
Forecasting Capital Markets Day Targets with Language Models

What an informed reader can anticipate, and whether the market already knows

Alistair Smallwood, Head of Applied AI at Primer

ABSTRACT

Before a capital markets day, experienced investors form a view of what management will announce: which targets, roughly where they will land and whether they will be raised. That view rarely predicts the outcome exactly, but it narrows the range of plausible outcomes and makes a genuine surprise easier to recognise. This paper asks whether a language model reading the same documents can form that view, and whether the view is worth anything once the market has seen it.

We test Claude Opus 4.5 on 768 capital markets and investor days held after its training data ends, using only filings, presentations and transcripts published before each event, with no web access. We score its forecasts against the announced targets, simple rules and sell-side consensus estimates, measure the share-price and trading-volume reaction for the 553 events with price data, and test whether it can also foresee the other announcements that move prices. Analysis plans were fixed before results were seen.

The model names the company's headline target 73% of the time, against at most 39% for repeating the last target's measure. Its median level error is 0.9 percentage points, about 1.1 points closer to the announced target than consensus, and it anticipates much of the other news, such as buybacks and cost programmes. It is weaker on direction, horizon and form, overconfident and persistently optimistic about the market's response. Capital markets days move prices, with two-day volatility 2.7 times normal, mostly on near-term guidance and capital returns rather than targets, but confident forecasts from the language model carry no alpha.

We interpret the model's forecasts as reproducing what the market already knows, which is why its highly confident forecasts carry no alpha. Their use is as a gauge, for human or agent investors, of whether a capital markets day announcement is a genuine surprise. Consistent with this, trading volume was higher when the model had not anticipated the target outcome (82% above normal when it was unsure and 69% when its confident call was wrong, against 55% when it was right), although these differences, and a broader test across all kinds of announcement, are not yet statistically conclusive.

1 Introduction

Capital markets days are among the few occasions when management sets out its ambitions in numbers. A company might commit to an operating margin of 18% by 2028, a revenue growth range or a return on capital, and explain how it intends to get there. Investors prepare for these events by reading the previous plan, recent results commentary, conference remarks and management's hints, and they form a view of what is coming.

That view is valuable even when it is imprecise. Knowing that a margin target is likely to rise by about a point, and that the horizon will probably move out two years, tells an investor what would count as a surprise. An analyst can prepare properly for a handful of events a season; a portfolio may follow hundreds of companies.

Language models can now read the same documents. If a model given only pre-event material reproduces the informed in-

vestor's view, it can provide that preparation for every company at negligible cost. If it sees something the market does not, it might also have trading value. Both claims are easy to make and hard to test, because a model may have read about the event during training, so a retrospective test can reward memory rather than reasoning (Glasserman and Lin 2023; Sarkar and Vafa 2024). Earlier work on language models and returns has mostly read news after it is published (Lopez-Lira and Tang 2023); here the model must anticipate an announcement before it happens.

We avoid that problem by testing a model only on events that took place after its training data ends, and by checking that it cannot recall them. The paper reports how well the model anticipates the targets, how it compares with sell-side consensus, whether its view is reflected in the market's reaction and which other announcements move prices.

2 Research questions and scope

Anticipation: from pre-event documents alone, can a model predict which targets a company will announce, where they will land, whether they will be raised, kept or cut, and their horizon and form, better than simple rules? **Against the sell side:** are its forecasts closer to the announced targets than

consensus? **Market value:** do its confident forecasts predict the share-price reaction, or behave like the market's existing expectation? **What moves prices:** which announcements drive the reaction, and could the model have foreseen them?

3 Out-of-sample design

Events. We found candidate events by sweeping the titles of presentations and transcripts in Primer's document store for capital markets days, investor days and strategy updates, then screening each for a numeric headline target. After removing events whose targets had already been published before the event and events with empty evidence packs, 768 events remain, from September 2025 to October 2026.

Training boundary. Claude Opus 4.5's training data runs to August 2025, so every event is outside it. Before forecasting, the model was asked what each company announced at the event; none of the events was recalled.

Evidence pack. For each event the model receives up to 110,000 tokens of filings, presentations and transcripts published before the day before the event, assembled automatically and audited for publication dates. It runs with no tools and no web search, and returns a fixed answer format.

Tasks. In the open task the model is told only the company and the event date, and lists the targets it expects, with probabilities. In the measure-given task it is also told the headline measure. Each task runs with and without the evidence pack, which separates reading from general knowledge.

Why dated primary documents matter. The design depends on knowing exactly what was public before each event.

Primer's document tools search company filings, presentations and transcripts by company, document type and publication date, returning titles and timestamps without opening the documents, and they also support full-text search with BM25 ranking across the whole corpus. Finding the events needed only titles: a sweep of 116,948 deck and transcript titles, after which only the matching documents were read. The same filters froze a point-in-time evidence pack for every event and retrieved each event's own materials for the answer key and the release time; the study made about 25,000 document requests in all. Full-text search was not needed for this design, but it lets the same approach find specific passages, such as every earlier statement of a target, across thousands of documents without processing each one. A model with live web search could not be tested this way, because the web cannot be frozen at a date: reporting of the outcome leaks into the inputs (Glasserman and Lin 2023; Sarkar and Vafa 2024).

Answer key and scoring. Announced targets were extracted from each event's own materials in two independent passes and reconciled, and each target's definition was checked against the previous one. A matcher blind to values decides whether the model named the headline target. A raise, keep or cut is scored against the previous target with tolerance bands of half a point for margins, growth and returns, and one point for cost-income ratios and EPS growth.

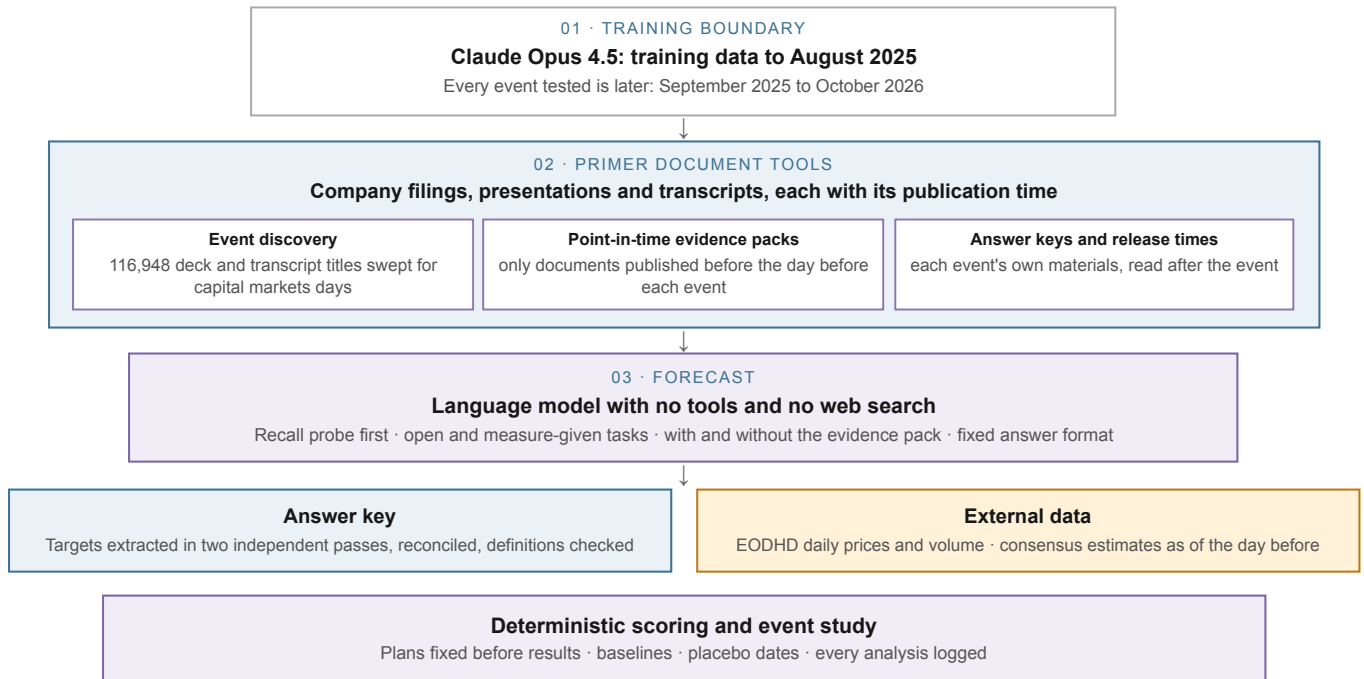


Figure 1. Architecture of the test. Primer's document tools supply every company document with its publication time, which makes it possible to find the events, freeze what was public before each one and read what was announced. The forecasting model sees only the frozen pack; scoring and the event study run in code on separately assembled data.

4 What the model anticipates

The model names the headline target in 73% of events (70% to 76%). The best a rule could do by repeating the company's last target measure is 39%: the share of events whose headline had a same-measure predecessor. The gap is +34 points (+30 to +38). Without the evidence pack the model still names the headline 67% of the time, so most of this skill reflects general knowledge of what companies in each industry target. The documents add the rest and, more importantly, the level.

Its median level error on the headline target is 0.9 points (mean 1.8). Without documents the mean error is 3.4 points. Against the strongest simple baseline, assuming no change from the previous target, the model is 0.29 points closer on average (0.09–0.52; 251 events). That is a modest edge on a demanding baseline, because targets usually change only a little.

Direction is harder. Of headline targets with a comparable predecessor, 60% were kept, 29% raised and 11% cut. The model

calls the direction correctly 64% of the time in the open task, against 60% for always predicting no change (+4.4 points; –2.6 to +11.3). When told the measure it reaches 66% against 58% (+8.0 points; +1.5 to +14.8). Without documents it falls to 43%–46%.

The model is weakest on presentation choices. It gets the horizon year right 43%–48% of the time and the form (floor, range or single figure) 49%–51%. Its probabilities are overconfident: targets rated 0.7 to 0.9 likely were presented 57% of the time. A newer model does better. On 62 events forecast by both models, Opus 5.5 called the direction correctly 71% of the time against 58% for Opus 4.5, with a mean level error of 1.3 points against 3.0. Some of those events precede the newer model's training cutoff, although they passed recall checks, so this comparison is indicative.

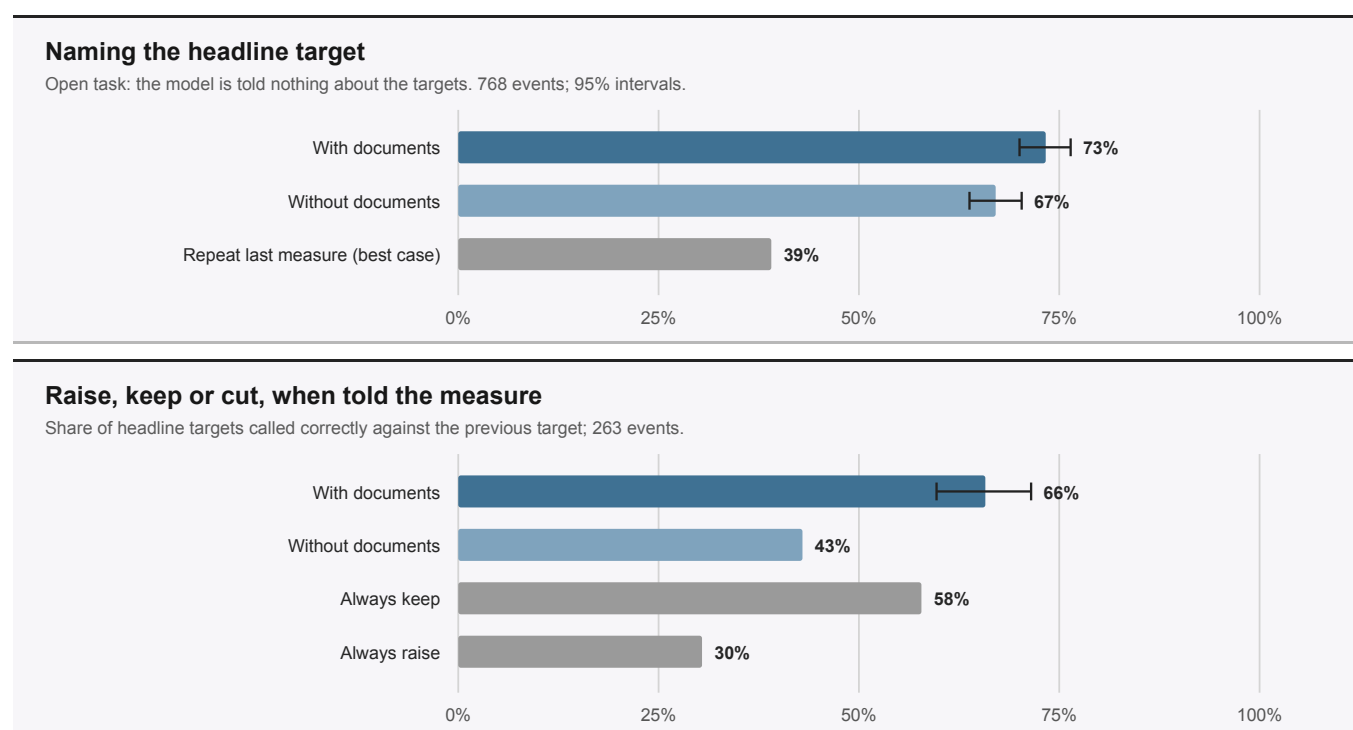


Figure 2. The model names the company's headline target far more often than a rule could, even its best case, but its raise/keep/cut calls only modestly beat always predicting no change.

Forecasting result (768 events)	With documents	Without	Baseline
Headline target named (open task)	73%	67%	Repeat last measure $\leq$ 39%
Share of all announced targets anticipated	53%	46%	–
Level error, median / mean (pp)	0.9 / 1.8	– / 3.4	No change: 0.29pp worse
Raise, keep or cut (open task)	64%	46%	Always keep 60%
Raise, keep or cut (measure given)	66%	43%	Always keep 58%
Horizon year right	48%	27%	–
Form right (floor, range, figure)	51%	40%	–

Table 1. Out-of-sample forecasting results for Claude Opus 4.5. Open task unless stated; events are the unit.

5 Against the sell side

Sell-side consensus is the obvious benchmark for an informed reader. We took sell-side consensus estimates as of the day before each event and converted them into the announced measure and horizon: margin as consensus profit over consensus revenue for the target year, and growth as the compound rate to the target year. This is possible for 145 events with prices. Many others fall away because consensus does not reach the target year or, for banks and insurers, comparable consensus lines are not available. These comparisons are exploratory.

Companies set targets above consensus in two-thirds of cases, by a median of about one point. That is unsurprising: consensus forecasts what a company will deliver, while a target states what management intends.

On the same 121 events, the model is closer to the announced target than consensus by 1.1 points on average (0.6–1.6), with a median miss of 1.0 against 1.4. Consensus is further from the target than simply assuming the previous target (1.2 points; 0.4–2.1; 61 events). In a regression of the announced target on all three forecasts, the model's forecast carries the weight (coefficient 0.86, t 2.9) and consensus adds nothing (0.03).

The model and consensus rank companies almost identically (rank correlation 0.95), but only 38% of forecasts sit within a point of each other. The model's view is a reading of management, not an echo of the sell side.

68% of announced targets sit above consensus for the same year (median about one point)	1.1pp closer to the announced target than consensus, on average (0.6–1.6pp)	0.03 weight on consensus once the model's forecast is known (model 0.86)
---	---	--

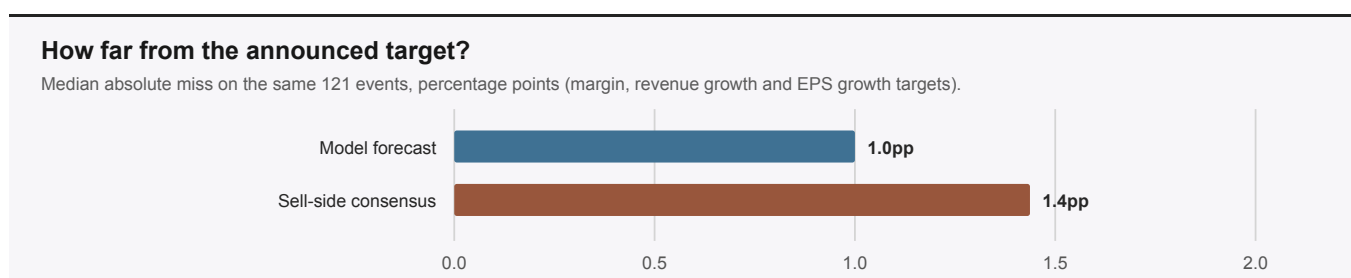


Figure 3. Consensus forecasts what a company will deliver; the model forecasts what management will say. Targets sit above consensus in most cases, and the model is about 1.1 points closer on average.

6 Does the market already know?

Design. We follow short-window event-study practice, which is well specified with daily data where long-horizon tests are not (Brown and Warner 1985; Kothari and Warner 2007). Prices are EODHD daily data, the source behind Primer's market-data tools. Abnormal returns come from a market model on each stock's local index, as recommended for multi-country samples (Campbell, Cowan and Salotti 2010), estimated over 250 trading days ending 11 days before the event. The main window is the event day and the next day: the usual choice when an announcement may come after the close and be priced only the next day (MacKinlay 1997). We also report windows from the day before to six days after, and the same tests on placebo dates 40 days later (Brown and Warner 1985).

We measure the reaction in two ways, both introduced by Beaver (1968) to test whether an announcement carries information: the variance of abnormal returns, which registers news whichever way it points, and abnormal trading volume (also Landsman and Maydew 2002; for investor days, Kirk and Markov 2016). Mean returns are tested with the standardised cross-sectional test of Boehmer, Musumeci and Poulsen (1991) adjusted for cross-correlation (Kolari and Pynnönen 2010), the generalised rank test (Corrado 1989; Kolari and Pynnönen 2011) and the generalised sign test (Cowan 1992). Of the 553 events with prices, 70 had results released within a day of the event and are excluded from the main sample of 483. Tokyo, Milan and India are not covered by EODHD, which removes 201 events.

Capital markets days are news. Two-day abnormal volatility is 2.7 times normal (2.3–3.1), against 1.1 on placebo

dates. The median two-day move is 2.8% against 1.5% on placebo dates; 27% of events move more than 5% (11% on placebo dates) and 6% more than 10%. Volume is 67% above normal. About 77% of the excess reaction falls on the event day and the rest on the next day; nothing unusual happens beforehand. Most first presentations or releases appear before the open (44%) or during trading (45%); only 4% appear the evening before.

No alpha in the model's confident calls. Events where the model confidently expected a raise (probability at least 75%; 94 events) returned -0.55 percentage points relative to the other 389 (95% interval -1.8 to $+0.7$). The slope of the two-day return on the model's net view, its probability of a raise minus its probability of a cut, is $+0.07$ points per unit (t 0.1). The result holds with simple market-adjusted returns (-0.58 ; -1.8 to $+0.7$), with all events including results days, and in every window from the event day alone to six days.

The target is not what moves the price. Even the actual outcome explains little. Raised targets averaged -0.04% over two days, kept targets -0.28% and cut targets -1.22% (raise minus cut $+1.2$ points; -0.9 to $+3.4$). Following Fried and Givoly (1982), who judged which forecast best represents the market's expectation by its link with the price reaction, we compared three yardsticks. Measured against the previous target, consensus or the model's forecast, a target above the yardstick coincided with a rising share price in 46%, 50% and 51% of cases: no better than a coin toss.

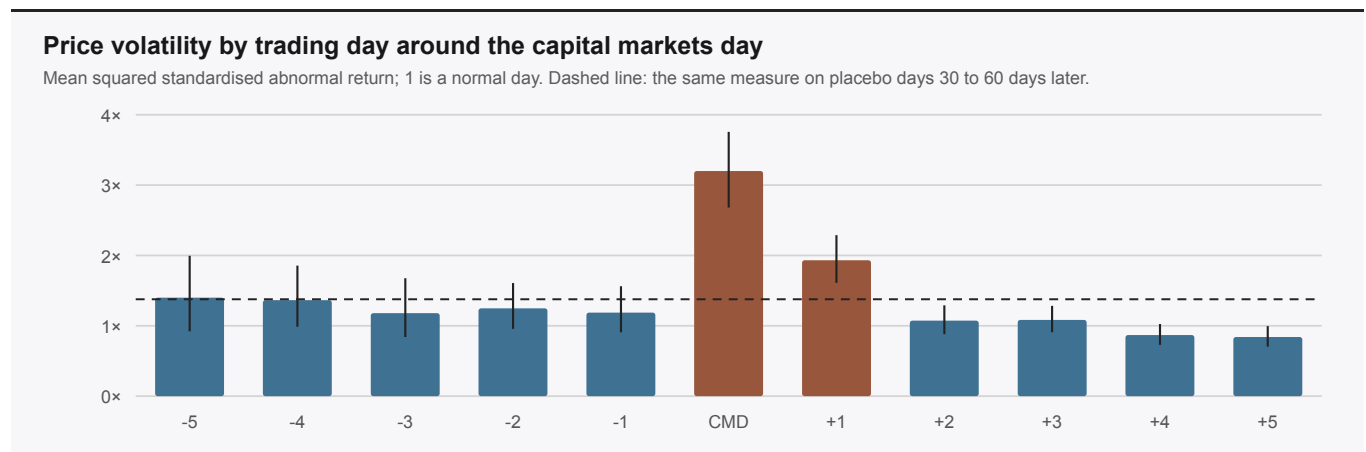


Figure 4. Volatility jumps on the event day and partly the next, with no build-up beforehand. About 77% of the excess two-day reaction falls on the event day.

Two-day abnormal return by the model's pre-event call

CMD day and the next day; market model on the local index; 95% intervals.

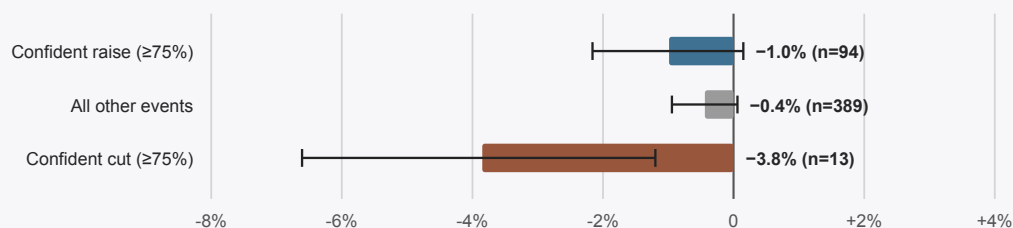


Figure 5. Confident raise calls did no better than other events (-0.55 points, -1.8 to $+0.7$). The handful of confident cut calls fell, a lead for the prospective test rather than a finding.

Market reaction (main sample, 483 events)	Result	Comparison or 95% interval
Two-day abnormal volatility vs normal	2.7×	placebo 1.1×
Median two-day move, ignoring sign	2.8%	placebo 1.5%
Events moving more than 5%	27%	placebo 11%
Trading volume vs normal	+67%	—
Confident raise minus other events	-0.55 pts	-1.8 to $+0.7$
Slope on the model's net view (per unit)	$+0.07$ pts	$t = 0.1$
Actual raise minus actual cut	$+1.2$ pts	-0.9 to $+3.4$
Target above expectation, shares up	46% / 50% / 51%	previous / consensus / model

Table 2. Event-study results. Main sample excludes events with results released within a day. Mean two-day abnormal return across all events: -0.55 points.

7 Expectations, volume and leads

If the model's forecast is the market's expectation, the market should react only to the gap between the outcome and that forecast. In a joint regression the two-day return rises with the actual direction ($+0.7$ points per step from cut to keep to raise) and falls with the model's expected direction (-1.0 points). That is the expected pattern, but neither coefficient is significant.

Trading volume gives a cleaner reading of whether something was news, because it rises when investors interpret the same announcement differently, even when good and bad news offset in the price (Kim and Verrecchia 1991; Kandel and Pearson 1995). Volume rose 55% when the model's confident call came true, 69% when a confident call was wrong and 82% when the model was unsure. The ordering is the one an expectations reading predicts, but the differences are within sampling error,

and only 30 events had confident wrong calls. The price measure shows no such pattern: two-day volatility is 2.9 times normal after confident correct calls and 2.5 after confident wrong ones.

Three further results are leads rather than findings. The 13 confident cut calls fell 3.8% over two days (-6.6 to -1.2), consistent with evidence that bad news is priced slowly (Hong, Lim and Stein 2000), but the sample is small. Among smaller companies, confident raises beat other events by $+1.1$ points, while among larger ones they lagged by 1.9 points; the difference between the halves is not significant ($t = 1.3$). Across markets there is no sign of the pre-announcement leakage documented for takeovers (Keown and Pinkerton 1981): pre-event volatility is 1.1 times normal, and pre-event drift does not anticipate the event-day move ($+0.5$ points; -0.3 to $+1.3$).

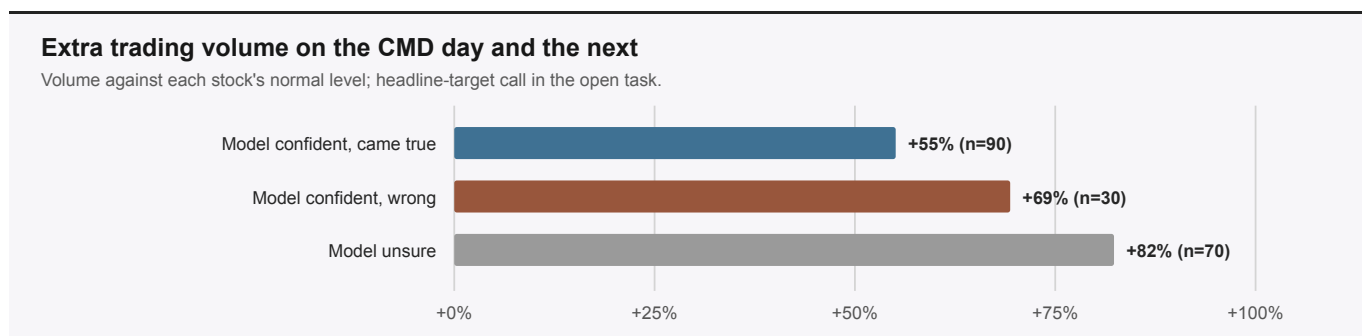


Figure 6. Less trading when the model's confident call came true, more when it was unsure: the ordering an expectations reading predicts, but the gaps are within sampling error.

8 What capital markets days move prices on

If the headline target is not what moves the price, something else announced on the day must be. To find out what, Claude Opus 5.5 read each event's own materials, its presentations, releases and transcripts, and listed every announcement by type: structural changes, capital returns, near-term guidance, medium-term targets, cost programmes, investment plans, segment reporting, management changes and financing. For each item it recorded whether it was new at the event, how material it was and whether it was good or bad for shareholders on its face, with a verbatim quote; 94% of 5,943 quotes match the source text exactly. This step reads outcomes rather than forecasting them, so the training boundary does not apply. The analysis plan was fixed before any tags existed.

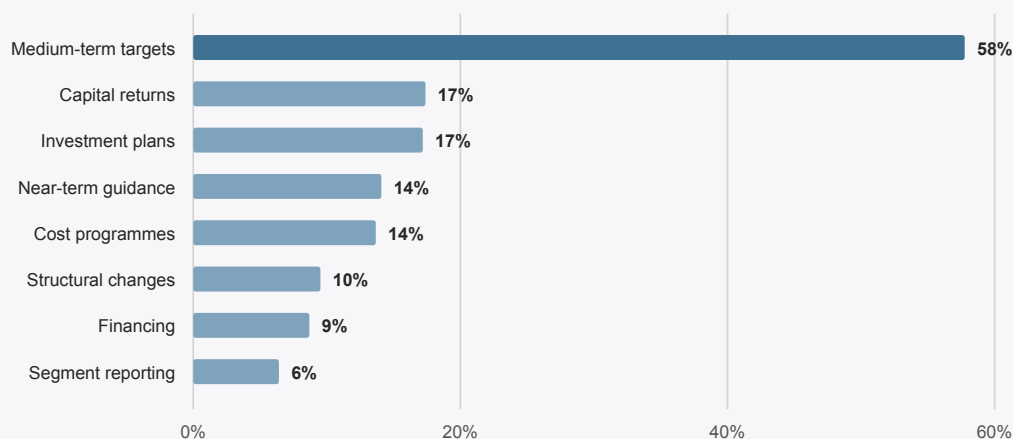
Targets dominate what capital markets days announce. 58% of events newly announce a material target, and targets are the headline announcement in 77%. But many events carry other news: 17% announce capital returns, 17% investment plans, 14% change near-term guidance, 14% launch cost programmes and 10% announce structural changes such as disposals, acquisitions or separations.

Near-term guidance is the clearest driver of the price. Events with new guidance changes moved more, with two-day volatility 3.9 times normal against 2.5 (95% interval for the difference +0.1 to +2.9), and guidance cuts fell 3.6% over two days (−6.8 to −0.7; 10 events). Including events with same-day results, the contrast is sharper: 7.0 against 2.4 times normal, with cuts down 4.7%. Capital returns pushed prices up (+1.2 points per unit of signed news, t 2.6). New targets raised trading volume by about 13% (t 2.3) but did not move the price either way, and structural announcements did not come with larger reactions. New deleveraging targets coincided with falls (−2.5%), which more likely reflects the stretched balance sheets of the companies that set them than the news itself.

Taken together, the tags explain about 4% of the direction of the two-day move and 4%–11% of its size. The market trades the near term, guidance and cash returns, rather than the multi-year ambition. Those are the announcements a forecaster would need to anticipate.

What capital markets days newly announce

Share of 483 events with a new, material announcement of each type.



Two-day abnormal return by kind of news

Events newly announcing each kind of news; 95% intervals.

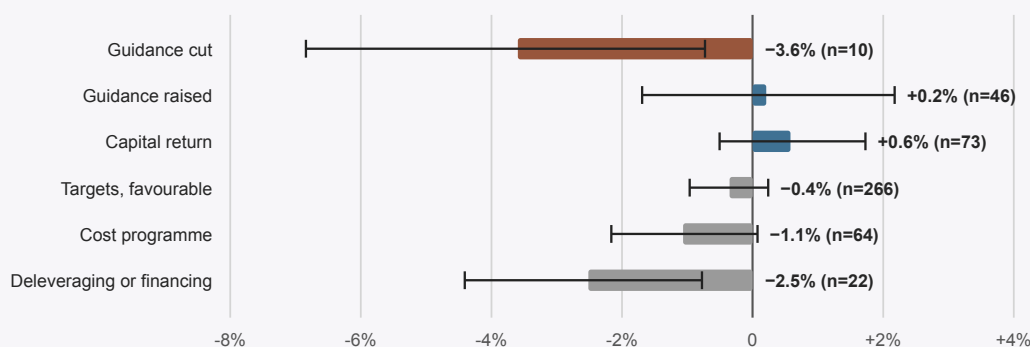


Figure 7. Targets are announced most often, but the price follows near-term news: guidance cuts fall and capital returns rise. Favourable target news, on its own, leaves the price unchanged on average.

9 Could the model see them coming?

If guidance and capital returns move prices, the useful question is whether a model could have seen them coming. We asked the same out-of-sample model, with the same evidence packs and no web access, for the probability that each event would newly announce each of twelve kinds of news, and which announcement would move the share price most, and in which direction. The answer key is the tagging in Section 8. The plan, including the tests and our predictions, was recorded before any forecast was made.

The model sees much of the news coming. Its probabilities rank events correctly about three times in four for cost programmes (AUC 0.78), capital returns (0.77) and new targets (0.77), and about seven times in ten for structural changes (0.72) and guidance raises (0.70). Guidance cuts, the most price-relevant category, are the exception: with only 12 in the main sample, an AUC of 0.62 (0.47–0.76) cannot be distinguished from chance (0.68 across all events, 19 cuts). Its probabilities are again too high: events it rated 50% to 70% likely happened 39% of the time.

None of this foresight carries alpha. The model's advance probabilities of a guidance cut, a guidance raise or a capital return do not predict the two-day return (the largest t-statistic is 1.0). Asked which announcement would move the price most, it expected a rise 97% of the time; the shares rose in 45% of cases (41%–51%). The model is a persistent optimist about how the market will receive company news.

Measuring surprise against the model does not, in aggregate, explain the price better than the news itself (R^2 0.019 against 0.021). For capital returns the pattern points the expected way, with announced returns adding 1.1 points (t 2.4) and expected ones subtracting 1.0 (t -0.4), but the second figure is far from significant. Including events with same-day results, trading volume rises with the amount of news (t 2.7) and with the size of the surprise against the model (t 2.3). The model anticipates the market-moving news about as well as the market appears to; what it cannot do is say how the market will take it.

How well the model foresees each kind of announcement

Area under the ROC curve for its advance probabilities: 0.5 is chance, 1.0 is perfect. 483 events; 95% intervals.

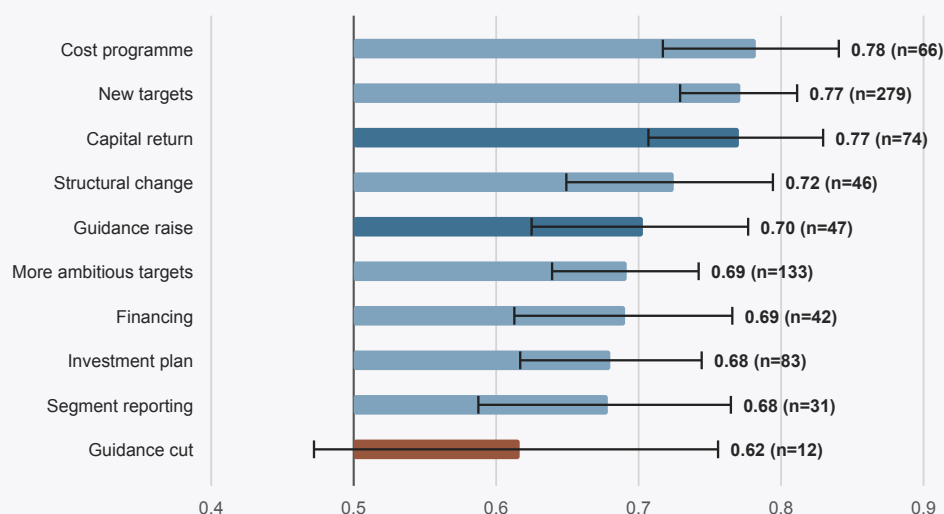


Figure 8. The model anticipates most kinds of capital markets day news well, including the capital returns that move prices. Guidance cuts, the other price-moving category, are rare and not yet distinguishable from chance.

10 Example 1: Pentair, anticipated and priced

COMPANY	EVENT	MODEL FORECAST BEFORE THE EVENT
Pentair	Investor day, 4 March 2026	Return-on-sales target raised to about 28% (95% probability of a raise)

INVESTOR INTERPRETATION

The model forecast both the raise and the level, and the market appears to have reached the same view. The shares moved -1.8% over two days, less than the median capital markets day move of 2.8% , on volume 56% above normal, close to the typical event.

Pentair had set a 24% return-on-sales target for 2026 at its previous investor day and was already guiding to 26% . Reading the evidence pack, the model inferred a new three-year target of $27\text{--}28.5\%$ and gave a 95% probability of a raise. At the event, management set a 2026–2028 target of more than 28% return on sales, alongside mid-single-digit revenue growth. Asked separately in advance which announcements to expect (Section 9), it put the chance of more ambitious targets at 70% .

What the model did. Its reasoning, recorded before the event: “Prior 24% target for 2026 already exceeded at 25.2% . 2026 guidance is 26% . Management confirmed transformation continues. New 3-year horizon suggests 28% target reflecting continued sourcing, operational, and organizational excellence initiatives.”

Boundary of the finding. A single event cannot show that a target was priced; the reaction also reflects everything else said on the day.

11 Example 2: Valeo, the guidance moved the price

COMPANY	EVENT	MODEL FORECAST BEFORE THE EVENT
Valeo	Capital markets day, 20 November 2025	2028 operating margin target of 5.5–7%

INVESTOR INTERPRETATION

The model anticipated the headline target almost exactly, yet the shares fell 12.7% over two days on volume 99% above normal. The news was the near-term outlook: sales in 2026 would be broadly flat, with growth returning only from 2027.

Under its Elevate 2028 plan Valeo targeted 2028 sales of €22–24bn, an operating margin of 6–7% and free cash flow after interest above €500m, and aimed to regain an investment-grade rating. The model had forecast a margin target of 5.5–7% for the same year. The same materials guided to flat sales in 2026, which the tagging records as a new negative guidance item. Asked separately in advance which announcements to expect (Section 9), it put the chance of a guidance cut at 5%.

What the model did. Its reasoning, recorded before the event: “Operating margin is Valeo’s headline KPI. With restructuring savings of €300M by 2026, lowered break-even, and better-priced order book, a target around 6% by 2028 is achievable vs 4.5–5.5% in 2025.”

Boundary of the finding. Attributing the fall to the 2026 outlook rests on the timing and the tags. Other content on the day may also have mattered.

12 Example 3: Restaurant Brands, the news was the buyback

COMPANY	EVENT	MODEL FORECAST BEFORE THE EVENT
Restaurant Brands International	Investor day, 26 February 2026	Comparable-sales growth target kept at 3% (80% probability)

INVESTOR INTERPRETATION

The model was right that the long-term growth algorithm would be kept, and the shares still rose 5.5% over two days on volume 165% above normal. The news was capital returns: RBI restarted share buybacks, expecting about \$500m in 2026, and planned more than \$1.6bn of total shareholder returns.

The investor day kept the growth algorithm, including about 3% comparable-sales growth, and set out free cash flow rising from about \$1.6bn in 2025 to more than \$2bn a year by 2028. RBI repeated the plan in a second presentation on 31 March 2026. With nothing new to absorb, the shares moved –0.3% over two days on volume 49% above normal. Asked separately in advance which announcements to expect (Section 9), it put the chance of a new or larger buyback at 22%.

What the model did. Its reasoning, recorded before the event: “Core algorithm component. Achieved 2.4% in 2025 (slightly below), but outperformed industry across largest segments. Management expects similar consumer backdrop in 2026.”

Boundary of the finding. The second presentation shows the same news drawing no reaction once known. It also shows that the event set includes some repeated presentations, which the tagging cannot distinguish from first announcements.

13 Limitations

Retrospective design. Every event postdates the model's training data, but the method was developed on earlier events, and several market analyses here were added after the main tests. Those are labelled exploratory.

One year, one model. The events cover roughly one year, and the market tests use one model generation. Events cluster in calendar time, which the tests address by adjusting for cross-correlation and clustering by week.

Coverage. EODHD does not cover Tokyo, Milan or India, so 201 events, mostly Japanese medium-term plans, are absent from the market tests. Japan is also the market where plans most often appear in the press before release. US stocks are measured against the S&P 500, and London and Johannesburg against index-tracking funds.

Consensus. Consensus comparisons rest on 145 events, because long-dated forecasts are thin and banks and insurers

lack comparable consensus lines.

Answer keys and harness. Announced targets were extracted by a model and checked against verbatim quotes, but some errors will remain. The command-line harness adds the date and account details to each conversation; the model was instructed to ignore them, and recall probes were negative.

Announcement tags. The answer key for Sections 8 and 9 comes from a model reading each event's materials. 94% of its quotes match the source exactly, but judgements of novelty and materiality can be wrong, and a repeated presentation of an earlier plan is read as new. Guidance cuts, the most price-relevant category, number only 12 in the main sample.

Horizon of the market tests. The market tests measure reactions over days. They do not test whether targets matter for valuation over longer periods, for example through credibility once they are delivered or missed.

14 Conclusion

A language model reading only pre-event documents reproduces much of the informed investor's view of what management will announce. It names the headline target in nearly three cases out of four, typically lands within a point of the level and is closer to the announced target than sell-side consensus. Asked in advance, it also anticipates much of the other news, such as buybacks and cost programmes. It remains weak on direction, horizon and form, its probabilities need recalibration, and it is persistently optimistic about how the market will respond.

The market evidence suggests that this view is either already priced or not what the market trades on the day. Capital markets days clearly move prices, mostly on near-term guidance and capital returns, but confident forecasts from the language

model carry no alpha, and neither the target surprise nor the surprise against its wider forecasts explains much of the reaction. Trading volume was higher when the model had not anticipated the outcome, which fits reading the model as a measure of what was expected, but the evidence is not yet conclusive.

The practical value therefore lies in preparation, coverage and judging surprise: a dependable view of the likely targets for every company, which narrows the range of outcomes and tells a human or agent investor whether an announcement is genuinely unexpected. The next step is a prospective test, fixed in advance, on upcoming events with a newer model. It will test the leads above: lower volume when confident calls come true, confident cut calls, guidance cuts and smaller companies.

R Evidence and references

1. Beaver, W. (1968). The information content of annual earnings announcements. *Journal of Accounting Research* 6, Supplement.
2. Boehmer, E., Musumeci, J. and Poulsen, A. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics* 30(2).
3. Brown, S. and Warner, J. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics* 14(1).
4. Campbell, C., Cowan, A. and Salotti, V. (2010). Multi-country event-study methods. *Journal of Banking & Finance* 34(12).
5. Corrado, C. (1989). A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23(2).
6. Cowan, A. (1992). Nonparametric event study tests. *Review of Quantitative Finance and Accounting* 2.
7. Fried, D. and Givoly, D. (1982). Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics* 4(2).
8. Glasserman, P. and Lin, C. (2023). Assessing look-ahead bias in stock return predictions generated by GPT sentiment analysis. arXiv:2309.17322.
9. Hong, H., Lim, T. and Stein, J. (2000). Bad news travels slowly: Size, analyst coverage and the profitability of momentum strategies. *Journal of Finance* 55(1).
10. Kandel, E. and Pearson, N. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103(4).
11. Keown, A. and Pinkerton, J. (1981). Merger announcements and insider trading activity: An empirical investigation. *Journal of Finance* 36(4).
12. Kim, O. and Verrecchia, R. (1991). Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29(2).
13. Kirk, M. and Markov, S. (2016). Come on over: Analyst/investor days as a disclosure medium. *The Accounting Review* 91(6).
14. Kolari, J. and Pynnönen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies* 23(11).
15. Kolari, J. and Pynnönen, S. (2011). Nonparametric rank tests for event studies. *Journal of Empirical Finance* 18(5).
16. Kothari, S. P. and Warner, J. (2007). Econometrics of event studies. In B. E. Eckbo (ed.), *Handbook of Empirical Corporate Finance*, Vol. 1. Elsevier.
17. Landsman, W. and Maydew, E. (2002). Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research* 40(3).
18. Lopez-Lira, A. and Tang, Y. (2023). Can ChatGPT forecast stock price movements? Return predictability and large language models. arXiv:2304.07619.
19. MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature* 35(1).
20. Sarkar, S. and Vafa, K. (2024). Lookahead bias in pretrained language models. Working paper.
21. Data: Primer document store (filings, presentations, transcripts); sell-side consensus estimates; EODHD end-of-day prices. Research records: projects/cmd-target-prediction (STATE.md, analysis plans and results), October 2026.

Research status. Prepared 4 October 2026. All results are retrospective; every event postdates the forecasting model's training data. Forecasting tests and the main market tests follow plans recorded before results were seen; analyses described as exploratory were added later and are labelled. The paper makes no claim of investment alpha or future stock performance. Licensed price and consensus values are not reproduced.