assisted information processing. These outages appear plausibly exogenous. They are typically caused by hardware failures, software misconfigurations, or faulty updates that are difficult to anticipate, even for the providers themselves, let alone for investors. Consistent with this view, outage incidence and duration are uncorrelated with market returns and changes in the VIX, and exhibit no systematic seasonality.

We leverage these plausibly exogenous outages to test how LLM availability affects market efficiency in the post-news drift setting of Tetlock (2007). Our approach compares return predictability for news released during versus outside LLM outages. Importantly, observable characteristics of the news articles and firms being covered are balanced across outage and non-outage periods, supporting the validity of this comparison.. If LLMs improve market efficiency by accelerating information processing, news released during outages should exhibit greater subsequent return predictability, as prices incorporate information more slowly. Conversely, if LLMs impair market efficiency, for example through correlated trading or collusive algorithms in (Dou et al., 2025), news released when LLMs are unavailable should exhibit weaker subsequent return predictability because prices deviate less from the efficient benchmark.[1]

Our findings suggest that LLMs enhance market efficiency. Following outages at any provider, a one-standard-deviation increase in news sentiment on day t predicts 25 basis points (bps) of abnormal returns on day $t + 1$, versus 13 bps during non-outage periods. Within our framework, this difference implies that LLM availability reduces news predictability by 46%, and the estimate is statistically significant at the 1% level. The elevated post-news return drift persists through day $t + 4$ with no reversal, indicating that LLMs enhance efficiency by reducing delayed incorporation rather than by offsetting overreaction. Importantly, news released during versus outside outages shows no differential pre-announcement return predictability, supporting the plausibly exogenous nature of LLM outages.

The effect magnitude varies systematically with outage intensity, consistent with a dose-response relationship. Restricting to severe outages increases the reduction in day $t + 1$ return predictability to 49%. The effect is even stronger when two or more major providers simultaneously experience outages (*multi-provider* outages), arguably the most complete disruption to LLM access. In this case, a one-standard-deviation increase in news sentiment predicts 34 bps of abnormal returns on day $t+1$ (versus 13 bps outside of outages), implying

[1]For example, if investors access the same public information and use the same LLM model to interpret it, their trades become correlated, inducing initial overreaction and subsequent reversal. Similar dynamics arise through informational cascades (Bikhchandani et al., 1992) and higher-order expectations (Allen et al., 2006).

a 61% reduction in news predictability. This monotonic increase in effect size as LLM access becomes more fully disrupted is difficult to reconcile with explanations unrelated to LLM availability and supports a causal interpretation.

All our specifications control for time-varying firm characteristics and past returns, as well as firm, news type, hour-of-day, day-of-week, month-of-year, and year-quarter fixed effects, absorbing persistent firm heterogeneity, seasonality, and common time trends.

Although LLMs enhance market efficiency, their centralized provision creates a new systemic fragility: outages temporarily remove this processing capability for a large fraction of investors, generating episodes of mispricing. To quantify the economic magnitude of these distortions, we implement a standard long-short strategy based on news sentiment released on day t, receiving returns from day $t+1$ onward. When LLMs are widely available, this strategy earns approximately 20 bps in abnormal returns on day $t+1$. During outages affecting multiple major providers, the strategy generates roughly 60 bps—a threefold increase. Severe outages at any single provider produce a smaller but still approximately twofold increase relative to baseline.

Contributions to the literature. While a large literature studies how information is incorporated into asset prices (Fama, 1970, Grossman and Stiglitz, 1980, Hong and Stein, 1999, Hirshleifer and Teoh, 2003, Tetlock, 2007, Engelberg and Parsons, 2011, Boudoukh et al., 2019), we are the first to examine how advances in *AI-driven* information processing affect market efficiency. Recent studies document that investors use LLMs to interpret financial disclosures (Blankespoor et al., 2024) and that outages at major AI providers reduce analyst activity (Bertomeu et al., 2025), yet whether AI availability affects how quickly public news is impounded into prices remains unexplored. Exploiting exogenous LLM outages, we provide the first causal evidence that LLMs enhance short-term price discovery.

We contribute to the emerging causal literature on market efficiency. Most closely related, Shive (2012) exploits large local power outages as an exogenous constraint on investor participation and shows that restricting potentially informed traders impairs price efficiency.

We extend this logic to AI-based information processing: temporary outages at major LLM providers hinder investors' ability to process public news and delay its incorporation into prices. We complement Cheng et al. (2025), who document reduced trading activity during ChatGPT outages, by showing that the impact extends beyond volume to the speed of price adjustment. More broadly, we add to a growing set of natural experiments—including Chen et al. (2020) and Chen et al. (2023)—that use exogenous shocks to identify the determinants of price efficiency.

Our findings also speak to a growing literature on how automated technologies reshape financial markets. Prior work shows that algorithmic trading enhances market quality through faster execution (Hendershott et al., 2011), improved price discovery (Brogaard et al., 2014), and strategic interaction among traders (Biais et al., 2015). Recent theory by Dou et al. (2025) shows, however, that reinforcement-learning algorithms may sustain collusive equilibria, so AI adoption need not unambiguously improve efficiency. Whether such concerns extend to generative AI—which is used primarily to process and interpret textual information rather than to execute trades autonomously—is an open empirical question, and our findings suggest that generative AI improves market efficiency.

Finally, a growing literature documents that large language models are capable of sophisticated financial analysis, including predicting returns from textual data (Lopez-Lira and Tang, 2023, Chen et al., 2025), measuring economic expectations and investor disagreement (Bybee, 2023, Bhagwat et al., 2025), analyzing corporate disclosures and policies (Li et al., 2026, Jha et al., 2024, Kim et al., 2023a,b, Beckmann et al., 2024), and replicating investor behavior (Chen et al., 2024, Fedyk et al., 2024, Ouyang et al., 2024). These studies primarily demonstrate what LLMs can achieve in controlled or quasi-experimental settings; we instead shift focus from capability to market impact, asking whether LLM availability meaningfully affects price discovery in real-world markets. We provide causal evidence that it does, and that the effect is positive.

2. Data

In this section, we describe our data sources and various outage definitions used in the paper.

2.1 LLM outages

We study outages at the three largest LLM providers: OpenAI, Anthropic (Claude), and Google (Gemini). Together, these providers account for nearly 80% of enterprise LLM usage and retail-facing AI chatbot traffic (Menlo Ventures, 2025, FirstPageSage, 2026). The high concentration of LLM provision implies that outages at these platforms constitute an economy-wide shock to investors' capacity to extract and act on value-relevant signals from unstructured text.

Outages typically arise from operational failures—software misconfigurations and the like—that are difficult to anticipate even for the engineers maintaining these systems, let alone for market participants. Because failures are provider-specific, they are also unlikely to proxy for broader infrastructure disruptions, such as cloud, internet, or power outages, that could independently affect financial markets. To see this, consider a incident from OpenAI's status history: on March 15, 2023, at 10:55 AM, a routine internal update introduced an erroneous configuration in a supporting system, rendering ChatGPT unavailable until engineers reverted the change. We provide systematic evidence in Section 4.1 supporting the plausibly exogenous nature of these events.

We obtain outage data from the official status pages maintained by each provider.[2] These pages provide real-time and historical records of service disruptions, including incident timestamps, durations, and affected services. We restrict the sample to outages on or after March 1, 2023, when OpenAI released the ChatGPT API, enabling programmatic access for institutional adoption. Prior to this date, only the public web interface was available (launched

[2]OpenAI status: https://status.openai.com/; Claude status: https://status.claude.com/; Google status: https://status.cloud.google.com/products/Z0FZJAMvEB4j3NbCJs6B/history

November 30, 2022), so institutional integration of LLMs into trading was likely limited. This restriction is not consequential: few outages occurred between the web interface launch and the API release, and results are unchanged using November 30, 2022 as the alternative start date.

Outages are classified into various types. Specifically, OpenAI categorizes incidents into four levels: *operational* (minor issues such as subscription-service hiccups), *degraded performance* (slowdowns that do not prevent core functionality), *partial outage* (significant disruptions affecting a substantial portion of users), and *full outage* (complete service unavailability). Anthropic classifies incidents analogously: *maintenance*, *minor* (limited impact on small subsets of users), *major* (widespread degradation affecting many users), and *critical* (severe or complete service failure). Gemini, a Google product, does not categorize their incidents and based on the few observations we have of Gemini outages, they are all severe outages that prevent users from using Gemini.

We restrict our attention to outages that persist beyond the NYSE close at 4:00 PM Eastern Time.[3] The rationale is that if an outage resolves intraday, investors may still process news using LLMs before market closes, limiting the outage's impact on price formation. News articles released during a qualifying outage are classified as treated. Figure 1 illustrates an example timeline for treatment definition. The timeline depicts an outage that begins after the NYSE open and lasts less than 24 hours but beyond the NYSE close. In practice, outages may start overnight and persist for multiple days. If an outage persists past the NYSE close on one trading day and resolves intraday on the next trading day, only articles published during the outage on the first day are classified as treated. If an outage persists past the NYSE close on consecutive trading days, articles published during the outage on each of those days are classified as treated.

For some analyses, we consider all reported incidents, regardless of severity, and refer to

[3] To address potential confounding this selection may bring, we include 30-minute time-of-day fixed effects in all specifications. This ensures that treated articles are compared with control articles released within the same intraday window, absorbing any mechanical differences in return dynamics driven by release timing.

these as *any* outages. These include operational disruptions at OpenAI and maintenance or minor incidents at Anthropic. While such events may not fully prevent access, they introduce frictions that capture the intensive margin of LLM availability. In other analyses, we focus on *severe* outages, defined as partial or full outages at OpenAI, major or critical incidents at Anthropic, and any outage at Gemini, which more substantially impair investors' ability to rely on LLM services. Finally, we examine *multi-provider* outages, defined as simultaneous disruptions at two or more providers. These episodes likely affect the broadest set of users and most closely approximate a systemic interruption of LLM availability.

2.2 Firm data

Our firm-level news data are based on Dow Jones Newswire accessed via RavenPack 1.0. We retain articles with ticker-specific relevance scores above 75 and measure sentiment using the RavenPack Event Sentiment Score.

We obtain daily stock returns from CRSP and compute abnormal returns by subtracting the DGTW characteristic-adjusted benchmark portfolio return from each firm's raw return (Daniel et al., 1997). Following Tetlock (2007), we also calculate the following control variables: the logarithm of book-to-market equity ($\log \text{BM}$), the logarithm of market capitalization ($\log \text{ME}$), and the logarithm of share turnover ($\log \text{Turnover}$).

2.3 Summary statistics

Table 1 Panel A reports summary statistics for LLM provider outages from March 2023 to November 2025. We identify 226 provider-level outage incidents: 158 for OpenAI, 63 for Claude, and 5 for Gemini. Of these, 90 are classified as severe, including 60 OpenAI outages and 25 Claude outages; all Gemini outages are severe. We also observe 28 multi-provider outages, of which 6 are severe. Given the small number of severe multi-provider outages, we do not analyze them separately.

Outages occur frequently, particularly for OpenAI and Claude. The average gap between consecutive outages is 6 trading days for OpenAI and 13 for Claude; restricting to severe outages increases these intervals to 16 and 24, respectively. Aggregating across providers, such an outage occurs roughly every 4 days on average, and a severe outage every 11 days. Multi-provider outages are less common, occurring approximately every 21 days.

Outage durations are right-skewed. Pooling across providers, the mean duration of any outage is 34 hours, compared to a median of 7 hours, with durations ranging from around 10 minutes to over 500 hours. The gap between the mean and median reflects occasional multi-day disruptions alongside more frequent short-lived incidents. Severe outages are slightly longer, with an average duration of 39 hours. Multi-provider outages are shorter on average (16 hours), reflecting that they are defined by the overlap of disruptions across providers, which mechanically limits their duration.

This variation in outage frequency, severity, duration, and cross-provider overlap generates substantial time-series heterogeneity in LLM availability, which we exploit in our empirical design.

3. Methods

We develop a simple and general framework to measure how LLM availability affects the speed of information incorporation into asset prices. Consider a market where firm-specific news arrives at discrete times. Let $P_{i,t}$ denote the price of asset i at time t, and define the raw return as $r_{i,t+1} = (P_{i,t+1} - P_{i,t})/P_{i,t}$. We suppress the asset subscript where no ambiguity arises. Prices follow a standard factor model: let $\mathbb{E}_t[r_{t+1}]$ denote the expected return conditional on time t information, determined by exposure to systematic risk factors. The abnormal return is the residual

$$AR_{t+1} = r_{t+1} - \mathbb{E}_t[r_{t+1}]. \tag{1}$$

Absent firm-specific news, expected abnormal returns are zero.

When news arrives at time t, it contains information relevant for future asset payoffs and thus prices. We denote by $I_t \in \mathbb{R}$ the information content of news released at time t, defined as the implied revision to fundamental value expressed in return units. The quantity I_t thus represents the price impact that would occur under instantaneous and complete processing of news. In our empirical implementation, we define I_t as our sentiment measure: RavenPack 1.0 Event Sentiment Score, but more broadly I_t can represent any informational content relevant for asset prices.

The semi-strong form of market efficiency implies that prices fully and immediately incorporate public information. Under this condition, abnormal returns on the day of the news release equal the full information content that has just arrived, $AR_t = I_t$, and no predictable drift remains on subsequent days, $AR_{t+k} = 0$ for $k \geq 1$. In practice, information incorporation is often incomplete within a single trading day. For example, a longstanding strand of the market-efficiency literature studies earnings announcements and documents post-announcement drift with evidence spanning several decades.(Ball and Brown, 1968, Hou and Moskowitz, 2005, DellaVigna and Pollet, 2009). In our news article based setting, similar studies would include Tetlock (2007), Ke et al. (2019), and Chen et al. (2022).

To model this friction, we define a market efficiency parameter $\lambda \in \mathbb{R}$, representing the fraction of information content incorporated into prices on the same day of news release. News with information content I_t thus generates abnormal returns

$$AR_t = \lambda \cdot I_t + \varepsilon_t, \tag{2}$$

$$AR_{t+1} = (1 - \lambda) \cdot I_t + \varepsilon_{t+1}, \tag{3}$$

where ε_t and ε_{t+1} are mean-zero noise terms orthogonal to I_t used to express noise in the lagged incorporation of information. When $\lambda = 1$, all information is incorporated on day t and no predictable drift remains; the market is efficient with respect to the news we study. When $\lambda = 0$, the entire price adjustment occurs with a one-period lag. Values of λ between

0 and 1 then capture partial contemporaneous incorporation.

Our framework also accommodates overreaction. When $\lambda > 1$, prices initially move by more than the information warrants on day t, generating expected returns on day $t+1$ in the opposite direction of the news—a partial reversal of the initial price movement. Empirically, both underreaction (delayed incorporation) and overreaction (excessive initial response) are plausible, and our framework nests both possibilities to let data determine which pattern prevails.

For tractability, we restrict attention to a two-period incorporation window, assuming information is fully incorporated by $t+1$: $AR_{t+k} = \varepsilon_{t+k}$ for $k \geq 2$. This assumption can be relaxed at the cost of additional notation and modeling complexity; our empirical analysis examines horizons beyond one day as a robustness check, though this is not central to the paper and avoided in the framework.

3.1 The Role of Large Language Models

Large language models have fundamentally altered how market participants process textual information, leading to their rapid adoption by both institutional and retail investors (Alves, 2025, Wang, 2025).

The effect of LLM adoption on market efficiency is theoretically ambiguous, however. On one hand, LLMs may enhance efficiency by lowering the cost of extracting value-relevant signals from unstructured text, allowing investors to respond more quickly to news (larger λ), thereby accelerating price discovery and reducing post-announcement drift. On the other hand, widespread LLM adoption could impair efficiency. For instance, Dou et al. (2025) theoretically examine the consequences of AI collusion, defined as "a scenario where autonomous, self-interested RL algorithms independently learn to coordinate their trading in a way that secures supra-competitive profits, without explicit agreements, communication, or pre-programmed intent." In our setting, such collusion could arise if investors all receive the same news and process it through the same LLMs, producing correlated signals I_t. Their simultaneous trading could then cause prices to overreact at time t ($\lambda > 1$), generating a reversal at $t+1$.

To test which effect dominates, we allow the efficiency parameter to depend on LLM availability. Let $O_t \in \{0, 1\}$ indicate whether major LLM providers experience service outages at time t, with $O_t = 1$ denoting an outage. The efficiency parameter becomes

$$\lambda(O_t) = \lambda_0 + \Delta\lambda \cdot (1 - O_t), \tag{4}$$

where λ_0 is baseline efficiency absent LLM access (the "baseline" regime) and $\Delta\lambda$ captures the change in efficiency attributable to LLM availability. LLMs improve efficiency if and only if prices deviate less from fundamentals when LLMs are available than when they are not: LLMs improve efficiency when $|1-(\lambda_0+\Delta\lambda)| < |1-\lambda_0|$ and impairs otherwise. Specifically, when information is incorporated incompletely in the baseline regime, $0 < \lambda_0 < 1$, LLMs improve efficiency if $0 < \Delta\lambda < 2(1-\lambda_0)$ and impair efficiency if $\Delta\lambda < 0$ or $\Delta\lambda > 2(1-\lambda_0)$. Under the alternative in which markets over-incorporate information in the baseline regime, $\lambda_0 > 1$, LLMs improve efficiency if $-2(\lambda_0 - 1) < \Delta\lambda < 0$ and impair efficiency if $\Delta\lambda > 0$ or $\Delta\lambda < -2(\lambda_0 - 1)$. Substituting (4) into (2)–(3) yields state-dependent abnormal returns. For instance, when LLMs are available ($O_t = 0$):

$$AR_t = (\lambda_0 + \Delta\lambda) \cdot I_t + \varepsilon_t, \tag{5}$$

$$AR_{t+1} = (1 - \lambda_0 - \Delta\lambda) \cdot I_t + \varepsilon_{t+1}. \tag{6}$$

and analogously, during outages ($O_t = 1$) we have:

$$AR_t = \lambda_0 \cdot I_t + \varepsilon_t, \tag{7}$$

$$AR_{t+1} = (1 - \lambda_0) \cdot I_t + \varepsilon_{t+1}. \tag{8}$$

To quantify how LLM availability changes market efficiency, we compute:

$$\phi = \frac{\Delta\lambda}{1 - \lambda_0}. \tag{9}$$

The numerator captures how LLM availability changes next-day return predictability. The denominator reflects the baseline deviation from the efficient benchmark at day t since $\mathbb{E}[AR_{t+1} \mid I_t] = (1 - \lambda_0)I_t$ during outages and $\mathbb{E}[AR_{t+1} \mid I_t] = (1 - \lambda_0 - \Delta\lambda)I_t = (1 - \lambda_0)(1 - \phi)I_t$ when LLMs are available. ϕ thus measures the proportional reduction in baseline inefficiency when LLMs are available. When the baseline regime features underreaction ($0 < \lambda_0 < 1$), $\phi > 0$ indicates that LLMs accelerate price discovery by reducing delayed incorporation, whereas $\phi < 0$ implies that they slow price discovery. When the baseline regime features overreaction ($\lambda_0 > 1$), $\phi > 0$ indicates that LLMs mitigate overreaction by reducing reversals, whereas $\phi < 0$ implies that they amplify overreaction. In either case, market efficiency improves if $0 < \phi < 2$ and is maximized at $\phi = 1$.

3.2 Identification

Identification of λ_0 and $\Delta\lambda$ exploits variation in abnormal returns across outage states. The key identifying assumption is that outages are exogenous to news and market conditions:

$$O_t \perp (I_t, \varepsilon_t, \varepsilon_{t+1}). \tag{10}$$

In other words, outage timing is uncorrelated with the content of firm news and other factors affecting returns. We view this assumption as plausible: outages typically arise from technical failures, infrastructure constraints, capacity limits, or software bugs—events that are difficult to anticipate and unrelated to firm fundamentals or aggregate market conditions. As we will show later, this assumption appears to hold.

Under the exogeneity assumption, the difference in expected day $t+1$ abnormal returns, or sensitivity of day $t+1$ returns to news content, across outage states identifies the LLM effect:

$$\mathbb{E}[AR_{t+1} \mid O_t = 1, I_t] - \mathbb{E}[AR_{t+1} \mid O_t = 0, I_t] = \Delta\lambda \cdot I_t. \tag{11}$$

3.3 Econometric Implementation

With the key object of interest defined (ϕ), we now show this model can be measured with a regression specification. We index news articles by i, article release date by t, and the firm mentioned by j. Article-level sentiment $S_{i,t}$ serves as a proxy for information content $I_{i,t}$, and O_t indicates whether an outage occurred when the article was published at time t. We estimate

$$AR_{i,t+1} = \alpha + \beta_1 S_{i,t} + \beta_2 (S_{i,t} \times O_t) + \gamma O_t + \mathbf{X}'_{i,t}\boldsymbol{\delta} + \eta_{i,t+1}, \tag{12}$$

where $\mathbf{X}_{i,t}$ contains controls including time-varying firm characteristics, and various fixed effects. $\eta_{i,t+1}$ is the error term. Standard errors are clustered by date. The coefficient β_1 captures the day $t+1$ return response to news when LLMs are available, corresponding to $1 - \lambda_0 - \Delta\lambda$ in our framework. The interaction coefficient β_2 captures the additional return response during outages, equal to $\Delta\lambda$. The sum $\beta_1 + \beta_2 = 1 - \lambda_0$ equals total delayed incorporation or reversal in the baseline regime. Therefore, $\phi = \frac{\beta_2}{\beta_1 + \beta_2}$ estimates the proportional change in the magnitude of next-day predictability attributable to LLM availability.

We note that our estimates represent a lower bound on the true effect of LLMs on market efficiency. Multiple LLM providers coexist, and many institutional investors rely on proprietary or locally deployed models that remain available during outages at major providers, albeit at lower capability or higher cost. As a result, an outage at major providers removes LLM access for only a subset of market participants. The effects we estimate therefore capture the marginal contribution of the specific providers studied, and understate the overall impact of LLMs as a technology class.

4. Results

This section proceeds in three steps. We first assess the exogeneity assumption of LLM outages. We then examine how LLM availability affects the speed at which public news is incorporated into prices. Finally, we quantify the economic magnitude of delayed incorporation through a news-sentiment-based trading strategy.

4.1 Exogeneity of LLM outages

A key identifying assumption of our analysis is that LLM outages are exogenous to market conditions and firm-level news. We assess this assumption in this section.

We begin by examining whether outages systematically relate to aggregate market conditions. Table 1 Panel B reports Pearson correlations between daily outage measures, both indicator variables and outage hours, and daily market returns and changes in the VIX. Across the board, the correlations are economically small and statistically insignificant.

Figure 2 reinforces this conclusion visually. The top panel plots cumulative outage hours (left axis) together with the market and VIX (right axis), both normalized to 100 at the start of the sample. Cumulative outages rise in discrete bursts over time and show no discernible co-movement with aggregate market performance or volatility. Notably, sharp VIX spikes, such as those in early 2024 and spring 2025, do not coincide systematically with outage episodes. The bottom panel presents daily outage hours. Outages are distributed throughout the sample rather than concentrated during periods of market stress. Figure A1 plots cumulative outage hours and average daily duration by outage type and finds similar patterns.

Furthermore, this dispersion of outages also holds across intraday and calendar dimensions. Figure A2 indicates that outages are broadly distributed across the trading day. Severe and multi-provider outages in particular exhibit no clustering around the market open or close. Table A1 reports outage rates by day of the week and month. Day-of-week variation is modest, with outage rates broadly similar across weekdays and only a slight mid-week uptick. Monthly variation is somewhat larger: rates are lower in mid-summer and higher toward year-end, peaking in November and December. But this seasonality does not align with known seasonal patterns in news flow or market activity and largely reflects the limited sample of LLM outages available at the time of writing.

While LLM outages are plausibly exogenous in timing, outage periods could still, by chance, differ from non-outage periods. A key diagnostic is therefore whether observable characteristics are balanced across the two regimes. Table 2 reports balance tests comparing news-level and firm-level characteristics for articles published during outages versus outside outages, using both raw mean differences and the normalized differences of Imbens and Wooldridge (2009). Across all outage definitions, normalized differences are small and close to zero. Article count per hour and average news sentiment are nearly identical across regimes, and firm characteristics, including log market equity, log book-to-market, and log turnover, exhibit negligible differences. The largest absolute normalized difference is 0.10 (log book-to-market for multi-provider outages), well below the 0.25 threshold suggested by Imbens and Wooldridge (2009) for covariate balance. Lagged abnormal returns at various horizons are also balanced, with no normalized difference exceeding 0.05.

Overall, the evidence indicates that outage occurrence is not systematically related to aggregate market conditions and does not display calendar variation likely to confound our empirical design, especially given the inclusion of hour-of-day, day-of-week, month-of-year, and year-quarter fixed effects.

4.2 Effect of LLM outages on post-news drift

To estimate the effect of LLM availability on market efficiency, we use the regression specification in equation (12). All specifications control for time-varying firm characteristics, past returns, and firm, news-type, hour-of-day, day-of-week, month-of-year, and year-quarter fixed effects, absorbing persistent firm heterogeneity, seasonality, calendar effect, and common time trends.

We begin with the broadest definition, exploiting *any* outage regardless of severity. Table 3 reports the results. Column (1) shows that a one-standard-deviation increase in news sentiment $S_{i,t}$ on day t is associated with a 39 bps increase in same-day abnormal returns ($t = 4.15$). The interaction term $S_{i,t} \times O_t$ is statistically insignificant, indicating that contemporaneous price reactions are unaffected by outages.

The difference emerges on day $t + 1$ (Column 2). For news released when LLMs are available, a one-standard-deviation increase in sentiment predicts a 13 bps abnormal return the following day. For news published during outages, however, the return predictability increases to 25 bps (13.48 + 11.42), almost twice as large. The implied efficiency share $\hat{\phi}$ is 45.9% and significant at the 1% level, implying that LLM availability eliminates roughly 46% of the delayed incorporation of news sentiment into prices.[4] Moreover, the positive coefficient of $S_{i,t}$ on day $t + 1$ is inconsistent with LLM-driven overreaction, which would predict subsequent reversal rather than drift. At longer horizons, as shown in columns (3)-(5), the interaction term remains stable and statistically significant. This suggests that the impact of outages concentrates primarily on day $t + 1$, consistent with the short duration of most outages: by the time cumulative returns over two to four days are measured, the outage has typically ended and LLM-assisted trading has resumed.

We also examine pre-announcement return dynamics. Appendix Table A2 reports the coefficients on the interaction term $S_{i,t} \times O_t$ for abnormal returns from day $t - 4$ to $t - 1$. The estimates are economically small and statistically insignificant throughout, indicating no differential return predictability prior to the event. This absence of pre-trends further supports the plausibly exogenous nature of LLM outages.

We next restrict our analysis to *severe* outages at any single provider. Table 4 reports the results. As expected, given the stronger disruption to LLM access, the estimated increase

[4]During these outages, some investors may retain access to alternative providers or locally deployed models. Our estimates therefore reflect a lower bound on the contribution of LLM-assisted processing to market efficiency.

in day $t + 1$ return predictability when multiple providers are unavailable is larger than in the any-outage case. In Column (2), the interaction coefficient is 12.74 ($t = 2.69$), implying that a one-standard-deviation increase in news sentiment generates an additional 13 bps of abnormal returns on day $t+1$ when news is released during a severe outage. The implied efficiency share is $\hat{\phi} = 49\%$ and significant at the 1% level, indicating that severe disruptions at a single major provider eliminate roughly half of next-day return predictability attributable to LLM availability.

The strongest effects emerge under *multi-provider* outages, during which LLM services are unavailable across multiple providers simultaneously. Because these events remove access to more than one major provider at once, they most closely approximate a market-wide loss of LLM-assisted information processing. Table 5 reports the results. In Column (2), the interaction coefficient is 20.83 ($t = 2.38$), with implied efficiency share, $\hat{\phi}$, equal to 61.1%. While the economic impact of multi-provider outages is the largest in our analyses, the statistical significance is slightly lower, mostly due to the limited number of such events (28 multi-provider outages versus 90 severe and 226 any outages). We note that $\hat{\phi}$ is unstable in Columns (4) and (5) because return drift at longer horizons is small following multiple-provider outages, driving the denominator $\hat{\beta}_1 + \hat{\beta}_2$ toward zero. These extreme values are imprecise, and we caution against over-interpreting them.

The estimated effects display a clear dose-response pattern: efficiency losses are 46-49% for single-provider outages and rise to over 61% for multi-provider outages. When only one provider is disrupted, investors can partially substitute toward alternative platforms; when multiple providers are simultaneously unavailable, a larger fraction of LLM-processing capacity is impaired, leading to a more pronounced delay in price discovery. The monotonic increase in effect size as LLM access becomes more limited is difficult to reconcile with explanations unrelated to LLM availability and strengthens a causal interpretation of the results.

4.3 LLM Outages and Trading Strategies

We now ask whether this delayed adjustment translates into trading profits for an investor who retains ability to quickly process news during LLM outages—for instance, through locally deployed models not exposed to provider disruptions. We construct a simple long-short portfolio that buys stocks with above-zero news sentiment and sells stocks with below-zero sentiment, holding each position in equal weight from day t through $t+4$, and compute average DGTW-adjusted daily abnormal returns separately for news released during and outside outages.[5]

Figure 3 plots the results. Panel A uses the broadest any-outage definition. Returns on the day of news publication are similar regardless of outage status, but a divergence appears on day $t+1$, with returns about 80% higher for news released during outages. Panel B applies the severe-outage definition and shows a larger gap, with day $t+1$ returns during outages roughly twice those outside outages. Panel C focuses on multi-provider outages and shows strongest divergence: day $t+1$ returns are roughly three times higher. This monotonic pattern, in which a stricter outage definition yields a larger return gap, is consistent with the dose-response relationship in Tables 3-5: more severe LLM disruptions produce larger deviations from efficient pricing.

Finally, in all three panels the outage and non-outage series converge by day $t+2$, indicating that the return differential is short-lived and dissipates as LLM access is restored. This is consistent with the interaction coefficients in Tables 3-5 shrinking toward zero beyond $t+1$: the outage effect reflects a discrete, one-day delay in price discovery rather than a persistent disruption.

These results point to an important tradeoff in the use of LLMs in financial markets. The same concentration of information processing among a small number of providers that allows us to identify the effect of outages on market efficiency also creates a source of market-wide

[5] Because article release times vary within the trading day, day t returns are not fully tradable in a realistic implementation.

fragility. When much of the market's information processing depends on just a few platforms, outages at any one of them can impair price discovery. The multi-provider results illustrate this clearly: simultaneous outages at only a few providers lead to a threefold increase in return predictability of public news, revealing a systemic vulnerability created by the current structure of LLM provision.

5. Conclusion

This paper shows that large language models accelerate price discovery in equity markets. Using outages at major LLM providers as plausibly exogenous shocks to investors' access to state-of-the-art language-processing tools, we find that LLM availability reduces post-news drift into the next trading day by 46% to 61%. The effect strengthens with outage severity and is largest during simultaneous disruptions across providers, consistent with a dose-response pattern in which greater impairment of LLM access leads to larger delays in price adjustment. We also find no evidence of subsequent reversal, suggesting that LLMs improve market efficiency by speeding the incorporation of public information into prices.

Our results also point to a broader implication for market structure. For investors that rely on unstructured text, access to state-of-the-art LLMs through major third-party providers has become increasingly indispensable: relative to dictionary-based methods or smaller local models, these services offer better performance without requiring firms to build and maintain comparable models in-house. But these same advantages have concentrated information processing in the hands of a small number of providers, tying market efficiency to their operational reliability. When major providers experience outages—which occur frequently in our sample—price discovery slows measurably across the market. More broadly, this suggests that AI can improve market efficiency while also creating a new source of systemic fragility.

References

Allen, Franklin, Stephen Morris, and Hyun Song Shin, 2006, Beauty contests and iterated expectations in asset markets, *The Review of Financial Studies* 19, 719–752.

Alves, Joice, 2025, "ChatGPT, what stocks should I buy?' AI fuels boom in robo-advisory market, Accessed: 2026-01-14.

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159–178.

Beckmann, Lars, Heiner Beckmeyer, Ilias Filippou, Stefan Menze, and Guofu Zhou, 2024, Unusual financial communication: Chatgpt, earnings calls, and financial markets, *Olin Business School Center for Finance & Accounting Research Paper* .

Bertomeu, Jeremy, Yupeng Lin, Yibin Liu, and Zhenghui Ni, 2025, The impact of generative ai on information processing: Evidence from the ban of chatgpt in italy, *Journal of Accounting and Economics* 101782.

Bhagwat, Vineet, J Anthony Cookson, Chukwuma Dim, and Marina Niessner, 2025, The market's mirror: Revealing investor disagreement with llms, *FEB-RN Research Paper* .

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2015, Equilibrium fast trading, *Journal of Financial economics* 116, 292–313.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of Political Economy* 100, 992–1026.

Blankespoor, Elizabeth, Joe Croom, and Stephanie M Grant, 2024, Generative ai and investor processing of financial information, *Available at SSRN 5053905* .

Boudoukh, Jacob, Ronen Feldman, Shimon Kogan, and Matthew Richardson, 2019, Information, trading, and volatility: Evidence from firm-specific news, *The Review of Financial Studies* 32, 992–1033.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-frequency trading and price discovery, *The Review of Financial Studies* 27, 2267–2306.

Bybee, Leland, 2023, Surveying generative ai's economic expectations, *arXiv preprint arXiv:2305.02823* .

Chen, Haiqiang, Ming Gu, and Bo Ni, 2023, How price limit affects the market efficiency in a short-sale constrained market? evidence from a quasi-natural experiment, *Journal of Empirical Finance* 73, 22–39.

Chen, Jian, Guohao Tang, Guofu Zhou, and Wu Zhu, 2025, Chatgpt and deepseek: Can they predict the stock market and macroeconomy?, *arXiv preprint arXiv:2502.10008* .

Chen, Shuaiyu, T Clifton Green, Huseyin Gulen, and Dexin Zhou, 2024, What does chatgpt make of historical stock returns? extrapolation and miscalibration in llm stock return forecasts, *arXiv preprint arXiv:2409.11540* .

Chen, Yifei, Bryan T Kelly, and Dacheng Xiu, 2022, Expected returns and large language models, *Available at SSRN 4416687* .

Chen, Yong, Bryan Kelly, and Wei Wu, 2020, Sophisticated investors and market efficiency: Evidence from a natural experiment, *Journal of Financial Economics* 138, 316–341.

Cheng, Qiang, Pengkai Lin, and Yue Zhao, 2025, Does generative ai facilitate investor trading? early evidence from chatgpt outages, *Journal of Accounting and Economics* 101821.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *The Journal of finance* 52, 1035–1058.

DellaVigna, Stefano, and Joshua M Pollet, 2009, Investor inattention and friday earnings announcements, *The Journal of Finance* 64, 709–749.

Dou, Winston Wei, Itay Goldstein, and Yan Ji, 2025, Ai-powered trading, algorithmic collusion, and price efficiency, *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper, The Wharton School Research Paper* .

Engelberg, Joseph E, and Christopher A Parsons, 2011, The causal impact of media in financial markets, *the Journal of Finance* 66, 67–97.

Fama, Eugene F, 1970, Efficient capital markets: A review of theory and empirical work, *The journal of Finance* 25, 383–417.

Fedyk, Anastassia, Ali Kakhbod, Peiyao Li, and Ulrike Malmendier, 2024, Ai and perception biases in investments: An experimental study, *Available at SSRN 4787249* .

FirstPageSage, 2026, Top generative ai chatbots by market share – january 2026, https://firstpagesage.com/reports/top-generative-ai-chatbots/, Accessed January 21, 2026.

Grossman, Sanford J, and Joseph E Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of finance* 66, 1–33.

Hirshleifer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337–386.

Hong, Harrison, and Jeremy C Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of finance* 54, 2143–2184.

Hou, Kewei, and Tobias J Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *The Review of Financial Studies* 18, 981–1020.

Imbens, Guido W, and Jeffrey M Wooldridge, 2009, Recent developments in the econometrics of program evaluation, *Journal of economic literature* 47, 5–86.

Jha, Manish, Jialin Qian, Michael Weber, and Baozhong Yang, 2024, Chatgpt and corporate policies, Technical report, National Bureau of Economic Research.

Ke, Zheng Tracy, Bryan T Kelly, and Dacheng Xiu, 2019, Predicting returns with text data, Technical report, National Bureau of Economic Research.

Kim, Alex, Maximilian Muhn, and Valeri Nikolaev, 2023a, Bloated disclosures: can chatgpt help investors process information?, *arXiv preprint arXiv:2306.10224* .

Kim, Alex, Maximilian Muhn, and Valeri Nikolaev, 2023b, From transcripts to insights: Uncovering corporate risks using generative ai, *arXiv preprint arXiv:2310.17721* .

Li, Kai, Feng Mai, Rui Shen, Chelsea Yang, and Tengfei Zhang, 2026, Dissecting corporate culture using generative ai, *The Review of Financial Studies* 39, 253–296.

Lopez-Lira, Alejandro, and Yuehua Tang, 2023, Can chatgpt forecast stock price movements? return predictability and large language models, *arXiv preprint arXiv:2304.07619* .

Menlo Ventures, 2025, 2025 mid-year llm market update: Foundation model landscape + economics, Menlo Ventures Perspective.

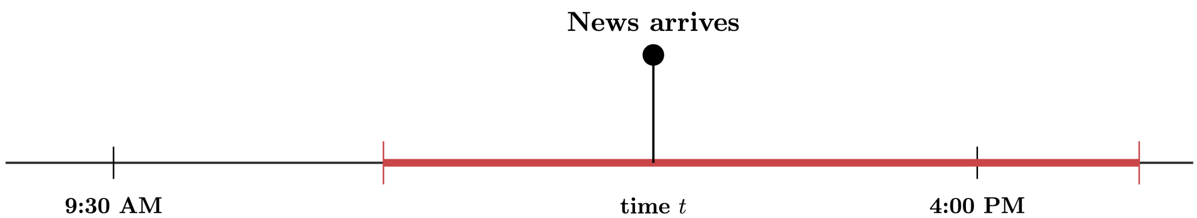
Ouyang, Shumiao, Hayong Yun, and Xingjian Zheng, 2024, Ai as decision-maker: Risk preferences of llms, *Available at SSRN 4851711* .

Shive, Sophie, 2012, Local investors, price discovery, and market efficiency, *Journal of Financial Economics* 104, 145–161.

Tetlock, Paul C, 2007, Giving content to investor sentiment: The role of media in the stock market, *The Journal of finance* 62, 1139–1168.

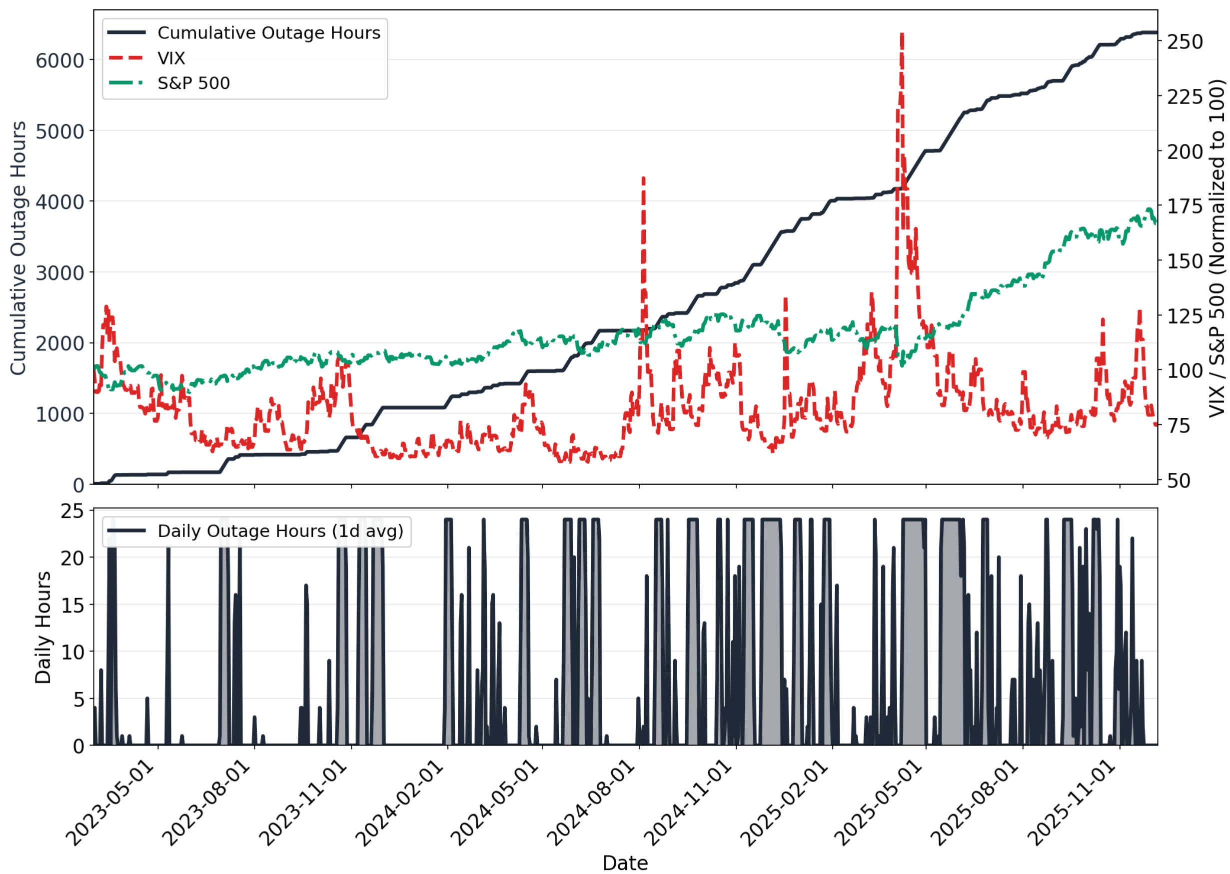
Wang, Lu, 2025, JPMorgan, Robeco Quietly Deploy AI in Daily Wall Street Routines, Accessed: 2026-01-14.

Figure 1: Outage Treatment Definition



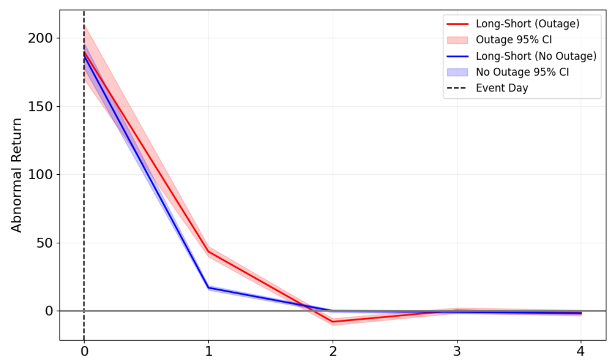
Note: This figure illustrates the intraday timing convention used to define treatment by an LLM outage. We focus on outages that remain ongoing through the NYSE close at 4:00 PM ET, shown by the red line segment extending beyond 4:00 PM. The rationale is that if an outage resolves intraday, investors may still process news using LLMs before placing closing trades, limiting its effect on price formation. News articles released while such an outage is in effect are classified as treated. Outages may begin before the NYSE opens and may persist for multiple days. If an outage persists past the NYSE close on one trading day and resolves intraday on the following trading day, only articles published during the outage on the first trading day are treated; if it persists past the NYSE close on consecutive trading days, articles published during the outage on each of those days are treated.

Figure 2: Time Series of LLM Outages

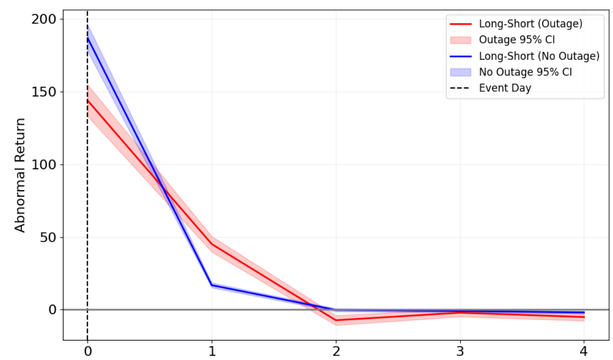


Note: The top panel plots cumulative LLM provider outage hours (left axis) together with the VIX and the market (right axis). The VIX and the market series are normalized to 100 at the start of the sample. The bottom panel shows average daily outage hours across OpenAI, Anthropic, and Google. The sample runs from March 2023 through November 2025. Appendix Figure A1 plots the time series by outage type.

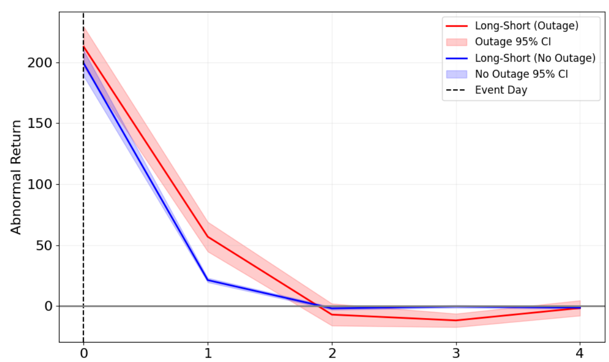
Figure 3: News Sentiment Portfolio Returns for News Released During versus Outside LLM Outages



(a) Any outages



(b) Severe outages



(c) Multi-provider outages

Note: This figure plots event-time daily abnormal returns for a long–short news sentiment–based trading strategy, separately for news released during outage and non-outage periods. Event time is measured in trading days relative to the news publication date, with day 0 (vertical dashed line) denoting the release day. The blue series reports abnormal returns on non-outage days, while the red series reports abnormal returns when the news is released during an outage. Shaded bands represent 95% confidence intervals. Panel (a) includes outages of any type at any provider. Panel (b) considers severe outages at any provider. Panel (c) focuses on outages during which at least two providers simultaneously experience an outage.

Table 1: Summary Statistics

Panel A: Frequency and duration of outages

	OpenAI		Claude		Gemini		Any provider		Multi-provider	
	All	Severe	All	Severe	All	Severe	All	Severe	All	Severe
Frequency										
Count	158	60	63	25	5	5	226	90	28	6
Gap	5.9	16.4	13.4	23.8	52.2	52.2	4.0	10.7	21.4	84.6
Duration										
Mean	39.7	46.5	18.6	18.6	46.3	46.3	34.0	38.8	16.0	17.2
Median	7.5	8.2	5.5	3.4	36.8	36.8	7.0	6.1	5.0	2.6
Std. Dev.	76.8	70.8	36.6	42.9	40.8	40.8	68.0	64.0	27.3	21.6
Min	0.2	0.4	0.3	0.3	6.0	6.0	0.2	0.3	0.6	1.5
Max	509.1	241.3	212.4	212.4	114.6	114.6	509.1	241.3	139.5	54.0

Panel B: Correlations between LLM outages and market outcomes

	Outage type		
	All	Severe	Multi-provider
Market returns			
Outage Indicator	−0.003	−0.062	0.026
	[−0.08]	[−1.64]	[0.70]
Outage Hours	0.016	−0.042	−0.019
	[0.42]	[−1.13]	[−0.50]
Δ *VIX*			
Outage Indicator	−0.010	0.057	−0.006
	[−0.27]	[1.51]	[−0.16]
Outage Hours	−0.028	0.018	0.013
	[−0.73]	[0.48]	[0.34]

Note. This table reports summary statistics for LLM outages from March 2023 to November 2025. Panel A reports summary statistics for outage frequency and duration. Frequency is measured by the number of outages (Count) and the average number of days between outages (Gap). Duration is measured in hours. Columns split statistics by provider and by outage type. "Any provider" reports statistics for the union outage state, i.e., periods when at least one provider experiences an outage. "Multi-provider" focuses on outages during which at least two providers experience simultaneous outages. Severe outages are defined as full or partial outages for OpenAI, critical or major incidents for Claude, and all observed Gemini outages (as Gemini only reports severe outages). Panel B reports Pearson correlations between daily outage measures (indicator and duration in hours) and daily market outcomes (Market returns and percent change in VIX), with t-statistics in square brackets by outage type. "All" refers to outages of any type at any provider; "Severe" refers to severe outages at any provider; and "Multi-provider" refers to periods during which at least two providers experience simultaneous outages. The sample includes $n = 704$ trading-day observations.

Table 2: Balance Tests between Outage and Non-Outage Periods

Variable		Outage	Non-Outage	Difference	Difference (z)
	All	556.92	532.30	24.62	0.03
Articles per Hour	Severe	550.63	532.30	18.33	0.03
	Multi-provider	555.93	532.30	23.63	0.03
	All	0.06	0.07	-0.01	-0.01
Sentiment	Severe	0.06	0.07	-0.01	-0.01
	Multi-provider	0.05	0.07	-0.02	-0.04
	All	16.97	16.96	0.01	0.00
log(ME)	Severe	16.86	16.96	-0.10	-0.02
	Multi-provider	16.94	16.96	-0.02	-0.00
	All	-0.84	-0.76	-0.07	-0.05
log(BM)	Severe	-0.82	-0.76	-0.06	-0.04
	Multi-provider	-0.92	-0.76	-0.15	-0.10
	All	-1.87	-1.86	-0.01	-0.00
log(Turnover)	Severe	-1.85	-1.86	0.02	0.00
	Multi-provider	-1.82	-1.86	0.04	0.01
	All	12.79	16.49	-3.70	-0.00
AR_{t-1}	Severe	15.04	16.49	-1.45	-0.00
	Multi-provider	49.93	16.49	33.44	0.03
	All	10.53	9.49	1.04	0.00
AR_{t-2}	Severe	13.69	9.49	4.20	0.00
	Multi-provider	32.14	9.49	22.65	0.02
	All	86.56	31.14	55.41	0.02
$\text{AR}_{t-3:t-21}$	Severe	90.00	31.14	58.86	0.02
	Multi-provider	-7.23	31.14	-38.37	-0.02
	All	287.40	198.84	88.56	0.01
$\text{AR}_{t-252:t-22}$	Severe	212.76	198.84	13.92	0.00
	Multi-provider	688.31	198.84	489.47	0.05

Note. This table compares characteristics of news articles and firms with news published during LLM outage and non-outage periods, separately by outage type. Rows correspond to outage definitions described in Table 1; the non-outage sample is common across rows. Articles per Hour is the average number of news articles per hour. Sentiment is the Dow Jones Newswire event sentiment score. Firm characteristics include log market equity (log(ME)), log book-to-market (log(BM)), log share turnover (log(Turnover)), and cumulative abnormal returns over the various horizons. "Difference" reports the difference in unconditional means between outage and non-outage periods. "Difference (z)" reports the normalized difference following Imbens and Wooldridge (2009), defined as $(\bar{y}_1 - \bar{y}_0)/\sqrt{(s_1^2 + s_0^2)}$; values below $|0.25|$ indicate adequate covariate balance.

Table 3: Return Predictability During Any Provider Outage

	AR_t (1)	AR_{t+1} (2)	$AR_{t+1:t+2}$ (3)	$AR_{t+1:t+3}$ (4)	$AR_{t+1:t+4}$ (5)
$S_{i,t}(z)$	39.15***	13.48***	13.63***	12.23***	11.35***
	[4.15]	[8.20]	[7.47]	[6.23]	[5.56]
$S_{i,t}(z) \times O_t$	17.55	11.42***	10.09**	10.15**	11.67**
	[0.95]	[3.09]	[2.18]	[2.08]	[2.19]
O_t	-7.74	-1.95	-0.38	-8.63	-5.84
	[-0.29]	[-0.27]	[-0.05]	[-0.99]	[-0.63]
AR_t		-0.03***	-0.03***	-0.01***	-0.01***
		[-8.25]	[-8.94]	[-4.11]	[-3.93]
AR_{t-1}	0.02	0.01	0.03**	0.03***	0.03***
	[0.15]	[0.88]	[2.26]	[2.64]	[2.61]
AR_{t-2}	0.11	0.01	0.00	0.01	0.01
	[0.85]	[0.80]	[0.43]	[1.00]	[0.91]
$AR_{t-21:t-3}$	-0.32	-0.01	-0.01	-0.01**	-0.02***
	[-1.64]	[-1.36]	[-1.50]	[-2.21]	[-3.01]
$AR_{t-252:t-22}$	-0.23*	-0.01	-0.01*	-0.01**	-0.01***
	[-1.65]	[-1.62]	[-1.92]	[-2.05]	[-3.07]
log(BM)	64.94	28.93**	48.92***	64.32***	70.05***
	[0.98]	[2.08]	[3.05]	[3.62]	[3.53]
log(ME)	399.07***	-77.23***	-96.01***	-108.32***	-121.65***
	[2.82]	[-8.95]	[-9.29]	[-9.16]	[-9.44]
log(Turnover)	418.42***	12.60	14.90*	2.99	5.02
	[4.61]	[1.64]	[1.73]	[0.32]	[0.50]
$\hat{\beta}_1 + \hat{\beta}_2$	56.71***	24.91***	23.71***	22.38***	23.02***
	[4.09]	[8.17]	[6.07]	[5.34]	[4.92]
$\hat{\phi}$	31.0	45.9***	42.5***	45.3***	50.7***
	[1.18]	[4.54]	[3.18]	[3.15]	[3.59]
Firm FE	Yes	Yes	Yes	Yes	Yes
Year × Quarter FE	Yes	Yes	Yes	Yes	Yes
Day-of-week Controls	Yes	Yes	Yes	Yes	Yes
Month-of-year Controls	Yes	Yes	Yes	Yes	Yes
Hour-of-day Controls	Yes	Yes	Yes	Yes	Yes
News Type Controls	Yes	Yes	Yes	Yes	Yes
# clusters (date)	463	463	463	463	463
Adj. R^2	0.050	0.019	0.020	0.008	0.009

Note. This table reports OLS estimates of return predictability from news sentiment of articles released during any LLM outage, using articles released outside of any outages as the control. The dependent variable is DGTW-adjusted abnormal returns (AR) measured over the horizon indicated in each column header. $S_{i,t}(z)$ is the article-level event sentiment score from Dow Jones Newswires, standardized to zero mean and unit variance. O_t is an indicator equal to one on days with an LLM provider outage. Controls include lagged abnormal returns, log book-to-market, log market equity, and log share turnover. $\hat{\beta}_1 + \hat{\beta}_2$ is the total sentiment coefficient during outages. $\hat{\phi} = \hat{\beta}_2/(\hat{\beta}_1+\hat{\beta}_2)$ measures the share of return predictability attributable to LLM-assisted trading only when $AR_{t+1:t+k}$ is the outcome; we also report it for AR_t specification for completeness. N = 4,185,598. Coefficients are in basis points. t-statistics based on standard errors clustered by date appear in brackets. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table 4: Return Predictability During Severe Outages

	AR_t	AR_{t+1}	$AR_{t+1:t+2}$	$AR_{t+1:t+3}$	$AR_{t+1:t+4}$
	(1)	(2)	(3)	(4)	(5)
$S_{i,t}(z)$	37.80***	13.47***	13.70***	12.14***	11.13***
	[4.00]	[8.34]	[7.59]	[6.27]	[5.50]
$S_{i,t}(z) \times O_t$	28.62	12.74***	10.56*	11.38*	11.97*
	[1.09]	[2.69]	[1.70]	[1.73]	[1.67]
O_t	-42.24	0.90	3.49	-7.11	-3.85
	[-0.99]	[0.09]	[0.31]	[-0.61]	[-0.30]
AR_t		-0.03***	-0.03***	-0.01***	-0.01***
		[-7.16]	[-7.79]	[-3.99]	[-3.91]
AR_{t-1}	0.01	0.01	0.03**	0.03***	0.04***
	[0.08]	[0.97]	[2.33]	[2.68]	[2.66]
AR_{t-2}	0.11	0.01	0.01	0.01	0.01
	[0.87]	[0.96]	[0.68]	[1.29]	[1.15]
$AR_{t-21:t-3}$	-0.30	-0.01	-0.01	-0.02**	-0.02***
	[-1.55]	[-1.45]	[-1.63]	[-2.18]	[-2.88]
$AR_{t-252:t-22}$	-0.22	-0.01	-0.01**	-0.01**	-0.01***
	[-1.57]	[-1.63]	[-1.98]	[-2.12]	[-2.95]
log(BM)	80.31	38.84***	57.75***	74.22***	80.20***
	[1.25]	[2.69]	[3.46]	[4.06]	[3.95]
log(ME)	363.67***	-76.02***	-94.84***	-105.99***	-119.75***
	[2.74]	[-8.40]	[-8.83]	[-8.76]	[-9.05]
log(Turnover)	372.10***	12.60	14.68	2.99	5.08
	[4.59]	[1.55]	[1.61]	[0.30]	[0.48]
$\hat{\beta}_1 + \hat{\beta}_2$	66.42***	26.21***	24.26***	23.53***	23.10***
	[3.16]	[6.18]	[4.34]	[3.93]	[3.47]
$\hat{\phi}$	43.1	48.6***	43.5***	48.4***	51.8***
	[1.60]	[4.42]	[2.68]	[2.93]	[3.00]
Firm FE	Yes	Yes	Yes	Yes	Yes
Year × Quarter FE	Yes	Yes	Yes	Yes	Yes
Day-of-week Controls	Yes	Yes	Yes	Yes	Yes
Month-of-year Controls	Yes	Yes	Yes	Yes	Yes
Hour-of-day Controls	Yes	Yes	Yes	Yes	Yes
News Type Controls	Yes	Yes	Yes	Yes	Yes
# clusters (date)	443	443	443	443	443
Adj. R^2	0.050	0.019	0.020	0.008	0.010

Note. This table reports OLS estimates of return predictability using the same specification as Table 3 but with severe LLM provider outages. N = 3,756,502. Coefficients are in basis points. t-statistics based on standard errors clustered by date appear in brackets. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table 5: Return Predictability During Multi-Provider Outages

	AR_t (1)	AR_{t+1} (2)	$AR_{t+1:t+2}$ (3)	$AR_{t+1:t+3}$ (4)	$AR_{t+1:t+4}$ (5)
$S_{i,t}(z)$	32.84***	13.28***	13.57***	11.86***	10.55***
	[2.99]	[8.69]	[7.98]	[6.32]	[5.38]
$S_{i,t}(z) \times O_t$	27.31	20.83**	-2.65	-8.37	-10.59
	[1.02]	[2.38]	[-0.12]	[-0.33]	[-0.36]
O_t	98.94	-30.56*	9.58	-25.59	-29.58
	[1.43]	[-1.78]	[0.35]	[-0.80]	[-0.71]
AR_t		-0.03***	-0.03***	-0.01*	-0.01**
		[-4.30]	[-4.05]	[-1.84]	[-2.25]
AR_{t-1}	0.02	0.02	0.04**	0.03***	0.04***
	[0.17]	[0.99]	[2.50]	[2.74]	[2.68]
AR_{t-2}	0.12	0.01	0.01	0.01	0.01
	[0.89]	[1.23]	[1.41]	[1.63]	[1.34]
$AR_{t-21:t-3}$	-0.35*	-0.01	-0.01	-0.01*	-0.02**
	[-1.72]	[-1.23]	[-1.58]	[-1.68]	[-2.38]
$AR_{t-252:t-22}$	-0.08***	-0.00	-0.01*	-0.01*	-0.02***
	[-3.51]	[-0.67]	[-1.72]	[-1.76]	[-2.75]
log(BM)	83.66	34.11**	54.60***	61.95***	64.83***
	[1.45]	[2.25]	[3.18]	[3.33]	[3.06]
log(ME)	362.02***	-77.48***	-92.64***	-108.53***	-125.16***
	[3.09]	[-7.55]	[-7.73]	[-7.76]	[-7.96]
log(Turnover)	386.28***	12.24	13.44	2.74	4.20
	[4.34]	[1.32]	[1.30]	[0.24]	[0.35]
$\hat{\beta}_1 + \hat{\beta}_2$	60.15**	34.11***	10.93	3.49	-0.04
	[2.33]	[3.95]	[0.49]	[0.14]	[-0.00]
$\hat{\phi}$	45.4	61.1***	-24.2	-239.6	26488.2
	[1.63]	[5.67]	[-0.09]	[-0.10]	[0.00]
Firm FE	Yes	Yes	Yes	Yes	Yes
Year $\times$ Quarter FE	Yes	Yes	Yes	Yes	Yes
Day-of-week Controls	Yes	Yes	Yes	Yes	Yes
Month-of-year Controls	Yes	Yes	Yes	Yes	Yes
Hour-of-day Controls	Yes	Yes	Yes	Yes	Yes
News Type Controls	Yes	Yes	Yes	Yes	Yes
# clusters (date)	397	397	397	397	397
Adj. R^2	0.047	0.020	0.020	0.008	0.009

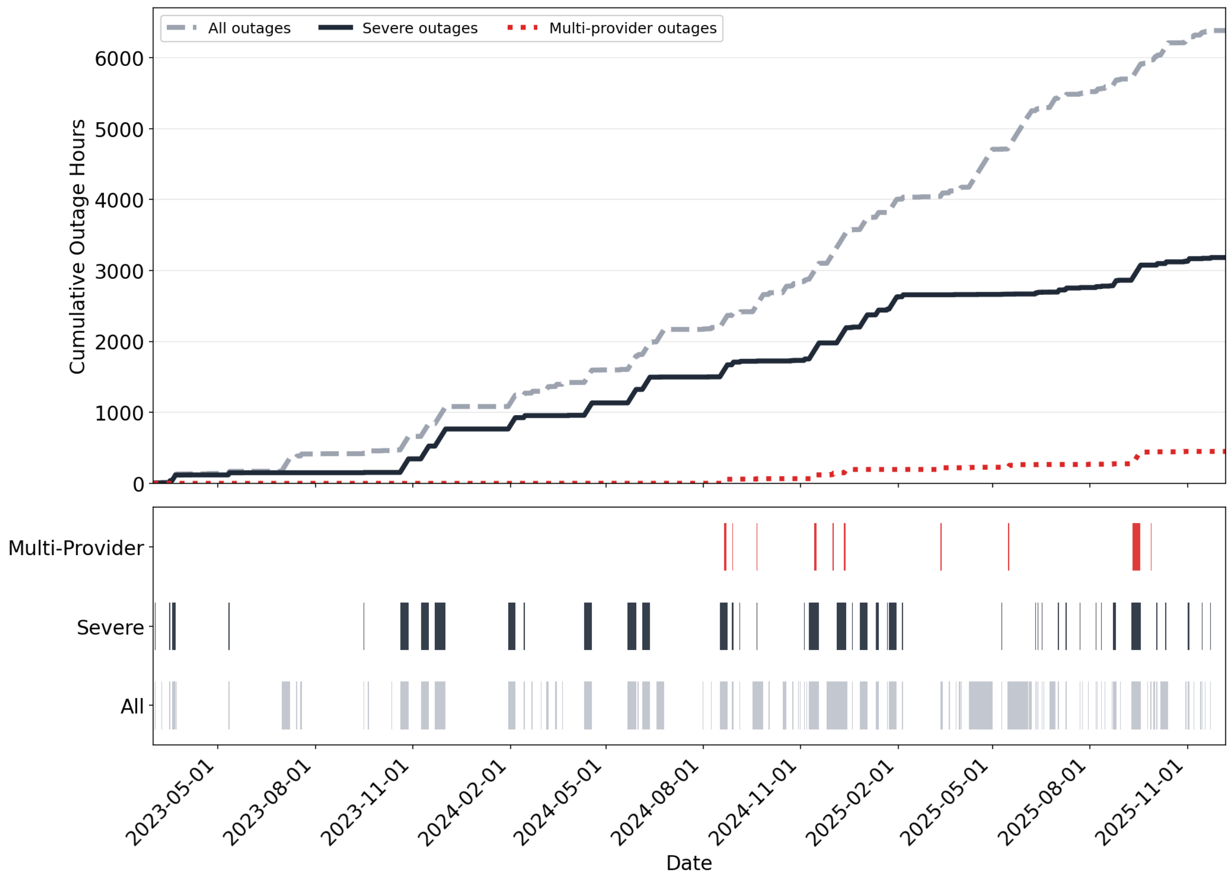
Note. This table reports OLS estimates of return predictability using the same specification as Table 3 but with multiple LLM provider outages. N = 3,179,438. Coefficients are in basis points. t-statistics based on standard errors clustered by date appear in brackets. * $p < .1$; ** $p < .05$; *** $p < .01$.

Internet Appendix:

Do LLMs make markets more efficient?

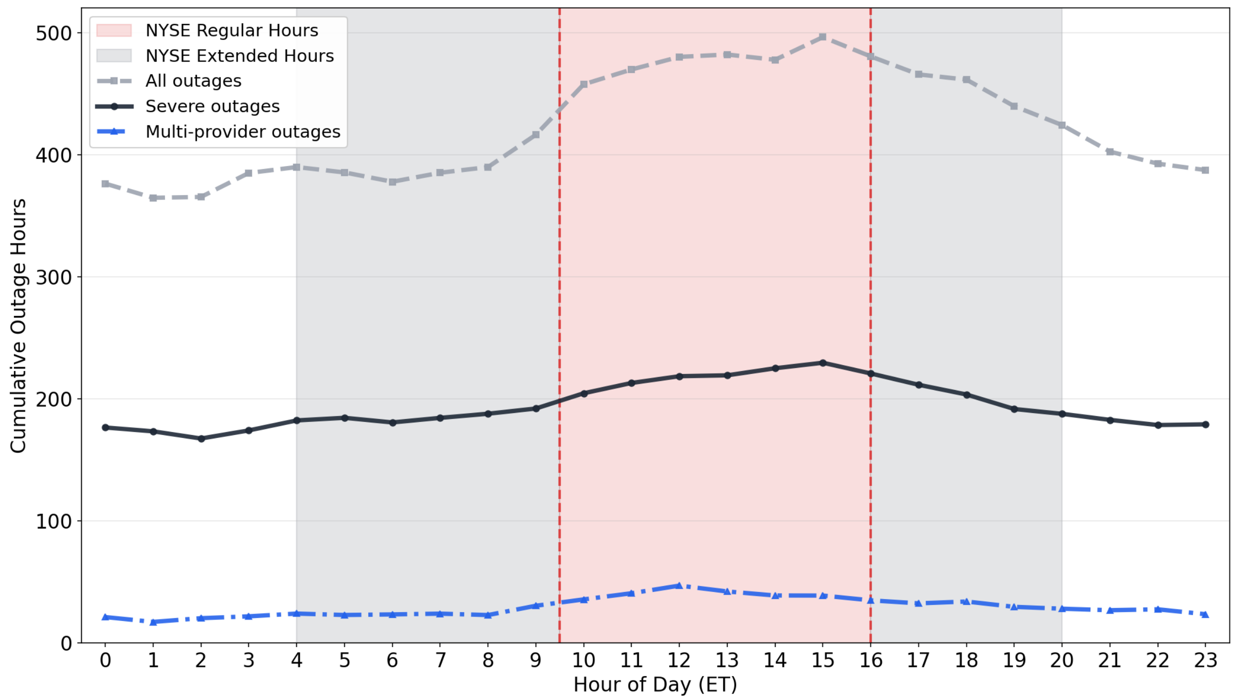
by Runjing Lu, Luka Vulicevic, and Yongxin Xu

Figure A1: Time Series of LLM Outages by Outage Type



Note: This figure repeats Figure 2, separately by outage type. See note to Table 1 for outage type definitions.

Figure A2: Intraday Distribution by LLM Outages



Note: This figure shows the intraday distribution of outage exposure by plotting cumulative outage hours by hour of day (Eastern Time). Outage duration is aggregated into hourly bins based on the clock time of occurrence. The gray dashed line represents any outages, the black solid line severe outages, and the blue dash-dot line simultaneous outages at two or more providers. Shaded regions indicate NYSE trading hours: regular session (9:30 AM-4:00 PM, light red) and extended hours (light gray). Vertical dashed lines mark the market open and close.

Table A1: Outage Seasonality

Panel A: Outage rate (%) by day of week

	Mon	Tue	Wed	Thu	Fri
All	23.2	25.5	33.8	32.1	31.0
Severe	14.3	15.4	20.8	18.2	16.3
Multi-provider	1.7	0.7	2.2	3.9	3.8

Panel B: Outage rate (%) by month of year

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
All	26.2	17.6	25.3	37.3	25.9	38.8	15.5	23.2	33.1	36.4	48.8	41.9
Severe	26.1	14.4	8.0	11.7	11.8	16.5	4.9	14.4	17.0	21.0	43.9	35.2
Multi-provider	0.0	0.0	2.4	0.8	2.2	0.3	0.9	4.0	8.3	0.0	6.3	7.9

Note. This table reports the outage rate across outage types (%) by day of week (Panel A) and month of year (Panel B). See note to Table 1 for outage type definitions.

Table A2: Dynamic $S_{i,t} \times O_t$ Coefficients by Outage Definition

	AR_{t-4}	AR_{t-3}	AR_{t-2}	AR_{t-1}
All	1.23	2.63	-1.12	-2.14
	[0.53]	[1.05]	[-0.44]	[-1.08]
Severe	0.94	4.66	-4.59	-2.66
	[0.31]	[1.41]	[-1.50]	[-1.11]
Multi-provider	-3.31	-5.14	1.74	4.40
	[-0.75]	[-0.94]	[0.27]	[0.84]

Note. This table reports dynamic coefficients on the interaction $S_{i,t} \times O_t$ from event-study regressions estimated separately for each horizon. Each column corresponds to the dependent variable AR_{t+k}, where k denotes event time relative to the outage date t (with $k \in [-4, -1]$). Rows differ by the outage definition used to construct O_t.See note to Table 1 for outage type definitions.