Do Markets Believe in Transformative AI?

Isaiah Andrews and Maryam Farboodi*

Abstract

We examine US bond yields around 30 major AI model release dates from four frontier labs between November 2022 and December 2025. Long-term Treasury and TIPS yields fall around these releases and remain lower for weeks, with the declines concentrated around releases that a crowdsourced forecast interprets as good news about the rate of AI progress. These changes are statistically significant according to a variety of small-sample valid statistical tests, and suggest that investors may take seriously the possibility of transformative impacts from AI.

1 Introduction

Since the debut of ChatGPT in November 2022, generative AI models have attracted intense interest from policymakers, researchers, and businesses. Some discussions of these models have raised the possibility that AI could lead to an increase, perhaps even a dramatic acceleration, in the rate of economic growth (Brynjolfsson et al., 2019; Trammell and Korinek, 2026; Acemoglu and Lensman, 2024; Jones, 2024; Korinek and Suh, 2024). Other discussions have suggested the possible gains from AI may be overstated, and argued that widespread AI adoption could potentially slow economic growth (Acemoglu and Restrepo, 2020). Some authors have even raised the possibility that poorly understood and controlled AI could pose an existential risk to humanity (Acemoglu and Lensman, 2024; Jones, 2024; Kokotajlo et al., 2025).

*This version: June 25, 2026. First version: July 14, 2025. Previous versions of this paper circulated with the title "Pricing Transformative AI." Andrews: MIT Department of Economics and NBER; iandrews@mit.edu. Farboodi: MIT Sloan School of Business, NBER, and CEPR; farboodi@mit.edu. We thank Marcellus Andrews, Dan Björkgren, Luigi Bocola, Markus Brunnermeier, Douglas Diamond, Mikhail Golosov, Pierre-Olivier Gourinchas, Valentin Haddad, Basil Halperin, Lars Hansen, Chad Jones, Peter Kondor, Lukasz Kowalik, Pablo Kurlat, Monika Piazzesi, Stefan Nagel, Tano Santos, Jesse Shapiro, and Cheryl Smith, as well as seminar participants at the Federal Reserve Bank of New York, Google, and MIT for very helpful discussions of and comments on this project.

It is potentially unclear to what extent the enthusiasm around AI reflects a genuine belief in its transformative potential, as opposed to belief in profit opportunities that may not translate into widespread or persistent prosperity. While the future impacts of AI are inherently unknown, understanding the beliefs of market participants is a potentially valuable input to both policy and research discussions. Financial market data has been used to infer market beliefs in other settings (see e.g. Jackwerth and Rubinstein, 1996; Wolfers and Zitzewitz, 2004; Gürkaynak et al., 2010; van Binsbergen et al., 2012, 2013), but empirical evidence about market beliefs on transformative AI is limited. In this paper, our goal is to use financial market data to provide systematic evidence regarding the beliefs of market participants about the possibility of transformative AI, by which we will mean AI technologies with a large and sustained impact on living standards, particularly through impacts on consumption or existential risk. The premise of our analysis is that if economic actors take seriously the possibility of transformative AI, this should be reflected in a wide range of forward-looking behaviors and, consequently, in long-term asset prices, including assets such as Treasury bonds which are not directly connected to AI.

That beliefs about transformative AI should affect investors' optimal choices is pointed out by e.g. Jones (2024). Chow et al. (2026) combine this observation with classic insights from consumption-based asset pricing to relate risk-free interest rates to market beliefs about transformative AI. The intuition is simple: if investors think AI will dramatically increase the rate of economic growth, then (on average across the economy) investors must expect to be richer in the future than they are today. This should decrease the marginal value of future consumption relative to present consumption, so real interest rates should rise in equilibrium. On the other hand, if investors think AI poses an existential risk, and so doubt that they will be alive in the future, this should reduce their desire to save and thus also drive up interest rates in equilibrium. Thus, consumption-based asset pricing models imply that both higher growth expectations and more concern for existential risk should increase real interest rates. Beyond expectations, uncertainty also matters: if AI increases investors' uncertainty about future consumption, this may fuel precautionary saving and so decrease real interest rates (see, e.g., Galindo Gil and Mihai, 2025).

We study the behavior of US government bond yields around major model release dates for four prominent AI labs (Anthropic, Google, OpenAI, and xAI) between November 2022 and December 2025, yielding a sample of 30 unique release dates. As shown in Figure 1, we find that Treasury yields decline around model release dates, with average declines in excess of ten basis points, or 0.1 percentage points, at some maturities. These declines

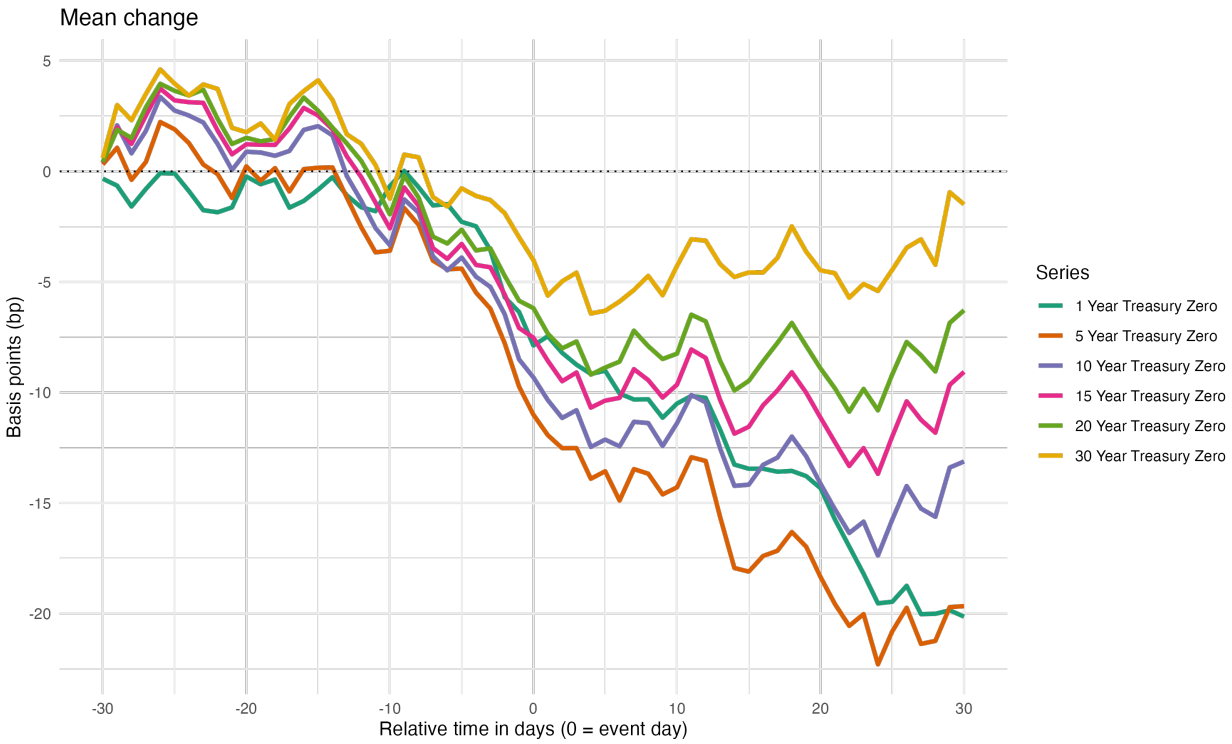

Figure 1: Mean change in yields (relative to thirty days before event) for zero-coupon US Treasury bonds. Mean taken across AI release events from November 2022 through December 2025.

are economically large and persist through roughly six weeks (30 trading-days) after the release. Yield movements appear to begin before the release of the model, which may not be surprising given that for at least some releases we know that models were made available to outside experts prior to the release date.

To assess whether movements of this size are larger than one would expect by chance, particularly given the many other market-moving events over our sample period, we use finite-sample inference procedures that are valid under a variety of maintained assumptions, including permutation tests which are valid if model release dates are as good as randomly assigned within a given calendar year, and time-shift tests which are valid if first-differenced bond yields are time-stationary. We also adopt a regression approach for our main analysis, rather than using simple averages as in Figure 1, to account for closely-spaced model releases particularly later in our sample. We find that the regression-analogs of the changes seen in Figure 1 are statistically significant at many maturities and horizons.

Different model releases may convey very different information, and pooling across

them potentially averages over both good and bad news for the rate of AI progress. To distinguish the two, we draw on data from the crowdsourced forecasting platform Metaculus, classifying each release by whether the median Metaculus forecast date for the arrival of (a particular definition of) general artificial intelligence moved earlier (reflecting good news about the rate of AI progress) or later (bad news) around the release. We find that yield declines are concentrated around, and statistically significant only for, releases which we classify as "good" news based on the Metaculus data, while "bad" releases are accompanied by systematically smaller (and statistically insignificant) changes.

We probe these directional results with a range of robustness checks. We drop subsets of releases to confirm that no one, two, or three events drive the pattern; we compare the yield changes around our release dates to those around other events (FOMC meetings, major technology-conference dates, macroeconomic data releases, and Treasury auctions); we control directly for proxies for other news, including measures of economic surprises, market volatility, and news sentiment; and we re-estimate on calendar-year subsamples. Across these checks the declines around good-news releases are directionally robust and economically large, and often remain statistically significant for longer maturities.

In order to interpret our empirical findings, we start from a complete-market consumption-based asset pricing model. In such a model, bond yields reveal the representative agent's expected marginal rate of substitution, so declines are consistent with some combination of lower expected consumption growth, greater consumption uncertainty or, potentially, a lower probability of extreme events like an existential disaster or the arrival of a post-scarcity society. A simplified version of this model, developed in Appendix A.2, separates these channels and finds that changes in consumption uncertainty appear to play little role. The model implies that the average release in our sample was associated with a decline of about 0.049 percentage points in expected annual consumption growth (assuming a relative risk aversion coefficient of two), or about 0.098 percentage points in the annual probability of an extreme outcome.

These magnitudes are large, and the model's assumptions restrictive, so it is natural to consider other interpretations, several of which have been proposed specifically in the context of AI. In heterogeneous-agent economies with incomplete markets, favorable AI news could shift wealth toward households who hold risky assets and lower expected consumption for the households who price safe bonds, and so imply falling yields even as aggregate growth expectations rise (Maresca, 2026; Moll et al., 2022). A related savings-glut channel works through the supply of capital, where AI-driven gains may accrue to wealthy

households with high saving rates and so raise the demand for safe assets (Caballero, 2026). A third possible explanation operates through the riskiness of US government debt, with favorable AI news improving the government's expected fiscal position and so making Treasuries safer (Kung et al., 2026a,b). These alternative mechanisms potentially rationalize falling yields without lower expected consumption growth.

None of the interpretations above directly speaks to the asymmetry we find for "good" vs. "bad" model releases, specifically that yields fall around "good" releases while not rising around "bad" releases. We find suggestive evidence, however, that while the "bad" releases in our sample were associated with negative updating about the arrival date for a particular definition of general artificial intelligence, if we consider other capability questions (e.g. whether an AI with human-level intelligence would be created before 2040) even the "bad" news model releases in our sample appear to convey good news on average, albeit less good than the "good" news dates. If, as seems likely, asset prices depend on beliefs about future AI capabilities beyond those captured by the particular Metaculus series we use, and these beliefs update asymmetrically across our "good" and "bad" news dates, this could explain the asymmetry in our results.

Taken together, our results suggest that investors, in aggregate, do take seriously the possibility of transformative AI, since AI model releases are associated with economically and statistically significant changes in non-AI-related asset prices. Moreover, these changes are concentrated around releases that appear to be good news about the rate of AI progress. What this updating implies for consumption and other beliefs is model-dependent: the complete-market reading points toward lower expected consumption growth or a lower probability of an extreme outcome, whereas alternative models can reconcile the same yield declines with rising growth expectations. Adjudicating among these explanations is beyond what our data can support, but any viable explanation must account for large and sustained yield declines, concentrated around AI model releases, in one of the most liquid markets in the world.

Literature Review While there is limited evidence about the aggregate impact of AI on the economy, or on market perceptions of that impact, there is a small but growing literature that uses data on job postings or asset prices to study the impact of AI on labor outcomes and compensation, as well as on firm growth (Webb, 2020; Acemoglu et al., 2022; Babina et al., 2025, 2024; Eisfeldt et al., 2026; Hampole et al., 2025). Most closely related to our analysis, Björkegren (2026), building on an earlier version of this paper, compares market responses

around open- and closed-weight model releases and finds that yields decline around closed-weight model releases, consistent with our results, while rising around open-weight releases.

The influence of growth prospects on financial markets is a widely discussed topic in the consumption-based asset pricing literature (see Mehra 2012 for a summary and Duffie 2010 for a textbook treatment). An important observation in this literature is that investors' discount factors, expected growth, and perceived growth uncertainty all influence prevailing interest rates. Chow et al. (2026) show that discounting (e.g. due to existential risk) and rising growth expectations impact interest rates the same way in the context of transformative AI. They further show that, consistent with their theory, real interest rates are positively correlated with both growth expectations and realized growth in a cross-country analysis. Other work, by contrast, finds a modest or negative relationship between growth and real interest rates (Hamilton et al., 2016; Hansen and Seshadri, 2013; Borio et al., 2017; Lunsford and West, 2019; Rogoff et al., 2024). Related to our work, Wachter and Wachter (2026) use a structural model to argue that the high level of AI-related capital expenditure from major US tech firms, together with industry forecasts, implies high forward-looking GDP growth expectations, albeit also with large uncertainty.

Our analysis also relates to the literature studying the impact of macroeconomic announcements on financial returns. In the context of Treasury yields and using FOMC announcements, Lucca and Moench (2015) document no statistically significant pre-FOMC announcement drift for Treasury bonds in the 1994-2011 period, while Savor and Wilson (2013) provide evidence of small announcement premiums for Treasury bonds, averaging about 3 basis points on announcement days, using data from 1961-2009.

The rest of the paper is organized as follows. Section 2 describes our data and empirical strategy, including the asset-pricing logic that motivates our analysis and the inference procedures we use to assess statistical significance. Section 3 reports our main empirical results. Section 4 interprets these results through both a simple complete-market asset-pricing model and a range of alternatives.

2 Data and Methods

We begin with a brief review of some results from the theory of asset-pricing which motivate our empirical strategy, then describe the AI model release series we consider, a dataset we use to proxy the information that arrived with these releases, and the statistical methods we employ.

2.1 Asset-Pricing Preliminaries

Financial theory shows that in the absence of arbitrage, regardless of whether or not markets are complete, asset prices can be expressed as the present value of future cash flows, discounted using the *stochastic discount factor* (SDF) (see Harrison and Kreps, 1979; Duffie, 2010).[1] Consider a discrete-time economy over periods $t=0,1,...,\bar{T}$, with uncertainty described by a probability space $(\Omega,\mathcal{F},\mathbb{P})$ where $\mathcal{F}_t \subseteq \mathcal{F}$ denotes the set of events which are known at period t. If we consider an asset that pays Y_{t+h} units of consumption in period $t+h$ and nothing at any other time, its period t price is given by

$$V_t(Y_{t+h}) = \mathbb{E}_t[M_{t,t+h} Y_{t+h}], \tag{1}$$

where $\mathbb{E}_t$ denotes the conditional expectation given $\mathcal{F}_t$ and $M_{t,t+h}$ is the SDF from t to $t+h$. For simplicity we write $M_{t+1} \equiv M_{t,t+1}$, so $M_{t,t+h} = \prod_{s=1}^{h} M_{t+s}$ cumulates the one-step-ahead SDFs. More generally, let $Y = \{Y_{t,h}\}_{h=0}^{\bar{T}}$ denote a general stream of payoffs $Y_{t,h}$ for periods $h=0,\cdots,\bar{T}-t$. The asset with this stream of payoffs has time-t price $V_t(Y) = \sum_{h=0}^{\bar{T}-t} V_t(Y_{t+h})$.

While the theory above applies to general assets, our empirical analysis will focus on bonds. We do this in order to isolate changes in the SDF: if market participants take seriously the possibility of transformative AI, this should change how they value long-term asset payoffs (for instance by changing beliefs about future economic growth, consumption possibilities, or existential risk) and thus be reflected in the SDF. Let 1_{t+h} denote a risk-free, h-period-ahead zero-coupon bond, i.e. the asset which pays one unit of consumption h periods in the future with certainty and nothing at any other time. By (1) this bond's time-t price is given by $V_t(1_{t+h}) = \mathbb{E}_t[M_{t,t+h}]$, and so is simply the expected h-period-ahead SDF. It is common to discuss bond prices in terms of yields, where the zero-coupon yield from t to $t+h$, $y_{t,t+h}$, is given by

$$y_{t,t+h} = V_t(1_{t+h})^{-1/h} - 1 = \mathbb{E}_t[M_{t,t+h}]^{-1/h} - 1. \tag{2}$$

Thus, a decrease in the h-period ahead zero-coupon bond yield reflects an increase in the expected h-period ahead SDF. By contrast, changes in the price of risky assets, such as equities, could reflect either changes in the SDF or changes in the expected flow of future payments. Thus, movements in equity prices are not, by themselves, indicative of changes in the SDF.

[1] If markets are complete the SDF is unique, while otherwise there may exist multiple SDFs which price all traded assets. For brevity we will refer to "the" SDF, though uniqueness is not important for our arguments except where otherwise noted.

Could the SDF Equal Zero? The absence of arbitrage is usually taken to imply that $M_{t,t+h} > 0$ with probability one, since otherwise an investor can purchase a sometimes-positive payment at time $t+h$ at zero time-t cost. Under some extreme scenarios discussed in relation to AI, however, it is possible that one might have $M_{t,t+h} = 0$. For instance, if investors take seriously the possibility of an existential disaster, then (presuming human investors do not value hypothetical asset payouts following human extinction) the SDF following such a disaster should be zero. At the opposite extreme, if investors think that AI could lead to a post-scarcity economy where all individuals enjoy unconstrained consumption possibilities, this could also imply a zero SDF.[2]

In either case, it is useful to separate changes in the probability that the SDF is exactly zero from changes in its distribution conditional on being nonzero. Fortunately, under the weaker no-arbitrage condition that $M_{t,t+h} \geq 0$, a simple diagnostic is available. Consider two assets Y_{t+h} and Y'_{t+h}, both of which pay off in period $t+h$ and at no other time. If we consider the ratio of their time-t prices, observe that

$$\frac{V_t(Y_{t+h})}{V_t(Y'_{t+h})} = \frac{\mathbb{E}_t[M_{t,t+h}Y_{t+h}]}{\mathbb{E}_t[M_{t,t+h}Y'_{t+h}]} = \frac{\mathbb{E}_t[M_{t,t+h}Y_{t+h}|M_{t,t+h}>0]}{\mathbb{E}_t[M_{t,t+h}Y'_{t+h}|M_{t,t+h}>0]}.$$

Thus, if we consider price ratios for two assets which pay off in the same period, this ratio is entirely determined by beliefs conditional on the SDF being strictly positive. Thus, if we see changes in such ratios, we know they are not driven by changes in the probability that $M_{t,t+h}$ is zero.

Empirical Strategy The model above suggests an empirical strategy for learning about changes in AI beliefs from asset prices: if we have a date t at which we believe information arrived about the future course of AI, changes in long-dated asset prices around this date should incorporate the impact of the new information about AI.

To fix ideas, again consider the price for an asset that pays Y_{t+h} units in period $t+h$. If we think new information about AI arrived at t, we may compare prices at t_- and t_+ for $t_- < t < t_+ \ll h$, and use the law of iterated expectations to write

$$V_{t_+}(Y_{t+h}) - V_{t_-}(Y_{t+h}) = V_{t_+}(Y_{t+h}) - \mathbb{E}_{t_-}\left[V_{t_+}(Y_{t+h})\right] - \mathbb{E}_{t_-}\left[\left(M_{t_-:t_+} - 1\right)V_{t_+}(Y_{t+h})\right].$$

[2]This is distinct from the rare-disaster channel (Barro, 2006; Gabaix, 2012), in which an asset continues to pay in states where aggregate consumption is low. In that case, disaster risk can raise the value of safe payoffs and lower risk-free rates.

If the time difference $t_+ - t_-$ is reasonably small we expect the final term to be negligible.[3] Hence, by (1) and the law of iterated expectations we can approximate

$$V_{t_+}(Y_{t+h}) - V_{t_-}(Y_{t+h}) \approx \mathbb{E}_{t_+}\left[M_{t_+:t+h} Y_{t+h}\right] - \mathbb{E}_{t_-}\left[M_{t_+:t+h} Y_{t+h}\right].$$

Thus the change in prices between t_- and t_+ gives us, approximately, the difference in conditional expectations for the discounted payoff $M_{t_+:t+h} Y_{t+h}$ at information sets $\mathcal{F}_{t_-}$ and $\mathcal{F}_{t_+}$. In particular, if we consider the risk-free asset $Y_{t+h} = 1_{t+h}$, changes in prices reveal the change in the conditional mean of the SDF $M_{t_+:t+h}$.

For a given pair t_- and t_+ the difference $V_{t_+}(Y_{t+h}) - V_{t_-}(Y_{t+h})$ reflects all information that arrives between those dates, not just information about AI. Hence, in our empirical analysis we will aggregate across a series of AI news dates. So long as there is not other price-relevant information which systematically arrived at the same time as AI news, comparing behavior around AI model release dates to that at other dates will isolate the effect of AI news, though it will be important to account for the likely presence of other news when assessing statistical uncertainty.

2.2 Event Date Series

To look for asset price changes around the arrival of AI news, we need a set of dates at which AI information may have arrived. While there are a variety of reasonable approaches one might take to this problem, we focus on release dates for new generative AI models from four major AI laboratories: Anthropic, Google, OpenAI, and xAI. We base this lab list on the Arena (formerly LMArena) leaderboard (Chiang et al., 2024), which ranks large language models by aggregating human pairwise preferences; these four organizations are most consistently in the leaderboard's top five and ten over the dates in our sample for which historical leaderboard snapshots are available (LM Arena, 2024). Appendix B reports the underlying frequencies and shows that our results are similar if we extend the lab list to include two additional labs, Meta and DeepSeek.[4] For each lab, we focus on major updates to the lab's flagship model series (for example, Claude for Anthropic and

[3] By Cauchy-Schwarz, $\mathbb{E}_{t_-}\left[\left(M_{t_-:t_+} - 1\right) V_{t_+}(Y_{t+h})\right] \leq \sqrt{\mathbb{E}_{t_-}\left[\left(M_{t_-:t_+} - 1\right)^2\right]} \sqrt{\mathbb{E}_{t_-}\left[V_{t_+}(Y_{t+h})^2\right] - \mathbb{E}_{t_-}\left[V_{t_+}(Y_{t+h})\right]^2}$.

[4] The flagship models from the labs in our primary sample are all closed-weight. By contrast, in the period we study both Meta and DeepSeek had open-weight flagship models. Motivated by the initial version of this paper, Björkegren (2026) argues that market responses around open- vs. closed-weight model releases appear systematically different.

GPT for OpenAI) and use the release date from the lab's website whenever possible. We consider releases between November 2022, when the public release of ChatGPT (powered by the model GPT 3.5) marked a substantial increase in public attention to AI capabilities, and December 2025. Table 1 lists the resulting 30 unique release dates.[5]

We use AI model releases as our event dates in order to capture new, forward-looking information about AI capabilities, rather than other aspects of technology or firm performance. Put differently, our hypothesis is that major model releases provide information not only about the current state of AI capabilities but also about the rate of progress, potentially causing market participants to update their beliefs about future AI development. These events are also less directly linked to financial outcomes than some other plausible dates, such as earnings announcements. At the same time, it is clear that information about AI capabilities arrives outside of new model releases for these particular AI labs. There are many other AI researchers and firms, and even the firms we study make numerous announcements and incremental model releases outside the set of major releases we consider. So long as some information arrived around the dates we study, such alternative information sources do not pose a threat to the validity of our analysis, though they may matter for interpretation.

More directly relevant for us, for at least some model releases we know outside experts had access to the model prior to the official release.[6] To partially capture such information "leakage" our empirical specifications will include a window of dates prior to the model release (30 trading-days, or approximately 6 weeks, for our preferred specifications). While this extended window is still unlikely to capture all information leakage, uncaptured leakage should tend to reduce the amount of information arriving in our event windows, which we expect to attenuate our results.

We do not detect strong cyclical patterns in the release dates. For instance, while Friday releases are underrepresented in our sample (representing two of the thirty releases), a chi-squared test does not reject the null that model releases are uniformly distributed across days of the week. The number of releases does trend upward over time, with one release date in 2022, six in 2023, seven in 2024, and sixteen in 2025; as discussed below, our inference approaches account for this trend.

[5]Anthropic's Claude 1 and OpenAI's GPT 4 were both released on March 14, 2023; we treat this same-day pair as a single event date in the main analysis.

[6]Mollick (2024) discusses having early access to Google's Gemini Advanced. METR reports having received access to a version of OpenAI's GPT 4.5 seven days before public release (METR, 2025). The US and UK AI Safety Institutes ran pre-deployment evaluations of OpenAI's o1 (UK Artificial Intelligence Safety Institute and U.S. Artificial Intelligence Safety Institute, 2024) prior to public release.

Table 1: AI Model Release Dates

Date	Lab	Model
2022 release		
11/30/2022	OpenAI	ChatGPT 3.5
2023 releases		
02/06/2023	Google	Bard
03/14/2023	OpenAI	GPT 4
03/14/2023	Anthropic	Claude 1
07/11/2023	Anthropic	Claude 2
11/03/2023	xAI	Grok 1
11/21/2023	Anthropic	Claude 2.1
12/06/2023	Google	Gemini 1.0 Nano
2024 releases		
02/15/2024	Google	Gemini 1.5 Pro
03/04/2024	Anthropic	Claude 3
05/13/2024	OpenAI	GPT 4o
05/15/2024	xAI	Grok 1.5
06/20/2024	Anthropic	Claude 3.5 Sonnet[a]
08/13/2024	xAI	Grok 2
09/12/2024	OpenAI	o1-preview
2025 releases		
01/30/2025	Google	Gemini 2.0 Flash
01/31/2025	OpenAI	o3-mini
02/17/2025	xAI	Grok 3[b]
02/24/2025	Anthropic	Claude 3.7 Sonnet
02/27/2025	OpenAI	GPT 4.5
03/25/2025	Google	Gemini 2.5
04/14/2025	OpenAI	GPT 4.1
05/22/2025	Anthropic	Claude 4
07/09/2025	xAI	Grok 4
08/05/2025	Anthropic	Claude Opus 4.1
08/07/2025	OpenAI	GPT 5
09/29/2025	Anthropic	Claude Sonnet 4.5
11/12/2025	OpenAI	GPT 5.1
11/17/2025	xAI	Grok 4.1
11/18/2025	Google	Gemini 3 Pro
12/11/2025	OpenAI	GPT 5.2

Notes: Major model release dates from the four frontier AI labs (Anthropic, Google, OpenAI, and xAI) used in our main analysis. The 31 entries correspond to 30 unique event dates after consolidating the 03/14/2023 same-day pair. Each model name links to the lab's announcement page where one was available. For a small number of releases the lab's own announcement page is dated one or two days after the model's public release (e.g., owing to time-zone conventions or a post-dated announcement); in these cases the table reports the public-release date, as detailed below.

[a] Claude 3.5 Sonnet's public release was June 20, 2024 - the date encoded in Anthropic's pinned model identifier `claude-3-5-sonnet-20240620` and reflected in Google Cloud's Vertex AI announcement of the same day - although the linked Anthropic post is dated June 21.

[b] Grok 3 was released via livestream on February 17, 2025 (see TechCrunch, February 17, 2025), although xAI's linked announcement post is dated February 19.

2.3 Forecast Data

While model releases give us dates around which we think AI-relevant information may have arrived, they do not indicate what that information, if any, was. To gauge whether a given release was viewed as positive or negative news about the rate of AI progress we use crowdsourced predictions from the forecasting platform Metaculus. Metaculus is a platform where participants make probabilistic predictions about future events, with predictions aggregated to produce community forecast distributions. We focus on a Metaculus question regarding the arrival of "general artificial intelligence," which asks users to predict the first date at which a unified AI system will be publicly known to satisfy a number of criteria (Metaculus, 2020b).[7]

Metaculus is not a betting market, so we expect that forecast updates there may be slower than financial market reactions. To summarize the forecast distribution, we focus on its median at each date in our sample. If the median date shifts down (i.e. closer to the present) around a given model release we interpret this as suggestive evidence that the release was viewed as positive news about the rate of technological progress on AI and shorthand these as FASTER releases, while if it shifts up (i.e. further into the future) we interpret this as suggestive evidence that the release was viewed as negative news about the rate of progress and shorthand these as SLOWER releases.

2.4 Regression Specification

With 30 events distributed across 38 calendar months, and preferred specifications which consider a ± 30 trading-day window around each event to capture some pre-event information arrival, our event windows overlap considerably, especially in 2025.[8] To address this overlap, we follow Björkegren (2026) and consider a regression-based event-study specification for our main results.

To construct our baseline regression specification for a given yield series, we first-difference to construct daily yield changes $\Delta y_t = y_t - y_{t-1}$. For $\mathcal{T}$ the set of AI model

[7] Specifically, the criteria involve: (1) passing an adversarial Turing test (2) having general robotic capabilities and (3) performing sufficiently well on a collection of knowledge benchmarks. The question explicitly requires these capabilities be demonstrated by a unified system rather than separate specialized models.

[8] To visualize this overlap, Appendix B.2 plots our event windows.

release dates, (e.g. consisting of 30 dates for our main sample) we then run the regression

$$\Delta y_t = \sum_{\tau \in \mathcal{T}} \sum_{j=-b}^{s} \alpha_j \mathbb{1}\{t = \tau + j\} + \varepsilon_t \qquad \text{for all trading-days } t, \tag{3}$$

which models the change in bond yields as linear in the set of active relative-time indicators, where we include indicators from b days before to s days after each model release. Thus, if a given date is within the event window for two or more events, the relative-time indicators for all of those events will be active.[9]

To report results, we compute the partial sum coefficients

$$\hat{\beta}_l = \sum_{j=-b}^{l} \hat{\alpha}_j \qquad \text{for } l \in \{-b, ..., +s\}, \tag{4}$$

These summarize the cumulative change since the start of the event window. In our baseline specification we take $b = s = 30$ trading-days, while we also report some results for shorter event windows with $b = s = 15$.

To incorporate the Metaculus data, we consider the change in the median forecast over the ± 15 day window around each model release, and partition the event dates $\mathcal{T}$ into a set $\mathcal{T}^{\mathrm{S}}$ corresponding to slower-than-expected AI progress (i.e. where the forecast median shifted further into the future, containing 13 events) and a set $\mathcal{T}^{\mathrm{F}}$ corresponding to faster-than-expected AI progress (where the forecast median shifted closer to the present, containing 17 dates). We then estimate a joint regression with separate relative-time indicators for the two subsets:

$$\Delta y_t = \sum_{\tau \in \mathcal{T}^{\mathrm{S}}} \sum_{j=-b}^{s} \alpha_j^{\mathrm{S}} \mathbb{1}\{t = \tau + j\} + \sum_{\tau \in \mathcal{T}^{\mathrm{F}}} \sum_{j=-b}^{s} \alpha_j^{\mathrm{F}} \mathbb{1}\{t = \tau + j\} + \varepsilon_t, \tag{5}$$

and report cumulated coefficient series $\hat{\beta}_l^S$ and $\hat{\beta}_l^F$.

[9] This fully addresses the event-window overlap if we think yield responses to model releases are homogeneous across events and take the additive form suggested by (3). If instead effects are heterogeneous across releases, then the coefficients α_j will reflect a linear combination of effects at different lags, similar to the issues highlighted by Sun and Abraham (2021) for dynamic difference-in-differences with heterogeneous effects. Nevertheless, the event-window overlap in our sample is sufficiently large that we think an imperfect solution is preferable to ignoring the issue.

2.5 Inference Methods

To gauge whether markets are responding to AI model releases, we need a way to judge whether the yield movements we observe around model releases are larger than one would expect due to chance, particularly since many other market-moving events took place over our sample period. Given our very limited sample size, it is important to use a method that is valid in small samples. We consider three finite-sample inference procedures, each valid under different maintained assumptions.

Permutation test The permutation testing approach is based on the assumption that the timing of model releases within a given calendar year is as good as random. To implement this approach, we draw *placebo event dates* $\mathcal{T}^*$ (or $\mathcal{T}^{S,*}$,$\mathcal{T}^{F,*}$ for our results that split by Metaculus update direction) by sampling trading dates uniformly at random within each calendar year, fixing the number of events per year (or events with a given direction, for the split results) to match the actual data, and re-estimate (3) for each draw of placebo dates to obtain a placebo horizon-l coefficient $\hat{\beta}_l^*$. Repeating this exercise 5000 times yields a placebo distribution for the coefficient at each horizon, and we compute a (two-sided) p-value by considering the fraction of placebo draws which fall further into the tails of the distribution than does the observed draw. Thus, for instance, if the coefficient $\hat{\beta}_l$ computed in our actual data falls at the 95th percentile of the placebo distribution, the corresponding two-sided p-value would be approximately 10%.[10]

The permutation approach jointly tests that (a) the model release dates are uniformly distributed within a given calendar year (where the frequency may vary across years) and (b) yields are unrelated to the model release date. In the terminology of the treatment effects literature, we are testing the "sharp null" of no yield response whatsoever. We summarize results using p-values, which measure the probability that we would observe an event as or more extreme were the null hypothesis true. Under (a) and (b), standard arguments show that for any horizon l, the p-value obtained from the permutation approach will be larger (in the sense of first-order stochastic dominance) than a uniform distribution.

[10]Formally, for $\left\{\hat{\beta}_{l,r}^*\right\}_{r=1}^{R}$ the placebo draws, we compute the p-value as

$$p = 2 \cdot \min\left\{ \frac{1+\sum 1\left\{\hat{\beta}_{l,r}^* < \hat{\beta}_l\right\}}{R+1}, \frac{1+\sum 1\left\{\hat{\beta}_{l,r}^* > \hat{\beta}_l\right\}}{R+1} \right\},$$

where we take $R = 5000$.

Thus, small p-values provide evidence against the null, and if we reject (a) and (b) when the p-value is less than 5%, we have no more than a 5% chance of obtaining a false rejection.

Time-shift test. One limitation of the permutation test is that it assumes no time-series dependence between model releases, e.g. ruling out that when one lab releases a new model, other labs accelerate their next release in response. To address this and other dependence issues in the model-release series, we consider a time-shift test introduced by Harris (2020) and Yuan and Shou (2024), which to our knowledge has not previously been used in economics or finance. To implement this test, we fix the model release series but replace the yield series $\{y_t\}$ with $\{y_{t+\kappa}\}$ for shifts $\kappa \in \{-R,...,-1,0,1,...,R\}$, re-estimating (3) for each shift. We use $R=199$ trading-days and compute one-sided p-values by taking the number of placebo estimates smaller (or larger) than the original estimate and dividing by $R+1$. Two-sided p-values are then obtained by taking the minimum of the two one-sided p-values and multiplying by two.[11]

The time-shift approach jointly tests that (a) the first-differenced yield series Δy_t is time-stationary, meaning that for any finite l and κ, $(\Delta y_t,...,\Delta y_{t+l})$ and $(\Delta y_{t+\kappa},...,\Delta y_{t+l+\kappa})$ have the same distribution ex-ante, and (b) the yield series is independent of the release-date series. Unlike the permutation approach, this test restricts the distribution of bond yields but makes no assumption on the model release series. It is also worth highlighting that the p-value computation in this case does not simply compare the observed statistic to the distribution of shifts obtained across different values of κ (which would correspond to a denominator of $2R+1$, rather than $R+1$) but is instead larger - Yuan and Shou (2024) show that this larger p-value is important to account for the dependence of $\hat{\beta}^*_{h,\kappa}$ coefficients computed for nearby values of κ. A limitation of the time-shift approach is that we must have at least 199 trading-days of data beyond the final date in our analysis sample in order to use this approach. Consequently, when we use this approach we must drop model releases which occurred after July 07, 2025, reducing our effective sample size. Given this smaller sample size, together with the weaker assumptions the shift approach imposes on the model

[11] Formally, for $\left\{\hat{\beta}^*_{l,\kappa}\right\}_{kappa=-R}^{R}$ the placebo draws, we compute the p-value as

$$p = 2 \cdot \min\left\{ \frac{1+\sum 1\left\{\hat{\beta}^*_{l,\kappa} < \hat{\beta}_l\right\}}{R+1}, \frac{1+\sum 1\left\{\hat{\beta}^*_{l,\kappa} > \hat{\beta}_l\right\}}{R+1} \right\},$$

where we take $R=199$.

release series, we expect this approach to have lower power than the permutation method.

Rank correlation test. In the tests discussed above, we use the Metaculus data only to partition the model release dates $\mathcal{T}$ into the slower and faster sets $\mathcal{T}^S$ and $\mathcal{T}^F$. Our final test instead examines how yield changes around different events covary with Metaculus forecast revisions. Specifically, we compute event-specific yield changes $\{y_{\tau+s}-y_{\tau-b}\}_{\tau\in\mathcal{T}}$ and event-specific forecast revisions $\{m_{\tau+s}-m_{\tau-b}\}_{\tau\in\mathcal{T}}$ around our event dates and then compute the Spearman (rank) correlation between the two. We again obtain a two-sided p-value by a permutation approach.[12]

The rank correlation approach jointly tests that (a) we can treat either the ordering of the Metaculus revisions or the yield changes across event dates as a uniform random draw from the set of possible orderings and (b) these two orderings are independent. Unlike both the permutation and time-shift approaches, it discards the time-series structure of both the yield and model release series, using only the (relative) magnitudes of changes around the various dates.

Multiple testing corrections The tests described above produce a separate p-value for each yield series considered. Since we will consider both nominal Treasuries and TIPS, and a variety of maturities for each, one might be concerned about multiple testing issues, e.g. that we could obtain spurious rejections due to the large number of p-values considered. To address this, in our tables below we report multiple-testing corrected p-values, which produce a single p-value for each testing-method-by-analysis-window cell we consider. Thus, for instance, we will have one p-value for permutation inference and the ± 15 trading day window, and another for shift inference and the ± 30 trading day window.[13] If one desires an overall p-value for each inference method and table, Boole's inequality implies that it suffices

[12]Formally, for $r \in \{1,...,R\}$ we randomize the order of $\{m_{\tau+s}-m_{\tau-b}\}_{\tau\in\mathcal{T}}$ across event dates, and compute the rank correlation ρ_r^* based on this randomized order. We again obtain the two-sided p-value by examining how far the observed rank correlation lies in the tail of the permutation distribution,

$$p = 2 \cdot \min\left\{ \frac{1+\sum 1\{\rho_r^* < \rho\}}{R+1}, \frac{1+\sum 1\{\rho_r^* > \rho\}}{R+1} \right\},$$

where we take $R = 5000$.

[13]Formally, for $\mathcal{I}(c)$ the set of assets considered for inference-method-by-analysis-window cell c, and $\hat{\beta}_{l,i}$ the end-horizon coefficient for a given asset, we define $T_c^+ = \max_{i\in\mathcal{I}(c)}\hat{\beta}_{l,i}$ and $T_c^- = \min_{i\in\mathcal{I}(c)}\hat{\beta}_{l,i}$ as the largest and smallest coefficients within the cell, and apply the inference methods discussed above to these max and min coefficients to obtain one-sided p-values, where as above we then take the min of the two p-values and multiply by two to construct a two-sided p-value.

to take the minimum of the ±15 and ±30 p-values and multiply by two, $2 \cdot \min\{p_{15}, p_{30}\}$.

3 Empirical Results

We next report our empirical results. We begin by examining whether there are statistically significant changes in yields around our model release dates, pooling all releases. We then incorporate the Metaculus forecast data as a proxy for what information arrived around each release.

3.1 Baseline Results

As our primary yield series, we consider zero-coupon yields on nominal Treasury and TIPS bonds computed from the Federal Reserve's fitted yield curve estimates (Gürkaynak et al., 2006, 2010). Treasury zero-coupon yields (Board of Governors of the Federal Reserve System, US, 2025a) are available at maturities from one to thirty years, while TIPS zero-coupon yields (Board of Governors of the Federal Reserve System, US, 2025b) are available from two to twenty years, both at daily frequency. The results in this section pool across all 30 release dates in Table 1.

Table 2 reports two-sided p-values for the end-horizon cumulative coefficient $\hat{\beta}_s$ from (4) at each maturity, side-by-side for the ±15 and ±30 trading-day windows and for permutation and time-shift inference. Recall that p-values measure the probability of observing a more extreme outcome were the null hypothesis true. Hence, small p-values correspond to outcomes that are unlikely to arise under the null (in our case, if AI model releases have no effect on yields, and the auxiliary assumptions that either model release dates are as good as random within year for the permutation results, or that first-differenced yields are stationary for the time-shift results). In particular, a 10% test rejects when the p-value is less than 0.1, and a 5% test rejects when the p-value is less than 0.05.

The results in Table 2 paint a consistent picture. The ±15 coefficients are statistically significant under permutation inference for both Treasury and TIPS at maturities five years and longer. That is, under the maintained assumption that AI release dates are as good as random within a given year, we reject the null hypothesis that AI model release dates are unrelated to Treasury and TIPS yields for maturities over five years. For the ±30 coefficients, additional days enter the event window, so the results are noisier and the p-values larger; Treasury permutation p-values run from 0.04 to 0.09 at five-to-thirty years,

Table 2: Two-sided p-values for yield changes around AI model releases

	Treasury Zero-Coupon				TIPS Zero-Coupon			
	Permutation		Shift		Permutation		Shift	
Maturity	±15	±30	±15	±30	±15	±30	±15	±30
1 Year	0.16	0.11	0.19	0.53	–	–	–	–
2 Year	0.04**	0.08*	0.12	0.17	0.12	0.16	0.16	0.15
5 Year	0.01***	0.05*	0.06*	0.19	0.02**	0.11	0.07*	0.22
10 Year	0.00***	0.04**	0.04**	0.19	0.01***	0.08*	0.07*	0.21
15 Year	0.00***	0.04**	0.03**	0.15	0.01***	0.08*	0.07*	0.19
20 Year	0.00***	0.04**	0.03**	0.15	0.01***	0.09*	0.07*	0.19
30 Year	0.01**	0.10*	0.09*	0.17	–	–	–	–

Notes: P-values for regression-based event study coefficients. Permutation inference uses 5000 year-stratified placebo draws. Shift inference uses ±199 day shifts. Multiple-testing corrected p-values: permutation $p = 0.039^{**}$ (±15) and $p = 0.132$ (±30); shift $p = 0.080^{*}$ (±15) and $p = 0.150$ (±30). Significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

and TIPS from 0.08 to 0.11 at five-to-twenty years.

Time-shift p-values are uniformly larger than the corresponding permutation p-values; under this approach the ±15 coefficients are marginally significant for both Treasuries and TIPS at the maturities between five and twenty years, while none of the ±30 coefficients reach significance at conventional levels. After multiple-testing correction, the ±15 day permutation and time-shift results remain significant at the 5% and 10% levels respectively, while the ±30 day results are not significant at conventional levels.

To further explore what is happening around AI model releases, Figure 2 plots cumulative regression coefficients $\hat{\beta}_l$ for Treasury (top panel) and TIPS (bottom panel) zero-coupon yields, with maturities shown as subplots. The blue line is the observed coefficient $\hat{\beta}_l$, while the dark grey solid line is the placebo mean across 5000 placebo draws, and the dashed grey lines are the 10%, 5%, and 1% two-sided permutation critical values. So, β_l is statistically significant at the 10% level, for instance, if and only if the blue curve falls outside the 10% lines at horizon l. Examining the figure, we see that the coefficients tend to track their respective placebo means up to 10-15 days prior to each model release, after which they decline, with the precise timing of the decline varying across series.[14] The overall decline in yields around model releases is economically (and in many cases statistically) significant,

[14]To the extent one is willing to embrace the suggestion that most market responses are occurring within 15 days before the model release, this motivates focusing on the shorter ±15 window.

particularly at longer maturities.

Viewed through the lens of the asset-pricing model discussed in Section 2.1, yield declines around model releases could reflect either an increase in the expected SDF around these model releases, conditional on the SDF being nonzero, or a decline in the probability that the SDF is exactly zero. Fortunately, we can distinguish these possibilities by considering price ratios for pairs of assets which both pay off at the same date. Specifically, let 1_{t+h}^{Treas} denote the nominal Treasury zero which pays off at date $t+h$, and 1_{t+h}^{TIPS} the TIPS zero paying off at the same date. Given the possibility of inflation, 1_{t+h}^{Treas} is not risk-free when measured in consumption terms.[15] If we consider the price ratio

$$\frac{V_t(1_{t+h}^{Treas})}{V_t(1_{t+h}^{TIPS})} = \left(\frac{1+y_{\text{Treas}}}{1+y_{\text{TIPS}}}\right)^{-h} \tag{6}$$

then as discussed in Section 2.1, this ratio should reflect only the distribution of the SDF conditional on the SDF being nonzero. Thus, if the ratio (6) moves around the model releases in our sample, this implies that the yield changes we see are not solely driven by changes in the probability that the SDF is zero.

Figure 3 plots the results obtained by using the price ratio (6), rather than the bond yield, as the outcome in regression (3). To highlight the maturity pattern, here we show the end-horizon (i.e. $l=30$, 6-week post-release) cumulated coefficients, plotted across the maturities from 2-20 years where we have both Treasury and TIPS yields. Examining this plot, we see little change at short maturities, while at longer maturities the coefficient climbs substantially above the placebo mean, and this departure is marginally significant from the 9-year maturity onwards, where the degree of significance increases with maturity. Thus, under the asset pricing model of Section 2.1 we strongly reject the hypothesis that the observed changes around model releases are driven solely by changes in the probability that the SDF is zero. Moreover, the direction of the deviation from the placebo mean suggests that long-horizon inflation expectations may be falling around these releases, since the price of nominal Treasuries is rising relative to TIPS.

[15] Note that even 1_{t+h}^{TIPS} may be subject to some risks, for instance if the inflation indexing is imperfect or there is a risk of default.

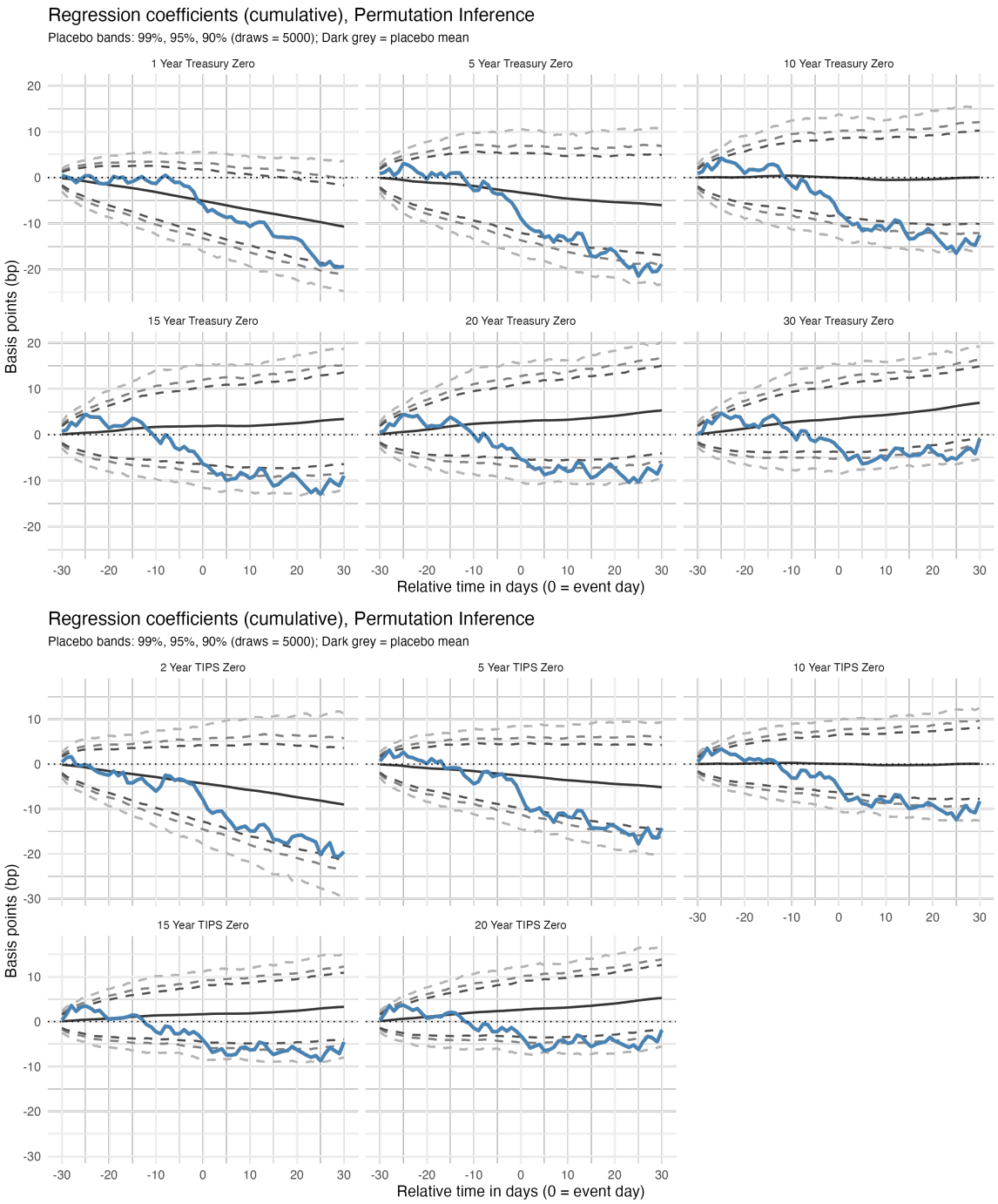

Figure 2: Regression-based event study for zero-coupon Treasury yields (top panel) and TIPS yields (bottom panel). Plots show cumulative regression coefficients from -30 to $+30$ trading-days relative to AI model release dates, with maturities shown as faceted subplots. Blue line is the observed cumulative coefficient; dark grey line is the placebo mean across 5000 year-stratified placebo draws; dashed grey lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

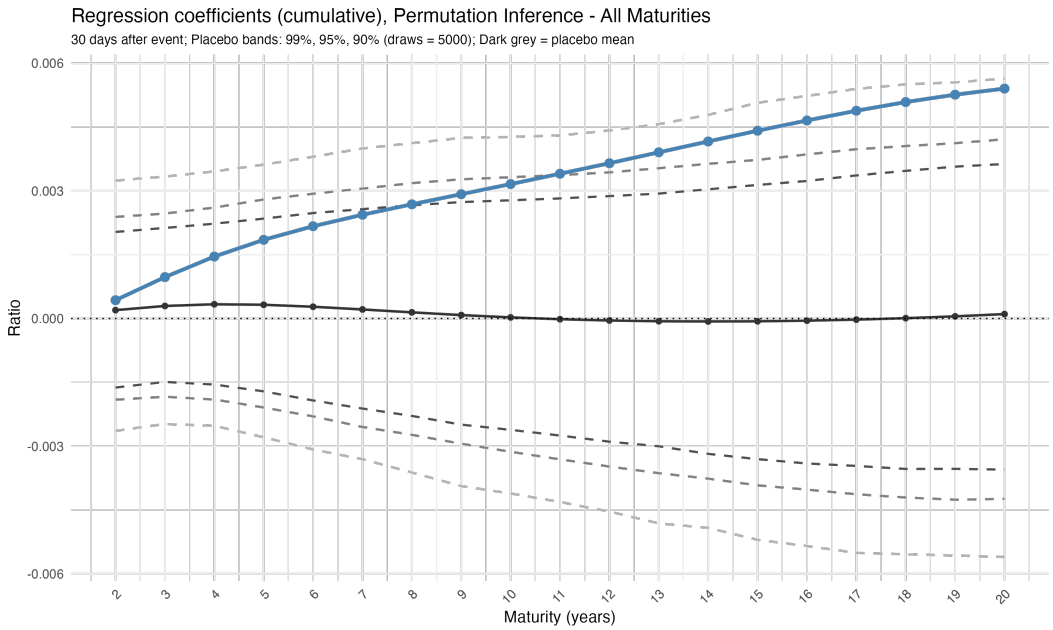

Figure 3: End-horizon change in the Treasury-to-TIPS price ratio across maturities, pooling across all 30 AI release events. Points show cumulative change at day +30 relative to day −30 in $((1+y_{\text{Treas}})/(1+y_{\text{TIPS}}))^{-h}$ for each maturity h. Under the identity of Section 2.1, this ratio depends only on beliefs about the SDF and the payoff conditional on the SDF being nonzero, so movement in the ratio is informative about the conditional-on-nonzero component of the SDF response. Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

3.2 Split by Metaculus Direction

We have shown that bond yields fall around AI model release dates, particularly at longer maturities, and that this response does not appear to be driven exclusively by changes in the probability that the SDF is zero. Different model releases may, however, convey very different sorts of information. For instance, some releases may show evidence of faster technological progress than investors expected, while others may be technological disappointments. We next extend our analysis by using changes in the Metaculus median prediction series (described in Section 2.3 above) as a proxy for the sort of information which arrived around each model release.

Specifically, for each event date τ we compute the change in the Metaculus median prediction over the ± 15 trading-day window around τ. Events with a positive change (forecast median pushed later, suggesting bad news for the rate of progress) are labeled SLOWER, while events with negative change (forecast median pulled earlier) are labeled

FASTER.[16] This split gives $|\mathcal{T}^{S}|=13$ and $|\mathcal{T}^{F}|=17$. The choice of window length trades off the possibility that Metaculus revisions arrive slowly (which favors a wider window) against the possibility of other AI-relevant news arriving (which favors a narrower one). The ± 15 window we use matches the shorter of the two windows we consider for the financial data, which also appears to contain most of the systematic yield change around our events.[17]

Tables 3 and 4 report two-sided p-values for the end-horizon cumulative coefficients in regression (5) at each maturity, separately for the SLOWER and FASTER coefficients. Splitting by direction sharpens the conclusions obtained above. In particular, the SLOWER coefficients are uniformly insignificant at conventional levels. By contrast, the FASTER coefficients are at least marginally statistically significant under permutation inference for almost all maturities and for both Treasuries and TIPS. Even for the shift inference approach, the ± 15 day coefficients are at least marginally significant. After multiple testing correction, both the ± 15 and ± 30 day permutation results remain significant at the 5% level or less, and the ± 15 day time-shift results are significant as well.[18]

Table 3: P-values for SLOWER coefficients (timeline lengthened; 13 events)

	Treasury Zero-Coupon				TIPS Zero-Coupon			
	Permutation		Shift		Permutation		Shift	
Maturity	± 15	± 30	± 15	± 30	± 15	± 30	± 15	± 30
1 Year	0.63	0.73	1.00	1.00	–	–	–	–
2 Year	0.98	0.50	1.00	0.84	0.40	0.64	1.00	1.00
5 Year	0.53	0.37	0.44	0.57	0.80	0.92	1.00	0.87
10 Year	0.45	0.42	0.29	0.59	0.83	0.78	0.67	0.81
15 Year	0.42	0.45	0.28	0.62	0.67	0.76	0.58	1.00
20 Year	0.37	0.51	0.23	0.64	0.58	0.80	0.49	1.00
30 Year	0.36	0.76	0.42	1.00	–	–	–	–

Notes: Events classified by direction of Metaculus prediction changes over ± 15 trading-days. Permutation uses 5000 year-stratified placebo draws. Shift uses ± 199 day shifts. Multiple-testing corrected p-values: permutation $p=0.976$ (± 15) and $p=0.744$ (± 30); shift $p=1.000$ (± 15) and $p=1.000$ (± 30).

[16] The forecast change is nonzero around all releases in our sample.

[17] Note, in addition, that our tests remain valid for the null of no relationship between AI releases and yields regardless of the choice of window.

[18] See Björkegren (2026) for parallel results which intersect the SLOWER vs. FASTER split with that between open- and closed-weight model releases.

Table 4: P-values for FASTER coefficients (timeline shortened; 17 events)

	Treasury Zero-Coupon				TIPS Zero-Coupon			
	Permutation		Shift		Permutation		Shift	
Maturity	±15	±30	±15	±30	±15	±30	±15	±30
1 Year	0.02**	0.08*	0.14	0.04**	–	–	–	–
2 Year	0.01***	0.09*	0.05*	0.10	0.00***	0.02**	0.03**	0.12
5 Year	0.01***	0.09*	0.05*	0.15	0.00***	0.05**	0.03**	0.15
10 Year	0.00***	0.06*	0.05*	0.13	0.00***	0.05**	0.04**	0.15
15 Year	0.00***	0.05**	0.05*	0.12	0.00***	0.05**	0.04**	0.12
20 Year	0.00***	0.04**	0.03**	0.10	0.00***	0.06*	0.05*	0.11
30 Year	0.02**	0.07*	0.08*	0.10	–	–	–	–

Notes: Events classified by direction of Metaculus prediction changes over ±15 trading-days. Permutation uses 5000 year-stratified placebo draws. Shift uses ±199 day shifts. Multiple-testing corrected p-values: permutation $p=0.007^{***}$ (±15) and $p=0.040^{**}$ (±30); shift $p=0.030^{**}$ (±15) and $p=0.120$ (±30). Significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

As we did for the pooled results, Figure 4 plots the cumulative coefficient paths for Treasury and TIPS zero-coupon yields, with the SLOWER coefficients in the top row of each panel and the FASTER coefficients in the bottom. Examining these plots, we see that the SLOWER coefficients remain statistically insignificant at conventional levels uniformly across horizons. By contrast the FASTER coefficients track the placebo mean until 10-15 days before the model release after which they decrease, similar to what we observed for the pooled specification but with larger end-horizon magnitudes. The yield declines indicated by the FASTER coefficients are again economically and statistically significant.

One limitation of the SLOWER vs. FASTER split is that we discard the magnitude of the forecast revision around each event, reducing the forecast information to a binary label. As an alternative, Table 5 reports results from our rank correlation approach, which as described in Section 2.5 focuses on the relative size of yield and forecast revisions around the model releases in our sample. We find strong statistical significance in the ±15 day specifications, across a range of maturities, for both Treasuries and TIPS, where this significance remains even after multiple testing correction. By contrast, the results are insignificant at conventional levels for the longer ±30 day windows.

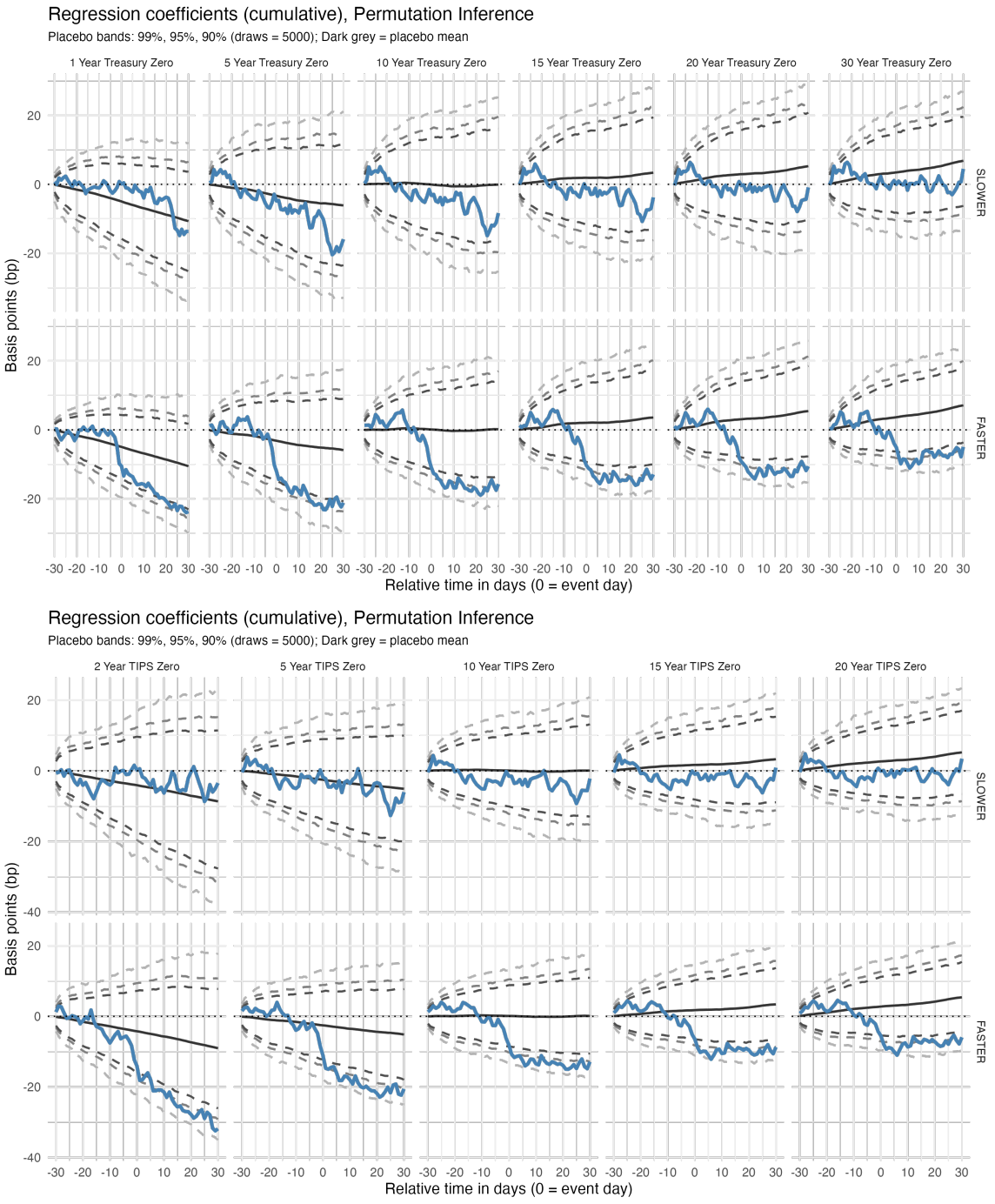

Figure 4: Regression-based event study for zero-coupon Treasury yields (top panel) and TIPS yields (bottom panel), split by direction of Metaculus prediction changes. Events classified as SLOWER (13 events) or FASTER (17 events) based on changes in the median prediction over ± 15 trading-days around each release. Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

Table 5: Rank correlations between yield and Metaculus changes

	Treasury Zero-Coupon				TIPS Zero-Coupon			
	±15 days		±30 days		±15 days		±30 days	
Maturity	ρ	p-value	ρ	p-value	ρ	p-value	ρ	p-value
1 Year	0.200	0.28	-0.084	0.67	–	–	–	–
2 Year	0.234	0.21	0.094	0.62	0.360	0.06*	0.253	0.19
5 Year	0.343	0.06*	0.119	0.53	0.502	0.00***	0.189	0.30
10 Year	0.497	0.01***	0.151	0.41	0.610	0.00***	0.219	0.24
15 Year	0.551	0.00***	0.179	0.34	0.567	0.00***	0.229	0.22
20 Year	0.591	0.00***	0.171	0.36	0.575	0.00***	0.254	0.18
30 Year	0.430	0.02**	0.179	0.34	–	–	–	–

Notes: Spearman rank correlations between yield changes (post minus pre) and Metaculus prediction changes over ±15 trading-days. 30 events. Multiple-testing corrected p-values: $p = 0.003^{***}$ (±15) and $p = 0.441$ (±30). Significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

3.3 Results for Other Financial Series

The SLOWER vs. FASTER asymmetry of Section 3.2 extends to a broader range of financial series. We briefly summarize the patterns here, while all figures are deferred to Appendix C. For brevity, we focus on permutation inference and the ±30 analysis window for all series.

Appendix C begins by showing end-horizon ($l=30$), cross-maturity plots for zero-coupon Treasuries (in Figure 8), TIPS (in Figure 9), and inflation breakevens (in Figure 10, Board of Governors of the Federal Reserve System, US 2025b). These plots give a clearer picture of the maturity structure of yield changes around our model release dates. Specifically, the SLOWER coefficients show small, insignificant declines for nominal Treasuries relative to the placebo mean, while the FASTER coefficients show large declines. The magnitude of the decline relative to the placebo mean does not vary strongly with maturity, but is somewhat larger for intermediate maturities. For TIPS, the SLOWER coefficients track the placebo means fairly closely, while the FASTER coefficients again show large declines relative to the placebo mean, now with the largest decline at short maturities. As might thus be expected, for inflation breakevens the SLOWER coefficients show uniform (and often marginally statistically significant) declines across maturities, with the largest declines at short maturities. By contrast, the FASTER coefficients show a (statistically insignificant) rise at short maturities, along with a decline at longer maturities, which is marginally statistically significant at the longest maturities. Taken at face value, these results for inflation breakevens suggest that dis-

appointing AI progress may be associated with lower inflation expectations across horizons, while faster AI progress may be associated with lower inflation expectations at long horizons.

We next turn to other financial series, where for each series we report regression results based on the SLOWER vs. FASTER split, showing the evolution of coefficients across our event windows, parallel to Figure 4 above. We begin with Figure 11, whose top panel shows the "over-time" results for the inflation breakevens, to complement the end-horizon results in Figure 10. The second panel plots our results from using Treasury trading volumes (Financial Industry Regulatory Authority, 2026), where we take the outcome in (5) to be the daily trading volume summed over on-the-run and off-the-run Treasuries and de-meaned over our analysis sample. The latter panel shows that neither the SLOWER nor FASTER releases are associated with systematically different Treasury trading volumes.

The top panel of Figure 12 plots results for investment-grade corporate bond indices (Ice Data Indices, LLC, 2025b), for maturity bins ranging from 1-3 to 15+ years. These series behave similarly to Treasuries and TIPS, with the SLOWER coefficients remaining insignificant and relatively close to the placebo mean, while the FASTER coefficients show economically (and at some horizons statistically) significant declines. The bottom panel of Figure 12 plots results for a measure of option-adjusted corporate bond spreads (ICE Data Indices, LLC, 2025a), again broken out by maturity. Here, we see a statistically significant increase in the SLOWER coefficients, accompanied by smaller and statistically insignificant declines in the FASTER coefficients. Thus, interestingly, it appears that corporate bond spreads tended to widen around model releases with slower-than-expected progress, despite trending down overall across our sample period as shown by the placebo mean.

To the extent yields and spreads move around AI model releases, one could reasonably ask whether this reflects changing confidence in US-based corporations, the US government, or both. To provide limited evidence toward this point, the top panel of Figure 13 plots results for a US dollar index (Board of Governors of the Federal Reserve System, US, 2025c), while the second panel plots results for 5-year US sovereign credit default swaps (CDS, Bloomberg L.P., 2026). For exchange rates, we observe small and largely insignificant downward deviations from the placebo mean in both the SLOWER and FASTER coefficients, where, unlike in our other results, the statistically significant deviations for the SLOWER coefficients occur early in the event window. Similarly, for the CDS the coefficients largely track the placebo mean, with the exception of some significant FASTER coefficients early in the event window. Together, these results do not suggest a strong pattern in either exchange rates or CDS around the model releases in our sample.

While we primarily focus on bonds for reasons discussed in Section 2.1, one might also wonder if equity prices change around the model release dates we consider. As initial evidence toward this question, Figure 14 plots the results from applying our approach to the S&P 500 (S&P Dow Jones Indices LLC, 2026, in the top panel) and the stock price of NVIDIA (Yahoo Finance, 2026, in the bottom panel), both measured as percentage changes relative to the start of the event window for interpretability. We focus on NVIDIA since their chips are used by multiple major AI labs, so one might expect its future revenues to be sensitive to a broad variety of AI-related events. In both cases, however, we see that the SLOWER and FASTER coefficients both tend to track their placebo means, up to limited periods of significance early in the event window.

3.4 Robustness

We conduct a variety of analyses to explore the robustness of our results. Here we briefly discuss four families of robustness checks, where we again focus on permutation inference unless otherwise noted. Figures, and additional details, are deferred to Appendix D. For brevity, in these robustness checks we focus on results for US Treasuries split by Metaculus direction.

3.4.1 Robustness to Dropping Events

Since we examine yield changes around a relatively small number of model releases, one might worry that our findings could be driven by one or a few extreme events. For instance, the March 14, 2023 model releases in our data occurred soon after the March 10 collapse of Silicon Valley Bank. To explore robustness of our results, Figures 15 and 16 in Appendix D report cumulative event-study paths obtained by dropping all subsets of one, two, and three event dates, plotted against the full-sample paths. We find that our results are quantitatively similar, and at some horizons remain at least marginally significant, when dropping any one date from our sample. Even when dropping two or three dates, the FASTER paths remain directionally similar to the full-sample estimate, though there exist date combinations such that they lose statistical significance.[19]

[19]To let us compare these results to a fixed placebo distribution, for these results we drop the stratification of event dates by year, and further assign each placebo date independently to the SLOWER and FASTER groups with probability proportional to their full-sample share.

3.4.2 Controlling for Other News

As a further robustness check, we directly control for proxies for certain non-AI news that arrived during our analysis period. Specifically, we add three series intended to capture other information that might have impacted bond yields: (i) the Citigroup US Economic Surprise Index (Citigroup Global Markets, 2025), which summarizes the deviation of economic data releases from forecasts; (ii) the Cboe VIX volatility index (Cboe, 2025), which is an option-implied measure of stock-market volatility; and (iii) the Federal Reserve Bank of San Francisco Daily News Sentiment Index (Shapiro et al., 2022; Federal Reserve Bank of San Francisco, 2023), which summarizes the economic sentiment of news articles from a variety of sources. We add each series, along with fifteen daily lags, as additional regressors in (5). Figure 17 in Appendix D shows that the SLOWER coefficients remain small and largely statistically insignificant, while the FASTER coefficients are directionally similar to our baseline results and remain statistically significant for longer maturities.

3.4.3 Alternative "Placebo" Dates

Our permutation approach is based on the assumption that AI model release dates are as good as random within a calendar year and, consequently, that systematic changes in bond yields around AI model releases may reflect beliefs about AI. Our inference results would thus be invalid if the timing of AI model releases were systematically related to yield movements for other reasons, for instance because AI labs attempt to time their releases around market movements directly, or because they time model releases around other, non-AI events which systematically move markets. While the leave-j-out analysis above shows that our findings are directionally robust to dropping any small set of model releases, it does not address the possibility of more pervasive timing correlation.

For any alternative date series, an extreme form of timing correlation would be for AI model releases to be drawn solely from that series. If the subset of dates selected was as good as random from within that series, we could repeat our permutation calculation to derive thresholds for statistical significance. Motivated by this observation, Figures 18, 19, and 20 in Appendix D report versions of Figure 4 which use one of (i) FOMC meetings, (ii) major tech-conference dates from the "Magnificent Seven" firms (Table 8), (iii) CPI release dates, (iv) BLS employment situation release dates, (v) retail sales report release dates, and (vi) Treasury auction dates for 10-, 20-, and 30-year bonds as the source of our

placebo dates, though in fact none of these series nests our AI model release series.[20] Our findings for the FASTER subset remain statistically significant relative to these alternative placebo distributions at long maturities.

3.4.4 Alternative Analysis Samples

To complement our full-sample results, we split the analysis sample into three calendar-time subsamples and re-estimate (5) within each: November 2022 through December 2023, January through December 2024, and January through December 2025. Figures 21 and 22 in Appendix D show that the FASTER point estimates are directionally similar across all three subsamples, though precision is reduced because each subsample contains fewer events than the full sample. The SLOWER results vary much more across subsamples, suggesting yield declines in some cases and yield increases in others.

4 Interpretation and Discussion

Section 3 documents a decline in US government bond yields, both Treasuries and TIPS, around AI model releases, and moreover shows that these declines are concentrated around the model releases which the Metaculus forecast suggests revealed faster-than-expected AI progress. In this section, we highlight economic mechanisms which might explain these yield declines. We begin by briefly discussing a simple complete-market model that allows us to quantitatively interpret our results, and then discuss alternative models from the literature. We close by revisiting the asymmetry between our SLOWER and FASTER estimates in Section 3, and suggest that it may be due to the importance of other AI-relevant information, beyond the forecast date we consider, for asset prices.

4.1 Complete Market Model: Representative Agent

As discussed in Section 2.1, under the absence of arbitrage the prices, and consequently yields, of risk-free bonds reflect the mean of any valid SDF. If markets are complete the SDF is unique, and in equilibrium every agent's marginal rate of substitution (MRS) between present and future consumption must be equal to the SDF. Moreover, the SDF is equal to the MRS for a representative agent who receives aggregate consumption (Lucas,

[20] Indeed, as illustrated in Appendix B.2 none of these series appears to have especially strong temporal correlation with our event dates.

1978; Hansen and Singleton, 1983). Thus, we can interpret changes in bond yields as changes in this representative agent's expected MRS, which is directly informative about the anticipated distribution of aggregate consumption.

Appendix A.1 makes the above precise in the context of a representative agent model with time-separable preferences, where there may exist some (stochastic) date T after which asset prices are irrelevant for utility (for instance following an existential disaster or the advent of a post-scarcity society). Under such a model, falling bond yields correspond to an increase in the expected marginal utility of consumption and thus are consistent with a decrease in the expected arrival rate for T, or, assuming concave utility, falling expected consumption. As highlighted in prior work by Chow et al. (2026) on the implications of transformative AI for real interest rates, similar conclusions also hold in some models allowing incomplete markets.

To explore the quantitative implications of our results through the lens of the representative agent model, Appendix A.2 imposes a number of further simplifications, including CRRA utility and lognormal consumption uncertainty in every period. Under these assumptions, a particular difference-in-differences of log forward yields can be written in closed form in terms of three components: the change in the expected arrival rate for T, the change in expected consumption, and the change in consumption uncertainty. Furthermore, this expression separates the component of the yield change due to a combination of the former two components and that due to the latter.

We find that, on average across our model releases, the uncertainty channel appears to play very little role in explaining our results, suggesting that (under the restrictions of this model) changes in consumption expectations or beliefs about extreme events (captured by T) are instead the leading explanations. Moreover, the belief changes on either dimension needed to explain our results are quite large. Taken at face value, the model suggests that the releases in our sample were associated with large declines in expected aggregate consumption growth (0.049 percentage points in annual growth per model release assuming a relative risk aversion coefficient of 2, or 1.47 percentage points if we scale by the 30 releases), dramatic declines in the expected arrival rate for T (0.098 percentage points in the annual arrival probability of T per model release, or 2.93 percentage points if we scale by the 30 releases), or some combination of the two. Given that these estimates may seem surprisingly large, and that the model used to derive them imposes a priori implausible assumptions (e.g. complete markets, a representative agent, strong functional form restrictions), it is natural to seek alternative explanations. We next turn to other models from the literature with a focus on

three classes of models which have been explicitly discussed as possible explanations for our finding (in a previous draft of this paper) that bond yields decline around AI model releases.

4.2 Heterogeneous Agent Incomplete Market Models

In reality markets are incomplete, while economic agents are heterogeneous, face idiosyncratic risks, and participate in risky asset markets to varying degrees. Even absent concerns about AI, these factors can generate precautionary demand for safe assets (Aiyagari, 1994) and can mean that risk-free assets are priced by different agents than risky assets (Vissing-Jørgensen, 2002; Guvenen, 2009). Under such circumstances, changes in bond yields may reflect distributional shifts rather than changes in aggregate consumption.

The implications of these factors for AI are explored by Maresca (2026), who develops a heterogeneous-agent model in which AI progress can raise aggregate growth while lowering the risk-free rate. The mechanism is that AI may shift wealth toward economic agents who invest in risky assets. If the agents who price the risk-free bond do not benefit from AI-related capital gains, or their labor income becomes less valuable, then their consumption may fall even as aggregate consumption rises, reconciling falling bond yields and rising aggregate output. The literature suggests the possibility of distributional impacts, with Acemoglu and Restrepo (2018) showing that automation can reduce the labor share and, during transitions, increase inequality. Krusell et al. (2000) show how capital-skill complementarity can account for changes in relative wages and inequality. Most closely related, Moll et al. (2022) study automation in a heterogeneous-agent model where automation affects income and wealth inequality through the return on wealth.[21]

Together, these results suggest that favorable AI news could lower Treasury yields, even while raising aggregate growth expectations, by changing beliefs about the future distribution of income and wealth and, in particular, lowering expected future consumption for the agents who price the risk-free asset.

4.3 Savings Glut

An alternative heterogeneous-agent explanation works through the supply of capital, under the empirically-supported assumption that wealthy households consume a smaller share of their income or wealth (Dynan et al., 2004; Ait-Sahalia et al., 2004; Saez and Zucman,

[21] Related concerns about the distributional impacts of AI have also motivated policy proposals, e.g. Sanders (2026).

2016). In models with non-homothetic preferences that imply this pattern, an increase in income or wealth at the top of the distribution can generate a large supply of savings which must be absorbed elsewhere, potentially leading to a fall in interest rates (Kumhof et al., 2015; Mian et al., 2025, 2021).

Similar to this perspective, Caballero (2026) models an AI boom where, building on prior work by Caballero et al. (2017), capital gains accrue disproportionately to households whose saving rate rises with wealth. Consequently, AI-related capital gains increase aggregate savings and lower interest rates, which in turn can help sustain high valuations and investment in AI.

Together, these results again suggest an alternative explanation for why positive AI news may lower Treasury yields, namely that such news increases the current and future wealth of investors, which in turn increases demand for safe assets. In contrast to the prior explanation, this one is compatible with broadly rising consumption expectations, and is in that sense more optimistic about the distributional implications of our empirical results.

4.4 Fiscal Backing and Changes in Treasury Risk

The preceding mechanisms treat Treasury or TIPS payoffs as risk-free. An alternative interpretation is that AI news changes the perceived riskiness of US government debt. Treasury yields could then fall because the fiscal-risk component of the yield declines, rather than because of changes in the SDF. This interpretation is consistent with the general logic of defaultable-bond pricing, in which bond yields incorporate expected losses and risk compensation (Eaton and Gersovitz, 1981; Duffie and Singleton, 1999; Arellano, 2008).

In the context of AI, Kung et al. (2026a,b) have explicitly proposed this channel as a potential explanation for falling bond yields in response to AI news. Their argument is that if favorable AI news increases the expected growth rate, this may in turn improve the expected fiscal position of the US government, and thus make US sovereign bonds safer. Their proposed mechanism is related to the fiscal reaction evidence in Bohn (1998), the government-budget-constraint valuation logic in Jiang et al. (2022), and recent work on fiscal redistribution risk and long-duration government claims (Gomez Cram et al., 2025; Miller et al., 2025).

Relative to the other explanations discussed above, our data offer more evidence on this point. Specifically, if yield changes were driven by changes in the perceived riskiness of US debt, we might expect sovereign CDS to shift around our AI news dates, contrary to what

we find in Figure 13 in the appendix.[22] Alternatively, if we thought that the risk of US corporate defaults were imperfectly correlated with the risk of a US sovereign default, we might expect positive AI news to lead to widening of corporate spreads, where Appendix Figure 12 again does not show evidence of this pattern. Instead, we observe a widening of spreads around our SLOWER dates. Interpreting these spread changes is delicate, however, since AI news might also affect the perceived riskiness of heavily AI-exposed US corporate bond issuers, masking changes in sovereign risk around FASTER dates or driving changes in spreads around SLOWER dates.

4.5 Other Mechanisms

An alternative mechanism, discussed in policy circles, works through changes in firms' investment demand. The canonical Q-theory mechanism suggests that good news about future productivity implies a higher expected return to capital, which in turn raises the desired investment and puts upward pressure on real rates (Jorgenson, 1963; Hayashi, 1982; Kurmann and Otrok, 2013).[23] This investment demand channel has been proposed as a reason for news about AI to increase interest rates (De Vere et al., 2026; Goodkind, 2026), contrary to what we observe in the data.

4.6 Interpreting the Directional Asymmetry

The preceding sections offer a range of candidate explanations for why bond yields might decline in response to positive news about the rate of AI progress. None of these explanations, however, suggests an asymmetry of the sort we observe in Section 3.2 where yields fall (relative to the placebo distribution) when AI progress is faster than expected, while not rising when AI progress is slower than expected.

One natural explanation is that the Metaculus timelines we consider are only a proxy for the AI-related information markets are responding to.[24] In particular, the question we

[22] Note that while the figure considers only the 5-year CDS, we observe declines in 5-year bond yields.

[23] In the other direction, with uncertain and irreversible investment, good news about future productivity can lower interest rates by leading investors to postpone investing, reducing current investment demand (Chetty, 2007). In this case, lower long rates are driven by declines in the term premium that follow an increase in demand for safe, long-term saving to finance the delayed investment.

[24] An alternative possibility is that we are losing information when we summarize the Metaculus forecast distribution by its median. An exploratory analysis suggests that this may not be the cause. Specifically, we find qualitatively similar results when we split dates based on the change in the 25th or 75th percentiles of the distribution, which are the other summary statistics readily available from the Metaculus data.

use asks when AI will meet a particular capabilities threshold, but does not speak to the capability level AI systems will ultimately reach. For further evidence on this point, we turn to other evidence from Metaculus. Many Metaculus questions that speak to AI capabilities or impacts beyond our baseline series (for instance, how long will elapse between the creation of weak AGI and superintelligence, Metaculus 2021, or when, if ever, gross world product will grow by more than 30% in a single year, Metaculus 2020a) unfortunately have many fewer participants than does our focal question.[25] One question with more participants than our baseline, however, is whether AI of human-level intelligence (HLI) will be created before 2040 (Metaculus, 2016).

The Metaculus series for this question varies up and down from the creation of the question until early 2022. From January 2022 the series increases steeply and near-monotonically through June 2022, and then again following the November 2022 release of ChatGPT. By fall 2023 the series is near-saturated, with forecast probability consistently exceeding 95%. Thus, while model releases in our sample led to both upward and downward updates about the rate of progress toward a particular definition of general artificial intelligence, no release saw a substantial decrease in the probability that HLI would be achieved by 2040.

Decomposing changes in the HLI series (again over a ± 15 day window) based on revisions of our focal AI series further clarifies the picture. Specifically, of the 13 SLOWER releases based on the general AI series the HLI probability decreased around 1 date and increased around 3 dates, with an average increase of 0.5 percentage points per release. By contrast, around the 17 FASTER releases the HLI probability decreased around 1 date and increased around 7 dates, with an average increase of 1.3 percentage points per release. Thus, from the perspective of whether HLI would be attained, the "bad" news dates in our analysis are not on average bad at all, but simply "less good." Saturation of the HLI series means these on-average differences are almost entirely driven by model releases in 2022-3 (hence the many zeros), but we nevertheless view this as suggestive evidence that the asymmetries observed in Section 3.2 may be due to aspects of AI progress beyond those captured in our primary Metaculus series.

[25] Specifically, 344 and 144 participants as of June 4, 2026, respectively, compared to over 1.9 thousand for our focal question.

5 Conclusion

We have shown that nominal Treasury and TIPS yields had statistically and economically significant declines around major AI model releases between the November 2022 release of ChatGPT and the end of 2025. These declines are concentrated around model releases which Metaculus forecasts suggest were positive surprises for the rate of technological progress. Taken together, these results suggest that investors take seriously the possibility of broad, economy-wide impacts from AI.

References

Acemoglu, Daron and Pascual Restrepo, "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

_ and _, "The Wrong Kind of AI? Artificial Intelligence and the Future of Labour Demand," *Cambridge Journal of Regions, Economy and Society*, 2020, *13* (1), 25–35.

_ and Todd Lensman, "Regulating Transformative Technologies," *American Economic Review: Insights*, 2024, *6* (3), 359–376.

_, David Autor, Jonathon Hazell, and Pascual Restrepo, "Artificial Intelligence and Jobs: Evidence from Online Vacancies," *Journal of Labor Economics*, 2022, *40* (S1), S293–S340.

Ait-Sahalia, Yacine, Jonathan A Parker, and Motohiro Yogo, "Luxury Goods and the Equity Premium," *The Journal of Finance*, 2004, *59* (6), 2959–3004.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Arellano, Cristina, "Default Risk and Income Fluctuations in Emerging Economies," *American Economic Review*, 2008, *98* (3), 690–712.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson, "Artificial Intelligence, Firm Growth, and Product Innovation," *Journal of Financial Economics*, 2024, *151*, 103745.

_, _, _, and _, "Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition," in Susanto Basu, Lucy Eldridge, John Haltiwanger, and Erich Strassner, eds., *Technology, Productivity, and Economic Growth*, Studies in Income and Wealth, University of Chicago Press, 2025, pp. 75–117.

Barro, Robert J., "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 2006, *121* (3), 823–866.

Björkegren, Daniel, "Market Beliefs about Open vs. Closed AI," 2026. Version 3, January 26, 2026.

Bloomberg L.P., "U.S. Sovereign Credit Default Swap Spreads," Bloomberg Terminal 2026. Accessed January 2026.

Board of Governors of the Federal Reserve System, US, "Market Yield on U.S. Treasury Securities at 1, 5, 10, 20, and 30-Year Constant Maturity, Quoted on an Investment Basis [DGS1, DGS5, DGS10, DGS20, DGS30]," FRED, Federal Reserve Bank of St. Louis 2025.

_ , "Market Yield on U.S. Treasury Securities at 5, 10, 20, and 30-Year Constant Maturity, Quoted on an Investment Basis, Inflation-Indexed [DFII5, DFII10, DFII20, DFII30]," FRED, Federal Reserve Bank of St. Louis 2025.

_ , "Nominal Broad U.S. Dollar Index [DTWEXBGS], Retrieved from FRED, Federal Reserve Bank of St. Louis," https://fred.stlouisfed.org/series/DTWEXBGS 2025.

Bohn, Henning, "The Behavior of U.S. Public Debt and Deficits," *Quarterly Journal of Economics*, 1998, *113* (3), 949–963.

Borio, Claudio, Piti Disyatat, Mikael Juselius, and Phurichai Rungcharoenkitkul, "Why So Low for So Long? A Long-Term View of Real Interest Rates," BIS Working Papers 685, Bank for International Settlements 2017.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press, 2019, pp. 23–60.

Caballero, Ricardo J., "Speculative Growth and the AI "Bubble"," Working Paper 34722, National Bureau of Economic Research 2026.

_ **, Emmanuel Farhi, and Pierre-Olivier Gourinchas**, "The Safe Assets Shortage Conundrum," *Journal of Economic Perspectives*, 2017, *31* (3), 29–46.

Cboe, "Cboe Volatility Index (VIX)," 2025. 8/28/2025.

Chetty, Raj, "Interest Rates, Irreversibility, and Backward-Bending Investment," *Review of Economic Studies*, 2007, *74* (1), 67–91.

Chiang, Wei-Lin, Lianmin Zheng, Ying Sheng, Anastasios Nikolas Angelopoulos, Tianle Li, Dacheng Li, Hao Zhang, Banghua Zhu, Michael Jordan, Joseph E. Gonzalez, and Ion Stoica, "Chatbot Arena: An Open Platform for Evaluating LLMs by Human Preference," 2024.

Chow, Trevor, Basil Halperin, and J. Zachary Mazlish, "Transformative AI, Existential Risk, and Real Interest Rates," Technical Report, Working Paper 2026. June 2026 version; first posted January 2023.

Citigroup Global Markets, "Citigroup Economic Surprise Index," Available via Bloomberg Terminal (ticker: CESIUSD Index) 2025.

Cram, Roberto Gomez, Howard Kung, Hanno Lustig, and David Zeke, "Fiscal Redistribution Risk in Treasury Markets," Working Paper 33769, National Bureau of Economic Research 2025.

Duffie, Darrell, *Dynamic Asset Pricing Theory*, 3 ed., Princeton University Press, 2010.

_ **and Kenneth J. Singleton**, "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies*, 1999, *12* (4), 687–720.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, "Do the Rich Save More?," *Journal of Political Economy*, 2004, *112* (2), 397–444.

Eaton, Jonathan and Mark Gersovitz, "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 1981, *48* (2), 289–309.

Eisfeldt, Andrea L., Gregor Schubert, and Miao Ben Zhang, "Generative AI and Firm Values," Working Paper 31222, National Bureau of Economic Research 2026. January 2026 revision.

Federal Reserve Bank of San Francisco, "Daily News Sentiment Index," `https://www.frbsf.org/research-and-insights/data-and-indicators/daily-news-sentiment-index/` 2023.

Financial Industry Regulatory Authority, "TRACE Treasury Aggregates: Daily Statistics," `https://www.finra.org/finra-data/browse-catalog/trace-treasury-aggregates/data` 2026.

Gabaix, Xavier, "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *Quarterly Journal of Economics*, 2012, *127* (2), 645–700.

Gil, Hamilton Galindo and Marius Mihai, "Asset Pricing Impacts of AI in a Model with Heterogeneous Risk Aversion," *SSRN Electronic Journal*, 2025. Available at SSRN 5277572; May 31, 2025 version.

Goodkind, Nicole, "Fed's Goolsbee Warns AI Could Produce Stagflation. "The Bigger the Hype, the Bigger the Concern."," `https://www.barrons.com/articles/fed-goolsbee-ai-stagflation-warning-e38a51e9` May 2026. Barron's, May 9, 2026.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, "The U.S. Treasury Yield Curve: 1961 to the Present," Finance and Economics Discussion Series 2006-28, Board of Governors of the Federal Reserve System 2006. Updated data available at `https://www.federalreserve.gov/data/nominal-yield-curve.htm`.

Gürkaynak, Refet S, Brian Sack, and Jonathan H Wright, "The TIPS Yield Curve and Inflation Compensation," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 70–92.

Guvenen, Fatih, "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica*, 2009, *77* (6), 1711–1750.

Hamilton, James D., Ethan S. Harris, Jan Hatzius, and Kenneth D. West, "The Equilibrium Real Funds Rate: Past, Present, and Future," *IMF Economic Review*, 2016, *64* (4), 660–707.

Hampole, Menaka, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Bryan Seegmiller, "Artificial Intelligence and the Labor Market," Working Paper 33509, National Bureau of Economic Research 2025.

Hansen, Bruce E and Ananth Seshadri, "Uncovering the Relationship between Real Interest Rates and Economic Growth," Technical Report, Working Paper 2013. Available at SSRN 2391449.

Hansen, Lars Peter and Kenneth J. Singleton, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 1983, *91* (2), 249–265.

Harris, Kenneth D., "A Shift Test for Independence in Generic Time Series," 2020.

Harrison, J. Michael and David M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 1979, *20* (3), 381–408.

Hayashi, Fumio, "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica*, 1982, *50* (1), 213–224.

ICE Data Indices, LLC, "ICE BofA 1-3, 3-5, 5-7, 7-10, 10-15, 15+ Year U.S. Corporate Index Option-Adjusted Spread [BAMLC1A0C13Y, BAMLC2A0C35Y, BAMLC3A0C57Y, BAMLC4A0C710Y, BAMLC7A0C1015Y, BAMLC8A0C15PY]," FRED, Federal Reserve Bank of St. Louis 2025.

Ice Data Indices, LLC, "ICE BofA 1-3, 3-5, 7-10, 10-15, 15+ Year U.S. Corporate Index Effective Yield, [BAMLC1A0C13YEY, BAMLC2A0C35YEY, BAMLC3A0C57YEY, BAMLC4A0C710YEY, BAMLC7A0C1015YEY, BAMLC8A0C15PYEY]," FRED, Federal Reserve Bank of St. Louis 2025.

Jackwerth, Jens Carsten and Mark Rubinstein, "Recovering Probability Distributions from Option Prices," *The Journal of Finance*, 1996, *51* (5), 1611–1631.

Jiang, Zhengyang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan, "Measuring U.S. Fiscal Capacity Using Discounted Cash Flow Analysis," *Brookings Papers on Economic Activity*, 2022, *53* (2), 157–209.

Jones, Charles I., "The AI Dilemma: Growth versus Existential Risk," *American Economic Review: Insights*, 2024, *6* (4), 575–590.

Jorgenson, Dale W., "Capital Theory and Investment Behavior," *American Economic Review*, 1963, *53* (2), 247–259.

Kokotajlo, Daniel, Scott Alexander, Thomas Larsen, Eli Lifland, and Romeo Dean, "AI 2027," April 2025. April 3, 2025.

Korinek, Anton and Donghyun Suh, "Scenarios for the Transition to AGI," Working Paper 32255, National Bureau of Economic Research 2024.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Kumhof, Michael, Romain Rancière, and Pablo Winant, "Inequality, Leverage, and Crises," *American Economic Review*, 2015, *105* (3), 1217–1245.

Kung, Howard, Hanno Lustig, and James D. Paron, "U.S. Treasury Investors' Massive Bet on AI," The Two Cents April 2026. Substack post, April 24, 2026.

_ , _ , and James Paron, "U.S. Treasury Investors Are Long in AI," May 2026. Working paper, May 14, 2026.

Kurmann, André and Christopher Otrok, "News Shocks and the Slope of the Term Structure of Interest Rates," *American Economic Review*, 2013, *103* (6), 2612–2632.

LM Arena, "Chatbot Arena Leaderboard," HuggingFace, `lmarena-ai/lmarena-leaderboard` 2024.

Lucas, Robert E., "Asset Prices in an Exchange Economy," *Econometrica*, 1978, *46* (6), 1429–1445.

Lucca, David O. and Emanuel Moench, "The Pre-FOMC Announcement Drift," *Journal of Finance*, 2015, *70* (1), 329–371.

Lunsford, Kurt G. and Kenneth D. West, "Some Evidence on Secular Drivers of U.S. Safe Real Rates," *American Economic Journal: Macroeconomics*, 2019, *11* (4), 113–139.

Maresca, Caleb, "AGI Could Lower Interest Rates," *SSRN Electronic Journal*, 2026. Available at SSRN 6514781; April 3, 2026 version.

Mehra, Rajnish, "Consumption-Based Asset Pricing Models," *Annual Review of Financial Economics*, 2012, *4* (1), 385–409.

Metaculus, "Will There Be Human-Machine Intelligence Parity before 2040?," `https://www.metaculus.com/questions/384/` 2016.

_ , "In Which Year Will the World's Real GDP First Exceed 130% of Its Highest Level from Any Previous Year?," `https://www.metaculus.com/questions/5159/date-of-gwp-growth-over-130/` 2020.

_ , "When Will the First General AI System Be Devised, Tested, and Publicly Announced?," `https://www.metaculus.com/questions/5121/date-of-artificial-general-intelligence/` 2020.

_ , "After a (Weak) AGI Is Created, How Many Months Will It Be before the First Superintelligent AI Is Created?," https://www.metaculus.com/questions/9062/time-from-weak-agi-to-superintelligence/ 2021.

METR, "METR's GPT-4.5 Pre-Deployment Evaluations," https://metr.org/blog/2025-02-27-gpt-4-5-evals/ 2025.

Mian, Atif, Ludwig Straub, and Amir Sufi, "Indebted Demand," *Quarterly Journal of Economics*, 2021, *136* (4), 2243–2307.

Mian, Atif R., Ludwig Straub, and Amir Sufi, "The Saving Glut of the Rich," Working Paper 26941, National Bureau of Economic Research 2025. July 2025 revision.

Miller, Max, James D. Paron, and Jessica A. Wachter, "Sovereign Default and the Decline in Interest Rates," Working Paper 34021, National Bureau of Economic Research 2025. July 2025 version.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, "Uneven Growth: Automation's Impact on Income and Wealth Inequality," *Econometrica*, 2022, *90* (6), 2645–2683.

Mollick, Ethan, "Google's Gemini Advanced: Tasting Notes and Implications," 2024.

Rogoff, Kenneth S, Barbara Rossi, and Paul Schmelzing, "Long-Run Trends in Long-Maturity Real Rates, 1311-2022," *American Economic Review*, 2024, *114* (8), 2271–2307.

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics*, 2016, *131* (2), 519–578.

Sanders, Bernie, "The Public Should Own Half of the Big A.I. Companies," https://www.sanders.senate.gov/op-eds/the-public-should-own-half-of-the-big-a-i-companies/ 2026. Op-ed.

Savor, Pavel and Mungo Wilson, "How Much Do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcements," *Journal of Financial and Quantitative Analysis*, 2013, *48* (2), 343–375.

Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson, "Measuring News Sentiment," *Journal of Econometrics*, 2022, *228* (2), 221–243.

S&P Dow Jones Indices LLC, "S&P 500 [SP500], Retrieved from FRED, Federal Reserve Bank of St. Louis," FRED, Federal Reserve Bank of St. Louis 2026.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Trammell, Philip and Anton Korinek, "Economic Growth under Transformative AI," *Annual Review of Economics*, 2026. Published online June 2, 2026.

UK Artificial Intelligence Safety Institute and U.S. Artificial Intelligence Safety Institute, "Pre-Deployment Evaluation of OpenAI's o1 Model," `https://www.aisi.gov.uk/blog/pre-deployment-evaluation-of-openais-o1-model` December 2024. December 18, 2024.

van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, "On the Timing and Pricing of Dividends," *American Economic Review*, 2012, *102* (4), 1596–1618.

_ , Wouter Hueskes, Ralph Koijen, and Evert Vrugt, "Equity Yields," *Journal of Financial Economics*, 2013, *110* (3), 503–519.

Vere, Hugo De, Srini Ramaswamy, and Seth Searls, "How AI Debt Financing Impacts Duration Supply and Interest Rates," Dallas Fed Economics February 2026. February 10, 2026.

Vissing-Jørgensen, Annette, "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 2002, *110* (4), 825–853.

Wachter, Jessica and Jonathan Wachter, "What Investment Data Implies About the AI Transition," Working Paper 35290, National Bureau of Economic Research 2026.

Webb, Michael, "The Impact of Artificial Intelligence on the Labor Market," *SSRN Electronic Journal*, 2020. Available at SSRN 3482150; January 2020 revision.

Wolfers, Justin and Eric Zitzewitz, "Prediction Markets," *Journal of Economic Perspectives*, 2004, *18* (2), 107–126.

Yahoo Finance, "NVIDIA Corporation (NVDA) Historical Data," Yahoo Finance 2026.

Yuan, Alex E. and Wenying Shou, "A Rigorous and Versatile Statistical Test for Correlations Between Stationary Time Series," *PLOS Biology*, 2024, *22* (8), e3002758.

A Representative Agent Asset Pricing Models

This appendix provides the assumptions, derivations, and empirical results underlying the discussion in Section 4.1.

A.1 Representative-Agent Model of Transformative AI and Bond Prices

The environment follows the standard consumption-based asset-pricing logic summarized in Duffie (2010), adapted to the possibility that transformative AI may either substantially change the growth path of the economy or lead to a date beyond which ordinary asset payoffs cease to matter. Specifically, following Jones (2024) and Chow et al. (2026), we allow AI to affect beliefs about both future aggregate consumption growth and a random date T after which asset holdings are irrelevant.

Again consider a discrete-time economy with dates $t=0,1,...,\bar{T}$. For simplicity, suppose there are finitely many states and finitely many agents. Agent i has time-separable utility

$$\mathbb{E}_0\left[\sum_{t=1}^{\bar{T}}\beta^t\big(1\{t\leq T\}u_i(C_{i,t})+1\{t>T\}U^*_{i,t}\big)\right],$$

where $T\leq\bar{T}$ is the random date after which asset holdings are irrelevant for utility (that is, we assume $U^*_{i,t}$ is independent of asset holdings). We further assume that u_i is concave with $u_i'(c)\to 0$ as $c\to\infty$.

Assume complete markets and absence of arbitrage. Then as discussed in Section 2.1 there exists a unique SDF that prices all assets. Recall that for simplicity, we write $M_{t+1}=M_{t,t+1}$ and $M_{t,t+h}=\prod_{s=1}^{h}M_{t+s}$. Complete markets imply that in equilibrium the SDF coincides with the marginal rate of substitution of a representative agent with utility

$$\mathbb{E}_0\left[\sum_{t=1}^{T}\beta^t u(C_t)\right],$$

where $C_t=\sum_i C_{i,t}$ is aggregate consumption and $u(C_t)=\sum_i\lambda_i u_i(C_{i,t})$ for Pareto weights $\lambda_i\geq 0$. Hence the SDF may be written as

$$M_{t,t+h}=\beta^h\frac{u'(C_{t+h})}{u'(C_t)}1\{t+h\leq T\}. \tag{7}$$

Equivalently, for every agent

$$M_{t,t+h} = \beta^h \frac{u_i'(C_{i,t+h})}{u_i'(C_{i,t})} 1\{t+h \leq T\}.$$

Equation (7) has two immediate implications. First, the flow of utility after T is irrelevant for asset prices. Thus, whether T corresponds to the arrival of an existential disaster or an infinite-consumption singularity, it has the same asset-pricing implications under the model. Second, higher future aggregate consumption lowers future marginal utility and lowers the SDF. Therefore, news that raises anticipated consumption growth tends, all else equal, to lower bond prices and raise yields. Conversely, news that lowers anticipated consumption growth tends to raise bond prices and lower yields.

To see the bond-pricing implications, again let 1_{t+h} denote a risk-free zero-coupon bond paying one unit of consumption at date $t+h$. By (1),

$$V_t(1_{t+h}) = \mathbb{E}_t[M_{t,t+h}] = \mathbb{E}_t\left[\beta^h \frac{u'(C_{t+h})}{u'(C_t)} 1\{t+h \leq T\}\right].$$

As in (2), the corresponding zero-coupon yield is

$$y_{t,t+h} \equiv V_t(1_{t+h})^{-1/h} - 1 = \mathbb{E}_t[M_{t,t+h}]^{-1/h} - 1.$$

Using the law of iterated expectations,

$$y_{t,t+h} = \frac{1}{\beta \mathbb{P}_t(t+h \leq T)^{1/h} \mathbb{E}_t\left[\frac{u'(C_{t+h})}{u'(C_t)} \,|\, t+h \leq T\right]^{1/h}} - 1 \tag{8}$$

Thus, zero-coupon yields are decreasing in the discount factor β, increasing in the probability that T arrives before the bond pays off, and decreasing in expected future marginal utility conditional on T not yet having arrived. Since u is concave, higher anticipated consumption growth tends to increase yields, while lower anticipated growth tends to decrease yields. Thus, as discussed in the main text this model implies that the yield declines we observe around AI model releases are consistent with e.g. a lower arrival probability for T or lower consumption expectations.

A.2 Simplified Representative Agent Model with Growth Uncertainty

The general representative-agent model discussed above suggests some possible explanations for our results but (a) does not distinguish between these explanations and (b) is purely qualitative. In this section we impose stronger assumptions which yield further traction in separating and quantifying these possible channels. Our assumptions in this appendix are restrictive, but deliver transparent empirical implications.

Assume that the representative agent in Appendix A.1 has CRRA flow utility,

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma},$$

so that the SDF is

$$M_{t,t+h} = \beta^h \left(\frac{C_{t+h}}{C_t} \right)^{-\gamma} 1\{t+h \leq T\}. \tag{9}$$

Suppose further that there exists a horizon $k \geq 0$ such that, for all $h \geq k$, the representative agent believes aggregate consumption evolves according to

$$C_{t+h+1} = (1+g) X_{t+h+1} C_{t+h}, \tag{10}$$

where g captures the long-run consumption-growth effect of AI and $\{X_s\}_{s=t+k+1}^{\bar{T}}$ captures non-AI determinants of consumption growth. The horizon k allows the long-term growth effects of AI to begin only after an initial transition period. Combining CRRA utility with (10), for $h \geq k$ the h-period-ahead SDF can be written as

$$M_{t,t+h} = \left(\frac{C_{t+k}}{C_t} \right)^{-\gamma} \beta^h (1+g)^{-(h-k)\gamma} \left(\prod_{s=k+1}^{h} X_{t+s} \right)^{-\gamma} 1\{t+h \leq T\}.$$

We further impose three restrictions conditional on information available at each observed date t and on $t+k \leq T$. First, the representative agent thinks that the process $\{X_s\}_{s=t+k+1}^{t+h}$, the term $(C_{t+k}/C_t)^{-\gamma}$, the date T, and the growth component g are mutually independent. Second, the representative agent believes that T arrives with probability δ_t in each period after $t+k$, so that

$$\mathbb{P}_t(t+h \leq T \,|\, t+k \leq T) = \prod_{s=k+1}^{h} \mathbb{P}_t(t+s \leq T \,|\, t+s-1 \leq T) = (1-\delta_t)^{h-k};$$

Third, $1+g$ is lognormally distributed under the representative agent's belief,

$$\log(1+g)\,|\,\mathcal{F}_t, t+k\leq T \sim N(\mu_t,\sigma_t^2).$$

Let $f_{t+k,t+h}$ denote the (gross) date-t forward yield from $t+k$ to $t+h$, defined by

$$f_{t+k,t+h}=\left(\frac{(1+y_{t,t+h})^h}{(1+y_{t,t+k})^k}\right)^{1/(h-k)}. \tag{11}$$

The next lemma (whose proof is deferred to the end of this appendix) provides a closed form for $\log(f_{t+k,t+h})$ under our assumptions.[26]

Lemma 1. *Under the above assumptions,*

$$\begin{aligned}\log(f_{t+k,t+h})=&-\log(\beta)-\log(1-\delta_t)+\gamma\mu_t-\frac{\gamma^2}{2}(h-k)\sigma_t^2\\&-\frac{1}{h-k}\log\left(\mathbb{E}_t\left[\frac{(C_{t+k}/C_t)^{-\gamma}}{\mathbb{E}_t[(C_{t+k}/C_t)^{-\gamma}]}\left(\prod_{s=k+1}^{h}X_{t+s}\right)^{-\gamma}\right]\right).\end{aligned} \tag{12}$$

To eliminate the term involving X_t, we employ a difference-in-differences strategy. Let $\mathcal{T}$ denote the set of AI event dates and $\mathcal{A}$ the set of all dates for which the pre- and post-event windows are observed. For each t, define $t_-=t-b$ and $t_+=t+s$, and let

$$\eta_{t_-,t_+,k,h}=\log\left(\frac{\mathbb{E}_{t_+}\left[\left(\frac{C_{t_++k}}{C_{t_+}}\right)^{-\gamma}\left(\prod_{s=k+1}^{h}X_{t_++s}\right)^{-\gamma}\right]}{\mathbb{E}_{t_-}\left[\left(\frac{C_{t_-+k}}{C_{t_-}}\right)^{-\gamma}\left(\prod_{s=k+1}^{h}X_{t_-+s}\right)^{-\gamma}\right]}\cdot\frac{\mathbb{E}_t\left[\left(\frac{C_{t_-+k}}{C_{t_-}}\right)^{-\gamma}\right]}{\mathbb{E}_t\left[\left(\frac{C_{t_++k}}{C_{t_+}}\right)^{-\gamma}\right]}\right).$$

If we difference the log forward yields at two dates $t_-<t<t_+$ we thus have

$$\log\big(f_{t_++k,t_++h}\big)-\log\big(f_{t_-+k,t_-+h}\big)=$$

$$-\log\left(\frac{1-\delta_{t_+}}{1-\delta_{t_-}}\right)+\gamma\big(\mu_{t_+}-\mu_{t_-}\big)-\frac{\gamma^2}{2}(h-k)\big(\sigma_{t_+}^2-\sigma_{t_-}^2\big)-\eta_{t_-,t_+,k,h}$$

We assume that $\eta_{t_-,t_+,k,h}$ has approximately the same average around AI event dates as

[26]These assumptions are restrictive, and appear unlikely to hold exactly. For instance, one might expect that more effective AI (i.e. AI yielding a higher g) would be associated with a closer arrival date for T. Similarly, if the growth effects of AI may "kick in" strictly before period $t+k$ then a higher g should lead to a higher C_{t+k}.

around dates in $\mathcal{A}$,

$$\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\eta_{t_-,t_+,k,h}\approx\frac{1}{|\mathcal{A}|}\sum_{t\in\mathcal{A}}\eta_{t_-,t_+,k,h}. \tag{13}$$

For instance, if we assumed that $\eta_{t_-,t_+,k,h}$ were stationary across time conditional on our event dates $\mathcal{T}$ and regularity conditions held, this would follow from the law of large numbers as $|\mathcal{T}|\to\infty$.

Motivated by this approximation, define the difference-in-differences operator

$$\mathrm{DID}(\log(f_{t+k,t+h});\mathcal{T},\mathcal{A})=\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\log\left(\frac{f_{t_++k,t_++h}}{f_{t_-+k,t_-+h}}\right)-\frac{1}{|\mathcal{A}|}\sum_{t\in\mathcal{A}}\log\left(\frac{f_{t_++k,t_++h}}{f_{t_-+k,t_-+h}}\right). \tag{14}$$

Using (12), the slope of (14) with respect to $h-k$ is approximately

$$-\frac{\gamma^2}{2}\mathrm{DID}(\sigma_t^2;\mathcal{T},\mathcal{A})\approx-\frac{\gamma^2}{2}\left(\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}(\sigma^2_{t_+}-\sigma^2_{t_-})-\frac{1}{|\mathcal{A}|}\sum_{t\in\mathcal{A}}(\sigma^2_{t_+}-\sigma^2_{t_-})\right). \tag{15}$$

An increase in growth uncertainty around AI releases relative to other dates therefore corresponds to a negative slope in the forward-rate difference-in-differences. Similarly, the intercept of (14) as $h\downarrow k$ is equal to

$$\gamma\mathrm{DID}(\mu_t;\mathcal{T},\mathcal{A})-\mathrm{DID}(\log(1-\delta_t);\mathcal{T},\mathcal{A}), \tag{16}$$

where

$$\mathrm{DID}(\mu_t;\mathcal{T},\mathcal{A})=\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}(\mu_{t_+}-\mu_{t_-})-\frac{1}{|\mathcal{A}|}\sum_{t\in\mathcal{A}}(\mu_{t_+}-\mu_{t_-}),$$
$$\mathrm{DID}(\log(1-\delta_t);\mathcal{T},\mathcal{A})=\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\log\left(\frac{1-\delta_{t_+}}{1-\delta_{t_-}}\right)-\frac{1}{|\mathcal{A}|}\sum_{t\in\mathcal{A}}\log\left(\frac{1-\delta_{t_+}}{1-\delta_{t_-}}\right).$$

Thus, under our assumptions the intercept reflects changes in expected growth and changes in the arrival probability of T, while the slope isolates the growth-uncertainty changes.

Taking the model to the data. The simplified model predicts the behavior of yields on risk-free zero-coupon bonds. We therefore apply the forward-rate transformation to the zero-coupon TIPS yields used in the empirical analysis. The event set $\mathcal{T}$ is chosen as in Section 3, and we consider ±30-day event windows. To choose the horizon k beyond which AI growth effects "kick in" we use the observation that, under the model, the difference-in-

differences based on the one-period-ahead forward curve should be linear beyond horizon k. As shown in the top panel of Figure 5, this appears approximately true for $k \geq 10$, so we use $k = 10$ in our analysis. Fixing this choice, we regress $\text{DID}(\log(f_{t+k,t+h}); \mathcal{T}, \mathcal{A})$ on $h-k$ for $h \in \{11,...,20\}$. This yields a slope of 0.0003 log points and an intercept of -0.098 log points, corresponding respectively to the 38.4 and 1.3 percentiles of the permutation distribution. Thus, we find strong evidence of a level shift in the forward curve on average around our event dates, but negligible evidence of a slope change.

From the intercept to magnitudes. Because the estimated slope is statistically indistinguishable from its placebo distribution, we focus on the intercept in (16). Two limiting cases bound the implied magnitudes. If the intercept reflects expected growth alone, then it equals $\gamma \text{DID}(\mu_t; \mathcal{T}, \mathcal{A})$. If we assume σ_t is small for all t and μ_t is close to zero, then $\text{DID}(\mu_t; \mathcal{T}, \mathcal{A})$ is approximately equal to the DID in expected growth, $\text{DID}(\mathbb{E}_t[g]; \mathcal{T}, \mathcal{A})$, so under $\gamma = 2$ the average release is associated with a roughly 0.049 percentage point decline in expected annual consumption growth, or a 1.47 percentage point decline cumulated over the 30 releases. If instead the intercept reflects the arrival probability alone, then assuming small δ_t the approximation $\log(1-\delta_t) \approx -\delta_t$ makes the intercept approximately equal to $\text{DID}(\delta_t; \mathcal{T}, \mathcal{A})$, so the average release is associated with a 0.098 percentage point decline in the annual arrival probability of T, or a 2.93 percentage point decline cumulated over the 30 releases in our sample.

Proof of Lemma 1. By the law of iterated expectations and the conditional independence assumptions above,

$$\begin{aligned}
\mathbb{E}_t[M_{t,t+h}] &= \mathbb{P}_t(t+k \leq T)\mathbb{E}_t[M_{t,t+h} \,|\, t+k \leq T] \\
&= \mathbb{P}_t(t+k \leq T)\beta^h \mathbb{E}_t\left[\left(\frac{C_{t+k}}{C_t}\right)^{-\gamma}\left(\prod_{s=k+1}^{h} X_{t+s}\right)^{-\gamma} | \, t+k \leq T\right] \\
&\quad \times \mathbb{E}_t\left[(1+g)^{-(h-k)\gamma} \,|\, t+k \leq T\right]\mathbb{P}_t(t+h \leq T \,|\, t+k \leq T).
\end{aligned}$$

Since $\log(1+g) \,|\, \mathcal{F}_t, t+k \leq T \sim N(\mu_t, \sigma_t^2)$,

$$\mathbb{E}_t\left[(1+g)^{-(h-k)\gamma} \,|\, t+k \leq T\right] = \exp\left(-\gamma(h-k)\mu_t + \frac{\gamma^2(h-k)^2\sigma_t^2}{2}\right).$$

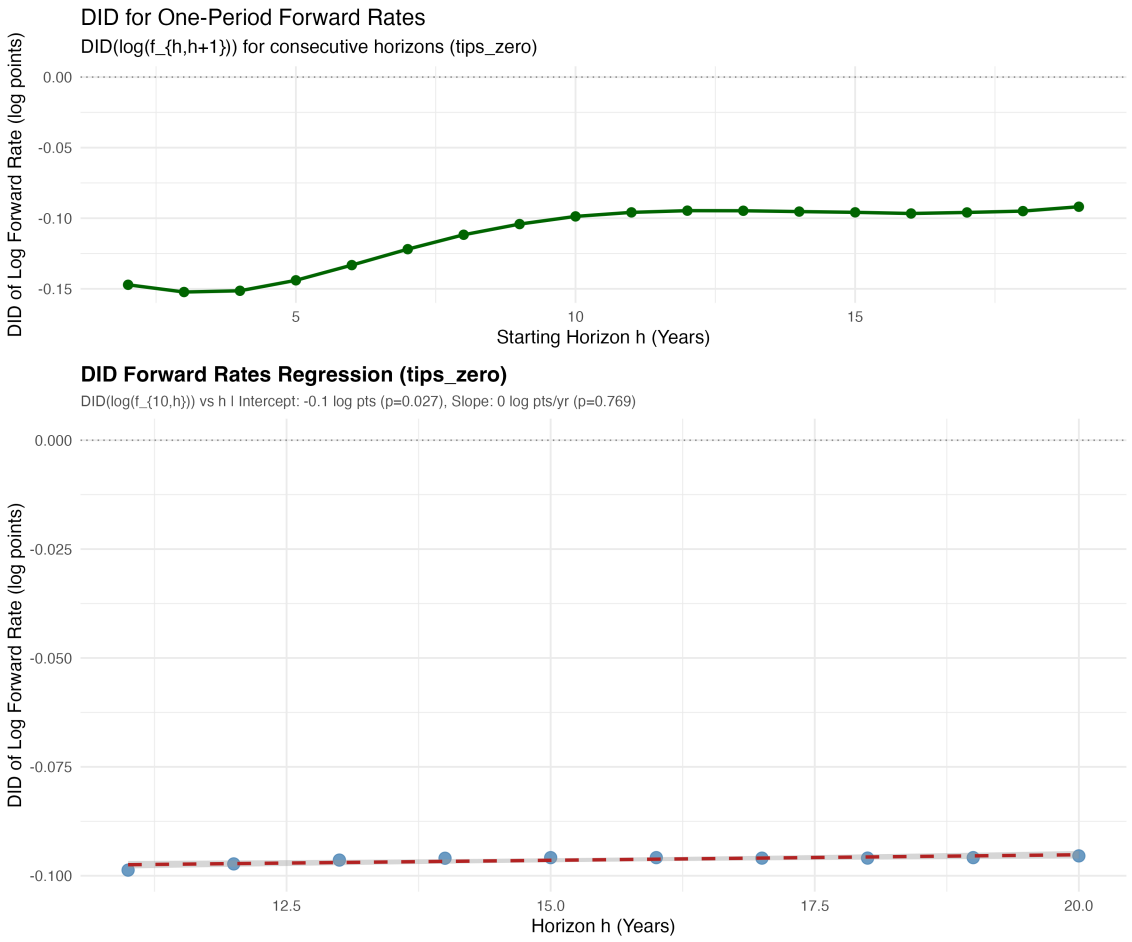

Figure 5: Difference-in-differences of TIPS zero-coupon log forward yields around AI model releases. The top panel plots the difference-in-differences of one-period log forward rates $\mathrm{DID}(\log f_{t+h,t+h+1};\mathcal{T},\mathcal{A})$ across starting horizons h. The bottom panel plots the difference-in-differences of log forward rates $\mathrm{DID}(\log f_{t+k,t+h};\mathcal{T},\mathcal{A})$ against the horizon h (with $k=10$), together with the fitted line, whose slope and intercept as $h\downarrow k$ are reported in the text. Sample: November 2022 through December 2025.

Moreover,

$$\mathbb{P}_t(t+h\leq T\,|\,t+k\leq T)=(1-\delta_t)^{h-k}.$$

Therefore,

$$\mathbb{E}_t[M_{t,t+h}] = \beta^h(1-\delta_t)^{h-k}\exp\left(-\gamma(h-k)\mu_t + \frac{\gamma^2(h-k)^2\sigma_t^2}{2}\right)\mathbb{P}_t(t+k\leq T) \times \mathbb{E}_t\left[\left(\frac{C_{t+k}}{C_t}\right)^{-\gamma}\left(\prod_{s=k+1}^{h} X_{t+s}\right)^{-\gamma}\right]. \tag{17}$$

Taking the difference between $\log(\mathbb{E}_t[M_{t,t+h}])$ and $\log(\mathbb{E}_t[M_{t,t+k}])$ gives

$$(h-k)\log(\beta)+(h-k)\log(1-\delta_t)-\gamma(h-k)\mu_t+\frac{\gamma^2}{2}(h-k)^2\sigma_t^2 + \log\left(\frac{\mathbb{E}_t\left[(C_{t+k}/C_t)^{-\gamma}\left(\prod_{s=k+1}^{h} X_{t+s}\right)^{-\gamma}\right]}{\mathbb{E}_t[(C_{t+k}/C_t)^{-\gamma}]}\right).$$

Since $y_{t,t+h} = \mathbb{E}_t[M_{t,t+h}]^{-1/h} - 1$, the forward yield satisfies

$$\begin{aligned}\log(f_{t+k,t+h}) &= \frac{\log(\mathbb{E}_t[M_{t,t+k}]) - \log(\mathbb{E}_t[M_{t,t+h}])}{h-k} \\ &= -\log(\beta) - \log(1-\delta_t) + \gamma\mu_t - \frac{\gamma^2}{2}(h-k)\sigma_t^2 \\ &\quad - \frac{1}{h-k}\log\left(\mathbb{E}_t\left[\frac{(C_{t+k}/C_t)^{-\gamma}}{\mathbb{E}_t[(C_{t+k}/C_t)^{-\gamma}]}\left(\prod_{s=k+1}^{h} X_{t+s}\right)^{-\gamma}\right]\right),\end{aligned}$$

which is (12).

□

B Additional Details on Date Series

B.1 LMArena Ranking Frequencies

We base our lab list on the Arena (formerly LMArena) leaderboard (LM Arena, 2024), which ranks large language models by aggregated human pairwise preference and updates as new models are added. We download historical snapshots of the Arena ranking from a GitHub repository (LM Arena, 2024) covering a selection of dates from May 2023 through August 2025. For each lab that ever had a model appear in the non-style corrected rankings

Organization	% in Top 5	% in Top 10	Best Rank
OpenAI	100.0	100.0	1
Google	69.6	82.6	1
Anthropic	47.8	62.9	1
xAI	47.1	76.6	1
LMSYS	5.9	23.9	4
Mistral AI	5.8	18.8	2
DeepSeek	4.1	23.7	5
Other	3.3	24.1	5
Meta	3.0	36.3	2
Alibaba	0.5	13.0	4
Microsoft	0.0	21.3	6
01.AI	0.0	15.1	7

Table 6: Rankings based on Arena (formerly LMArena) leaderboard data from May 2023 through August 2025. "% in Top 5/10" is the percentage of snapshots where the organization had at least one model in the top 5/10, among those snapshots where they had any model. "Best Rank" is the highest rank achieved by any model from the organization across the sample of snapshots.

(which we use since the style-corrected rankings are only available from August 2024) we compute the percentage of snapshot dates (among those for which the lab has any model in the ranking at all) where one or more of the lab's models appears in the top 5 or top 10. We also compute the top ranking ever observed for one of the lab's models.

OpenAI, Google, Anthropic, and xAI each have at least one model in the top ten in more than 60% of leaderboard snapshots, and a model in the top 5 for over 40%, far higher than the next highest lab. Consequently, we view model releases from these four labs as plausibly informative about the rate of frontier AI progress.

B.2 Event Window Overlap

Given that our main analysis sample considers 30 model release dates in a period of 38 months, and our leading specifications use a ± 30 trading day window around each release, there is naturally some overlap in our event windows, with a higher degree of overlap in 2025 given the large number of releases in that year. To illustrate the extent of overlap, as well as the variation over time, Figure 6 plots black lines for the event dates, and shaded regions for the ± 15 day event window (in the first panel) and the ± 30 day event window (in the second panel). The overlap between these event windows, particularly in 2025,

motivates the regression specification in Section 2.4. For comparison, each panel also shows the alternative event series discussed in Section 3.4.3 of the text and Appendix D.3.

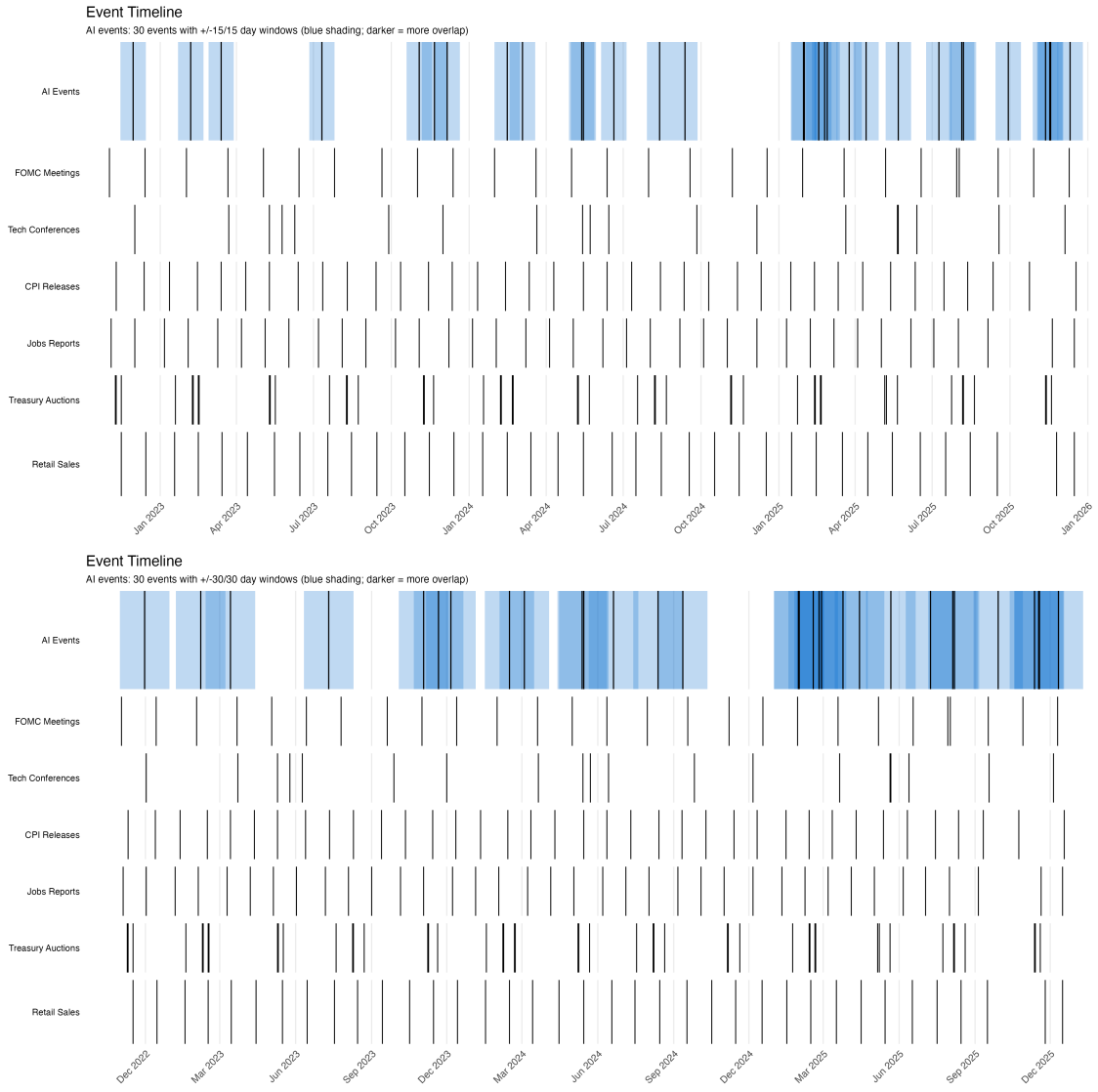

Figure 6: Event-window coverage at the ±15 (top) and ±30 (bottom) trading-day windows. In each panel the top row shades each release's event window, with darker shading marking calendar days covered by more overlapping windows; the lower rows mark the alternative date series (FOMC meetings, "Magnificent Seven" conference dates, CPI and employment releases, retail sales reports, and Treasury auctions) used as placebo sources in Appendix D. Sample: November 2022 through December 2025.

Table 7: AI Model Release Dates – Additional Releases in Extended Sample (with announcement URLs)

Date	Lab	Model
2023 releases		
02/24/2023	Meta	Llama
07/18/2023	Meta	Llama 2
2024 releases		
04/18/2024	Meta	Llama 3
05/06/2024	DeepSeek	DeepSeek V2
07/23/2024	Meta	Llama 3.1
09/05/2024	DeepSeek	DeepSeek V2.5
09/25/2024	Meta	Llama 3.2
11/20/2024	DeepSeek	DeepSeek R1-Lite
12/06/2024	Meta	Llama 3.3
12/26/2024	DeepSeek	DeepSeek V3
2025 releases		
01/20/2025	DeepSeek	DeepSeek R1
04/05/2025	Meta	Llama 4
08/21/2025	DeepSeek	DeepSeek V3.1
12/01/2025	DeepSeek	DeepSeek V3.2

Notes: Additional event dates included in the extended sample beyond the four-lab main sample of Table 1. These 14 Meta and DeepSeek releases combine with the 30 main-sample dates to give 44 unique extended-sample event dates.

B.3 Extended-Sample Event Study

Table 7 lists the flagship model release dates for Meta and DeepSeek in our analysis sample. Combined with the dates listed in Table 1, these give 44 unique extended-sample event dates. Figure 7 reports results, parallel to those in Figure 4, for this extended event list. The qualitative pattern from the main-sample analysis carries over: the FASTER coefficients decline relative to the placebo mean (albeit with less statistical significance), while the SLOWER coefficients remain closer to the placebo mean and statistically insignificant.

Figure 7: Regression-based event study for zero-coupon Treasury yields (top panel) and zero-coupon TIPS yields (bottom panel) using the extended event list of Table 7. Events are split by the direction of Metaculus prediction changes over a ± 15 trading-day window around each release. Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025; 44 unique event dates.

C Additional Data Series

This appendix reports results for additional financial series beyond the zero-coupon Treasury and TIPS yields analyzed in the main text. All results use the regression specification (5) with $b = s = 30$ trading-days, split by Metaculus prediction direction, with 5000 year-stratified permutation draws for inference. Data sources are discussed in the main text.

C.1 End-Horizon Long Differences by Maturity (Split by Direction)

Figures 8, 9, and 10 report the end-horizon cumulative change in zero-coupon Treasury, TIPS, and inflation breakeven yields, split by Metaculus direction, with maturity on the horizontal axis. The top panel of each figure gives the SLOWER results. Each point is the cumulated regression coefficient at day $+30$ relative to day -30, and the dashed lines show 10%, 5%, and 1% two-sided permutation critical values.

C.2 Other Series (Event-Study Time-Series, Split by Direction)

Figures 11, 12, 13, and 14 report the cumulative regression coefficients from (5) traced from $l = -30$ to $+30$ trading-days, separately for the SLOWER and FASTER subsets, for the remaining series we examine.

D Robustness Checks

This appendix reports the robustness checks summarized in Section 3.4. All checks re-estimate the joint regression (5) on the main sample of 30 release dates with the Metaculus SLOWER vs. FASTER split.

D.1 Dropping Events

Figure 15 reports cumulative event-study paths for Treasury zero-coupon yields obtained by dropping all subsets of one (top panel) or two (bottom panel) event dates from each Metaculus direction subset, plotted in green against the full-sample paths in blue. Figure 16 reports the analogous paths when dropping all subsets of three event dates.

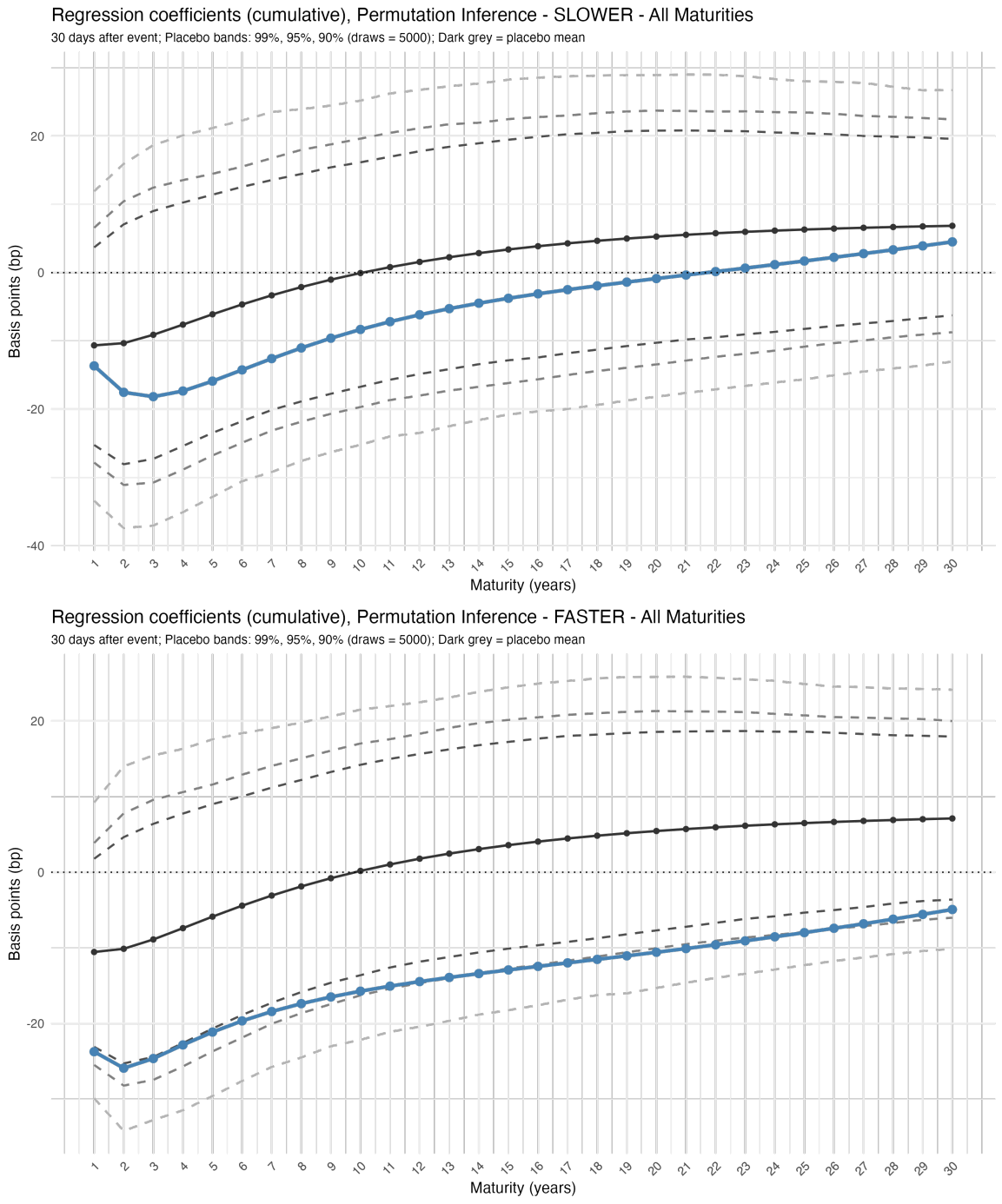

Figure 8: End-horizon ($l = 30$) change in zero-coupon Treasury yields by maturity, split by Metaculus prediction direction. Top panel: SLOWER coefficients (13 events). Bottom panel: FASTER coefficients (17 events). Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

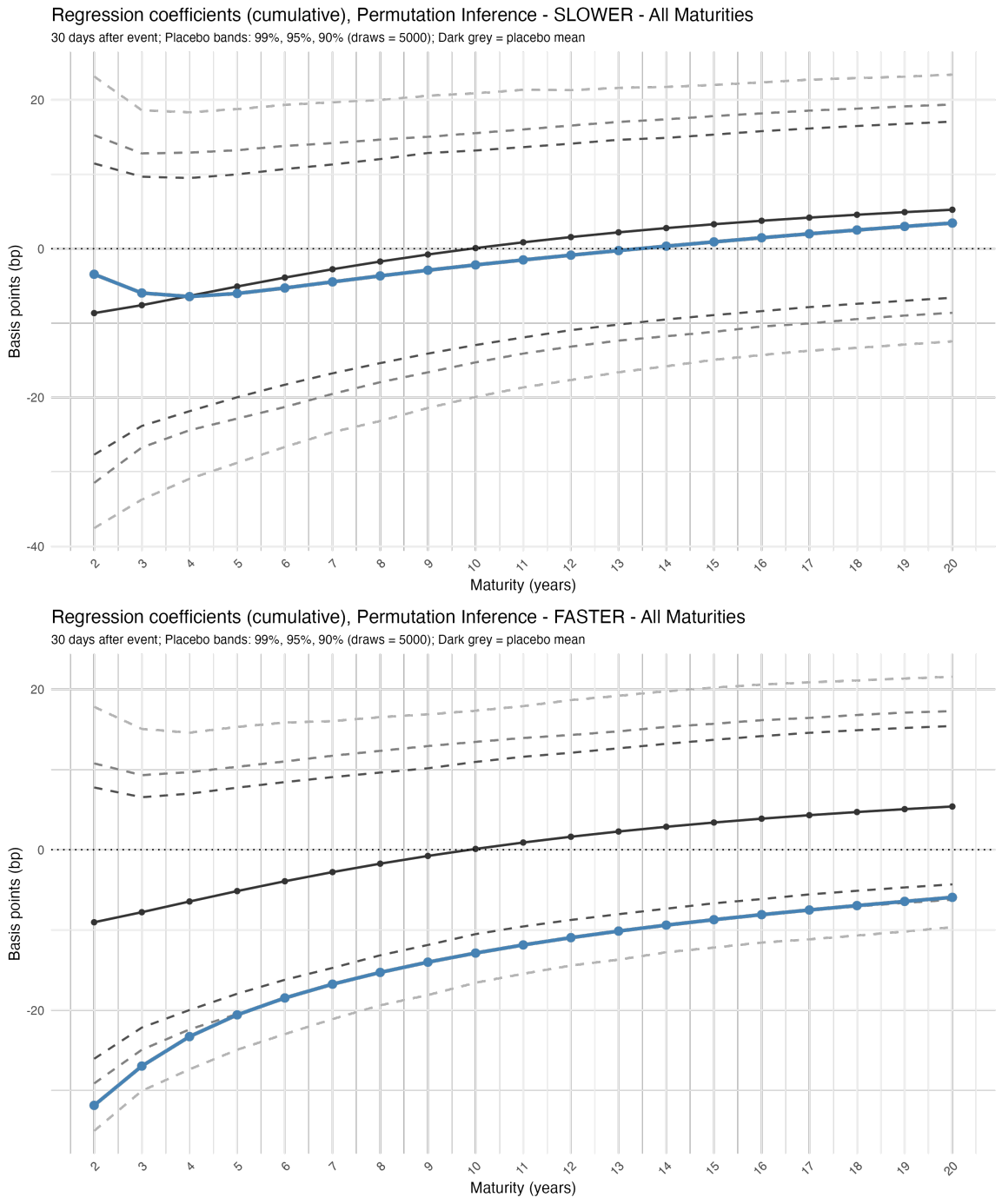

Figure 9: End-horizon ($l = 30$) change in zero-coupon TIPS yields by maturity, split by Metaculus prediction direction. Top panel: SLOWER coefficients (13 events). Bottom panel: FASTER coefficients (17 events). Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

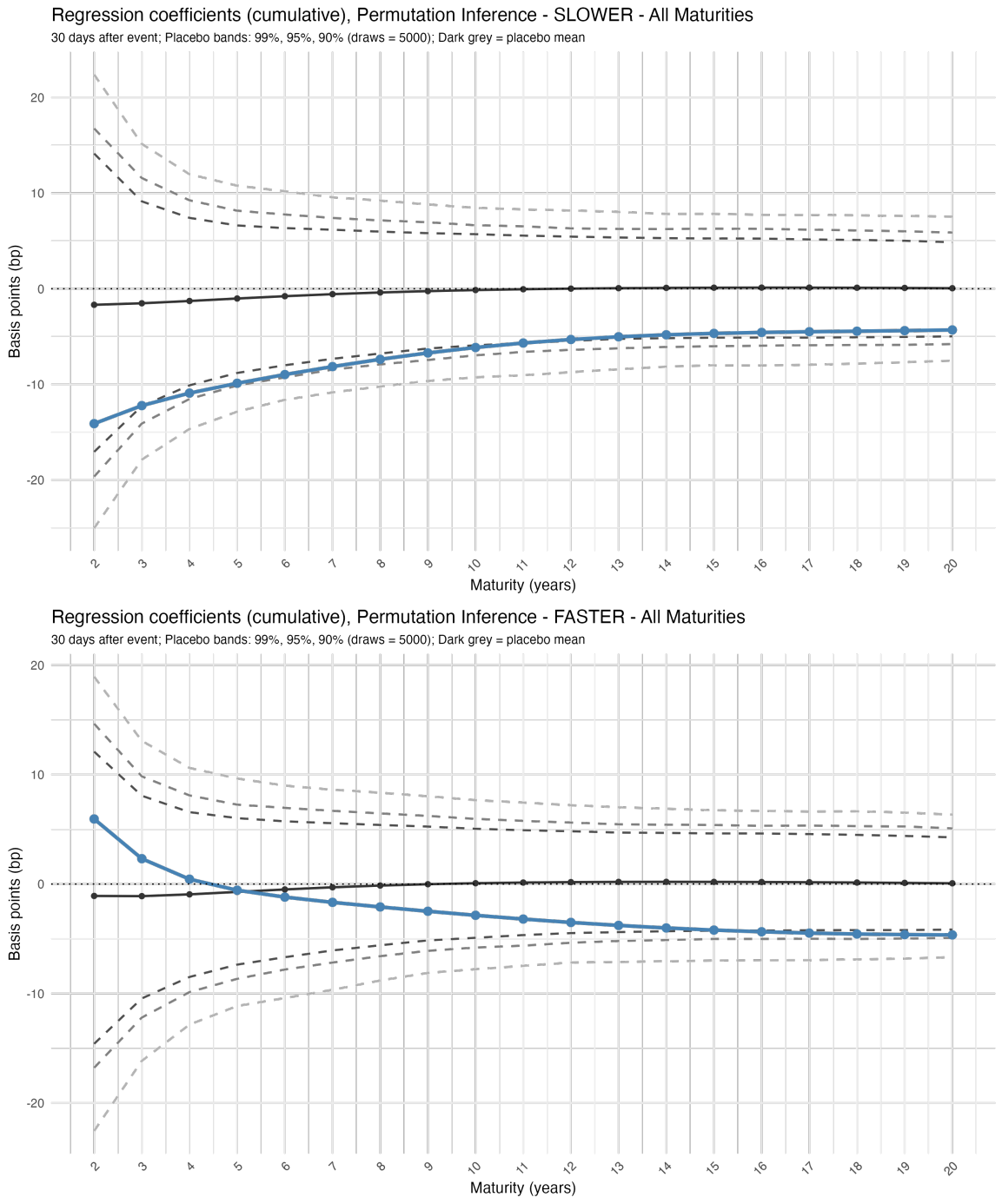

Figure 10: End-horizon ($l = 30$) change in zero-coupon inflation breakevens by maturity, split by Metaculus prediction direction. Top panel: SLOWER coefficients (13 events). Bottom panel: FASTER coefficients (17 events). Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

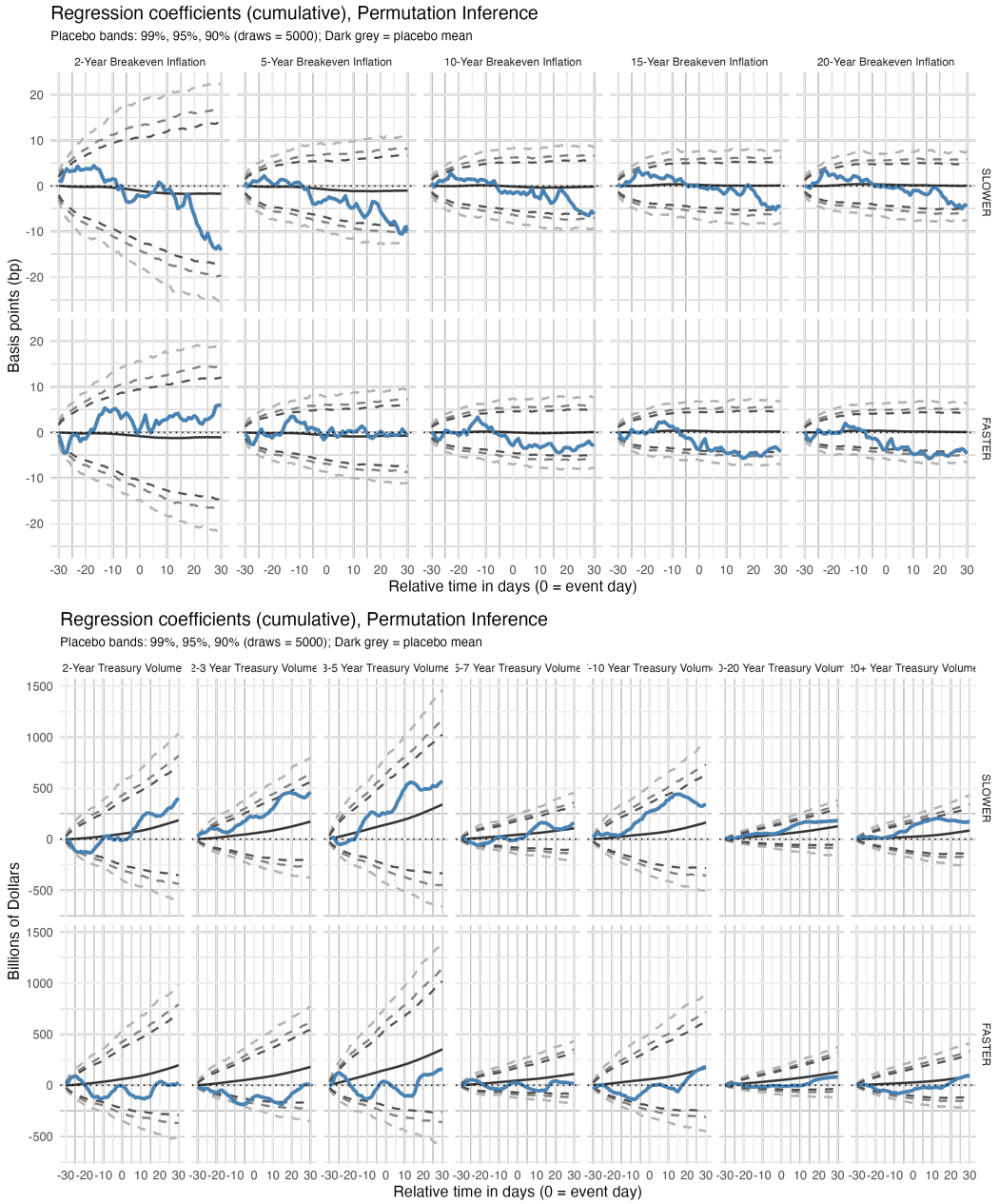

Figure 11: Regression-based event study for zero-coupon inflation breakevens (top panel) and demeaned cumulative Treasury trading volume (bottom panel), split by Metaculus prediction direction. Sample: November 2022 through December 2025.

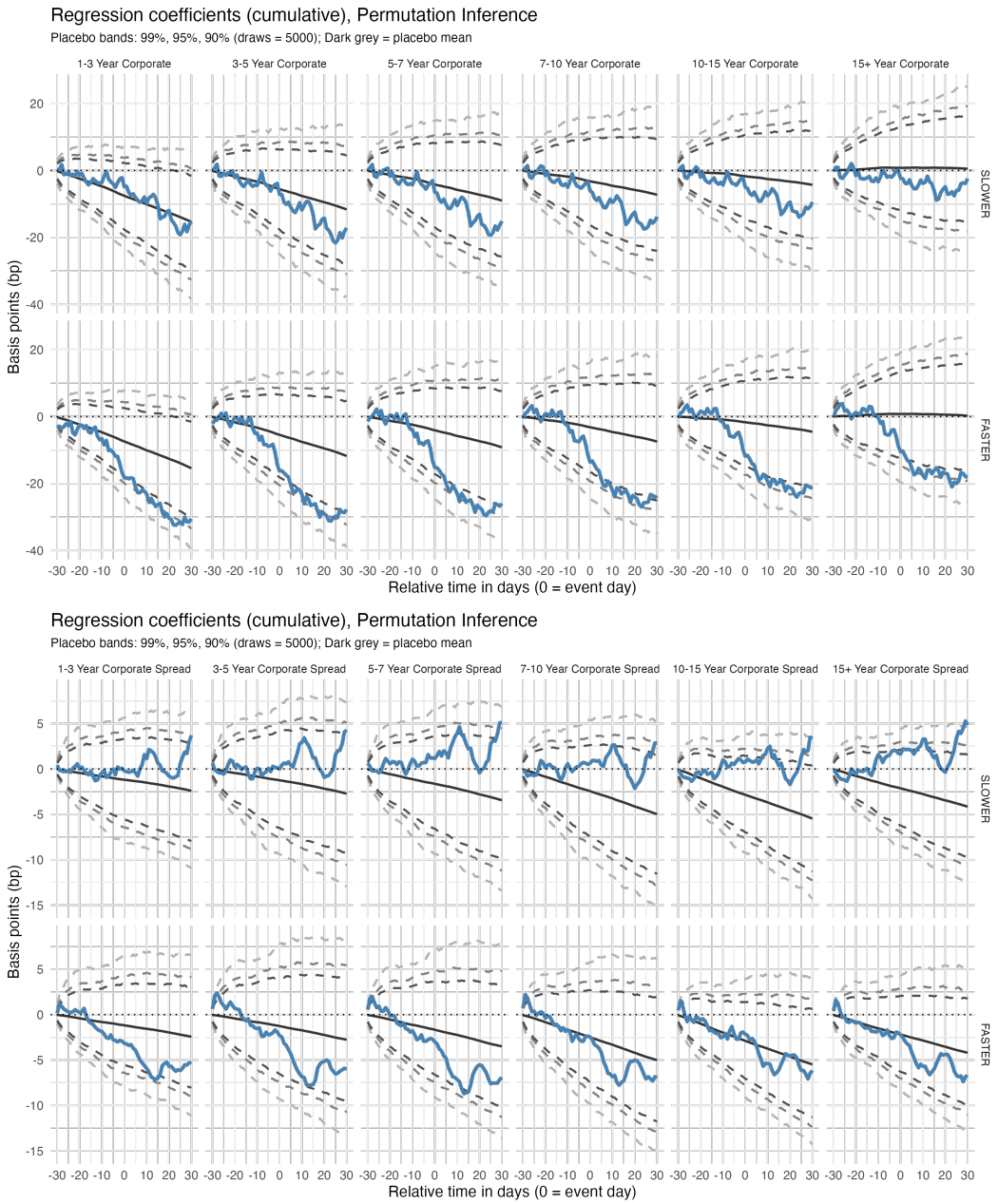

Figure 12: Regression-based event study for corporate bond yields (top panel) and option-adjusted corporate spreads (bottom panel) by maturity bucket, split by Metaculus prediction direction. Sample: November 2022 through December 2025.

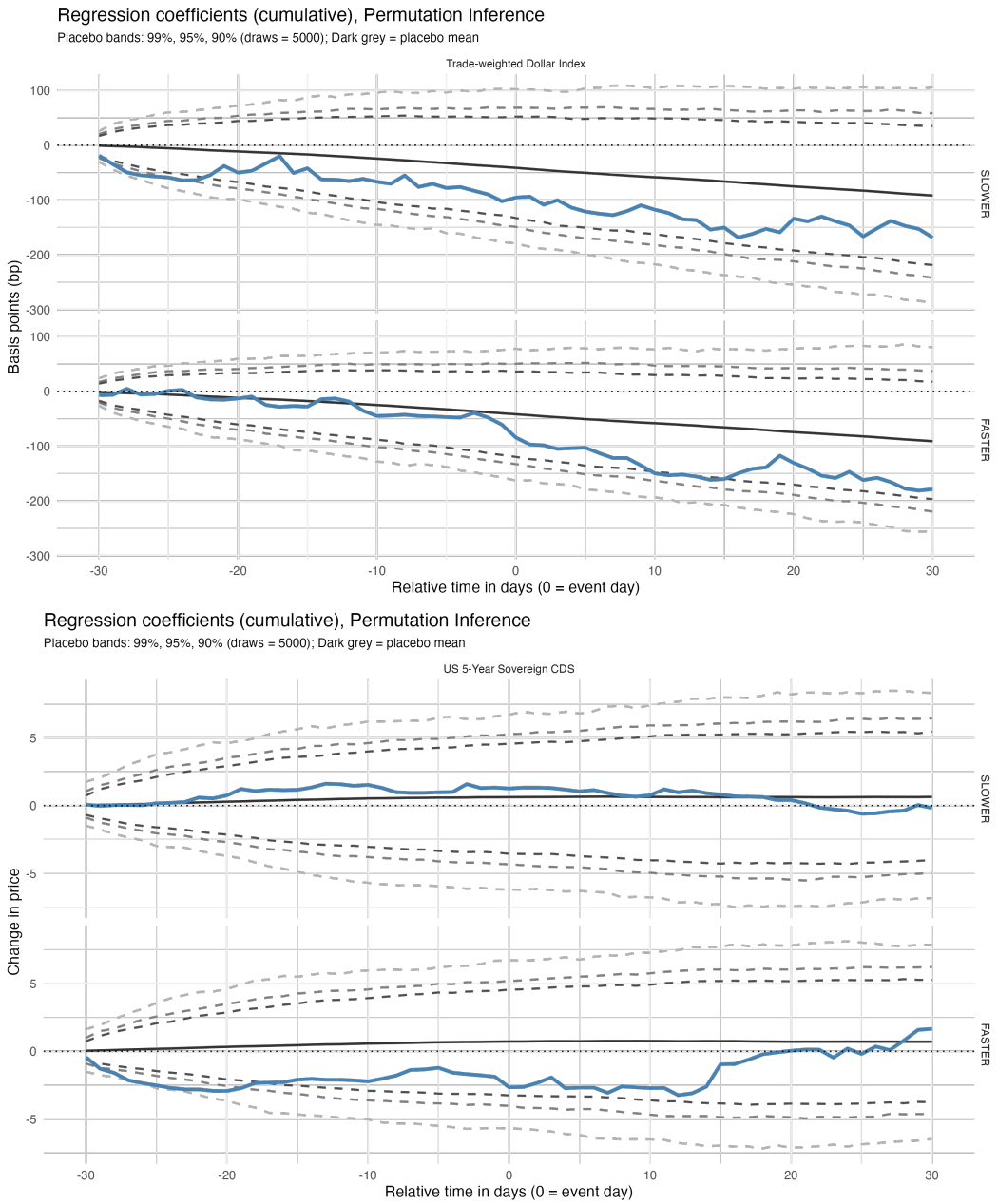

Figure 13: Regression-based event study for the broad trade-weighted US dollar index (top panel) and US sovereign CDS spreads (bottom panel), split by Metaculus prediction direction. Sample: November 2022 through December 2025.

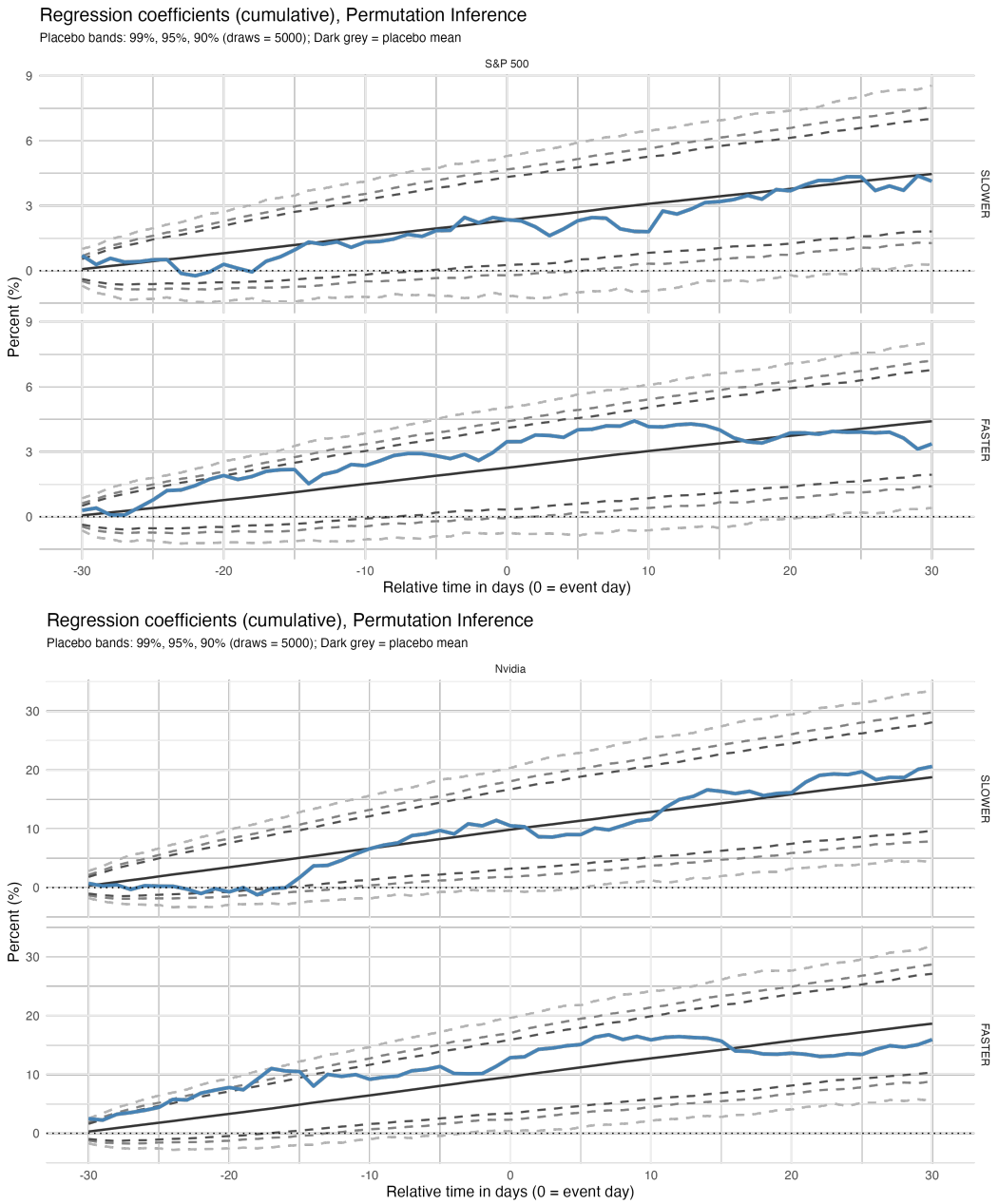

Figure 14: Regression-based event study for the S&P 500 (top panel) and NVIDIA share price (bottom panel), measured as cumulative percent return, split by Metaculus prediction direction. Sample: November 2022 through December 2025.

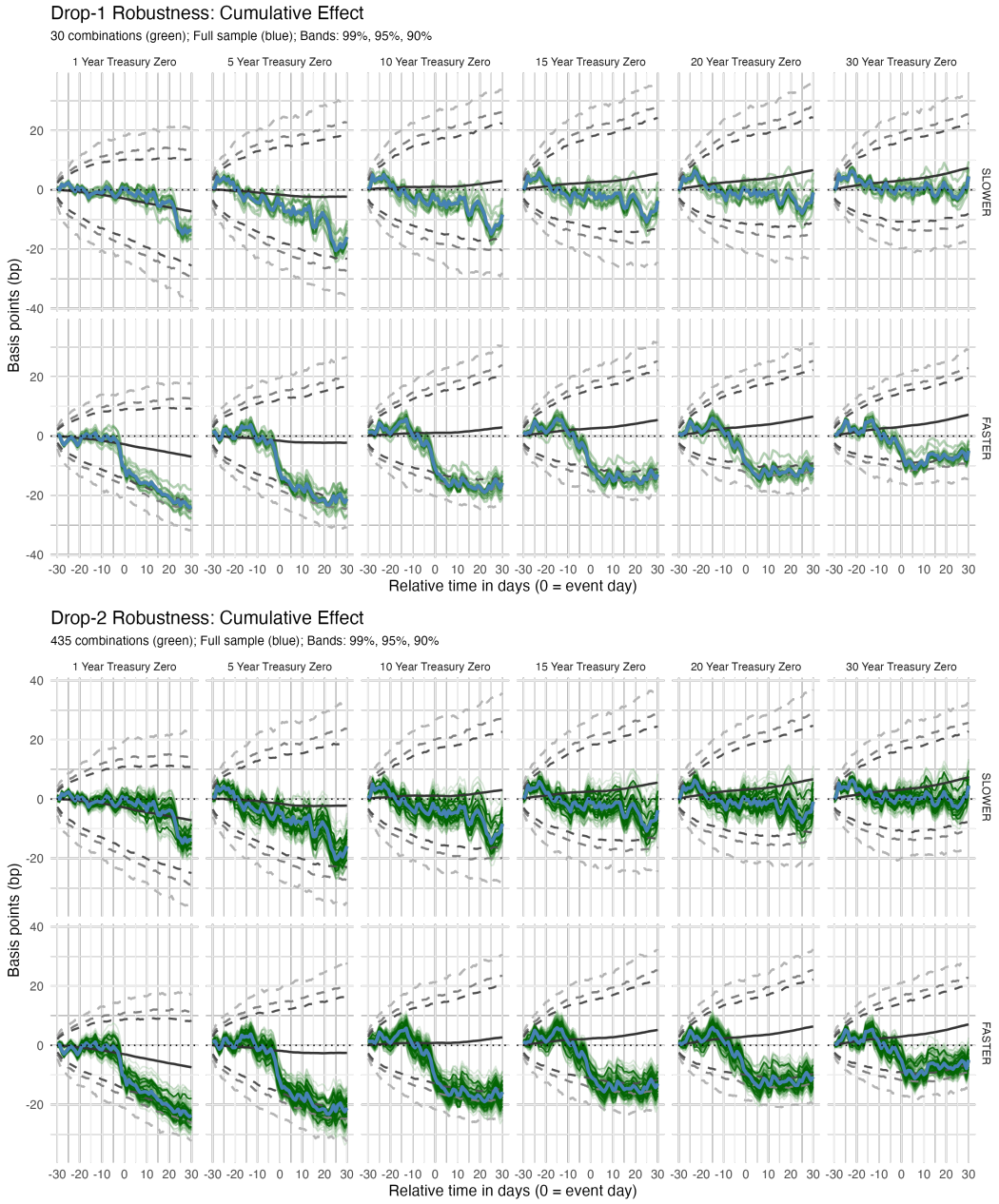

Figure 15: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, dropping one (top panel) or two (bottom panel) event dates at a time. Blue lines: full-sample estimate; green lines: each leave-j-out estimate. Sample: November 2022 through December 2025.

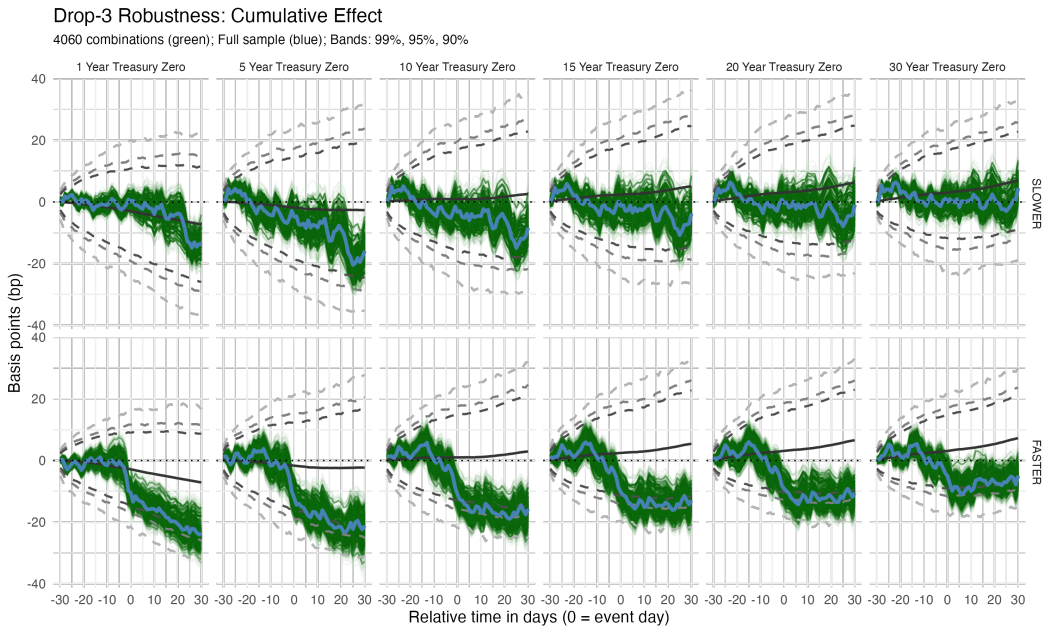

Figure 16: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, dropping three event dates at a time. Blue line: full-sample estimate; green lines: each leave-three-out estimate. Sample: November 2022 through December 2025.

D.2 Controlling for Economic Factors

To explore whether the patterns observed for the FASTER coefficients reflect other news that happened to occur near the AI release dates in our sample, we add the Cboe VIX volatility index (Cboe, 2025), the Federal Reserve Bank of San Francisco Daily News Sentiment Index (Shapiro et al., 2022; Federal Reserve Bank of San Francisco, 2023), and the Citigroup US Economic Surprise Index (Citigroup Global Markets, 2025) as additional regressors in (5), including their current values and fifteen daily lags.

D.3 Alternative Placebo Dates

The baseline permutation p-values reported throughout Section 3 draw placebo event dates uniformly at random within each calendar year, matching the per-year count to the actual data. In each check here we instead draw placebo dates from a specific alternative family of release dates, so the placebo distribution matches the timing pattern of that family rather than "any trading day in that year." The six placebo families are FOMC meeting dates, major tech-conference dates from the "Magnificent Seven" firms (Table 8), CPI release dates, BLS employment situation release dates, retail sales release dates, and US Treasury auction

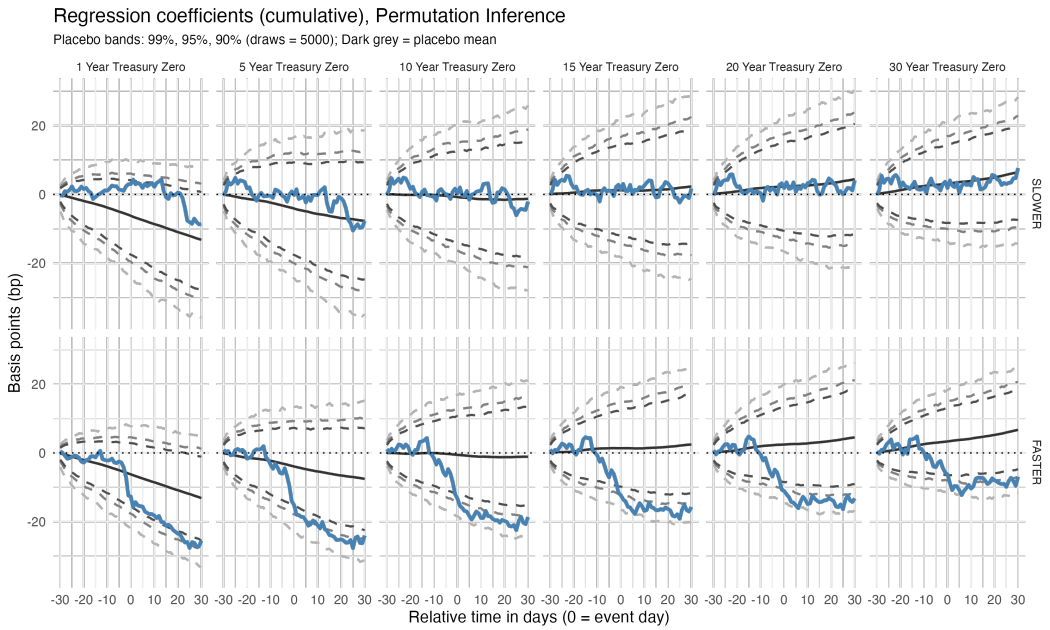

Figure 17: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, with VIX, Daily News Sentiment Index, and Economic Surprise Index (and 15 daily lags of each) added as additional regressors. Dashed lines show 10%, 5%, and 1% two-sided permutation critical values. Sample: November 2022 through December 2025.

dates for ten-, twenty-, and thirty-year bonds. Note that some of these series contain fewer dates than we have model releases, in which case the randomness of the placebo distribution arises solely from random assignment of the SLOWER and FASTER labels.

D.4 Alternative Sample Periods

Lastly, we split the analysis sample into three calendar-time subsamples and re-estimate the joint regression (5) within each subsample. Figure 21 reports results for the November 2022–December 2023 subsample (top panel; 5 FASTER and 2 SLOWER events) and the January–December 2024 subsample (bottom panel; 3 FASTER and 4 SLOWER events); Figure 22 reports the January–December 2025 subsample (9 FASTER and 7 SLOWER events). Each subsample contains fewer events than the full sample, so individual-subsample p-values are typically larger than the full-sample values.

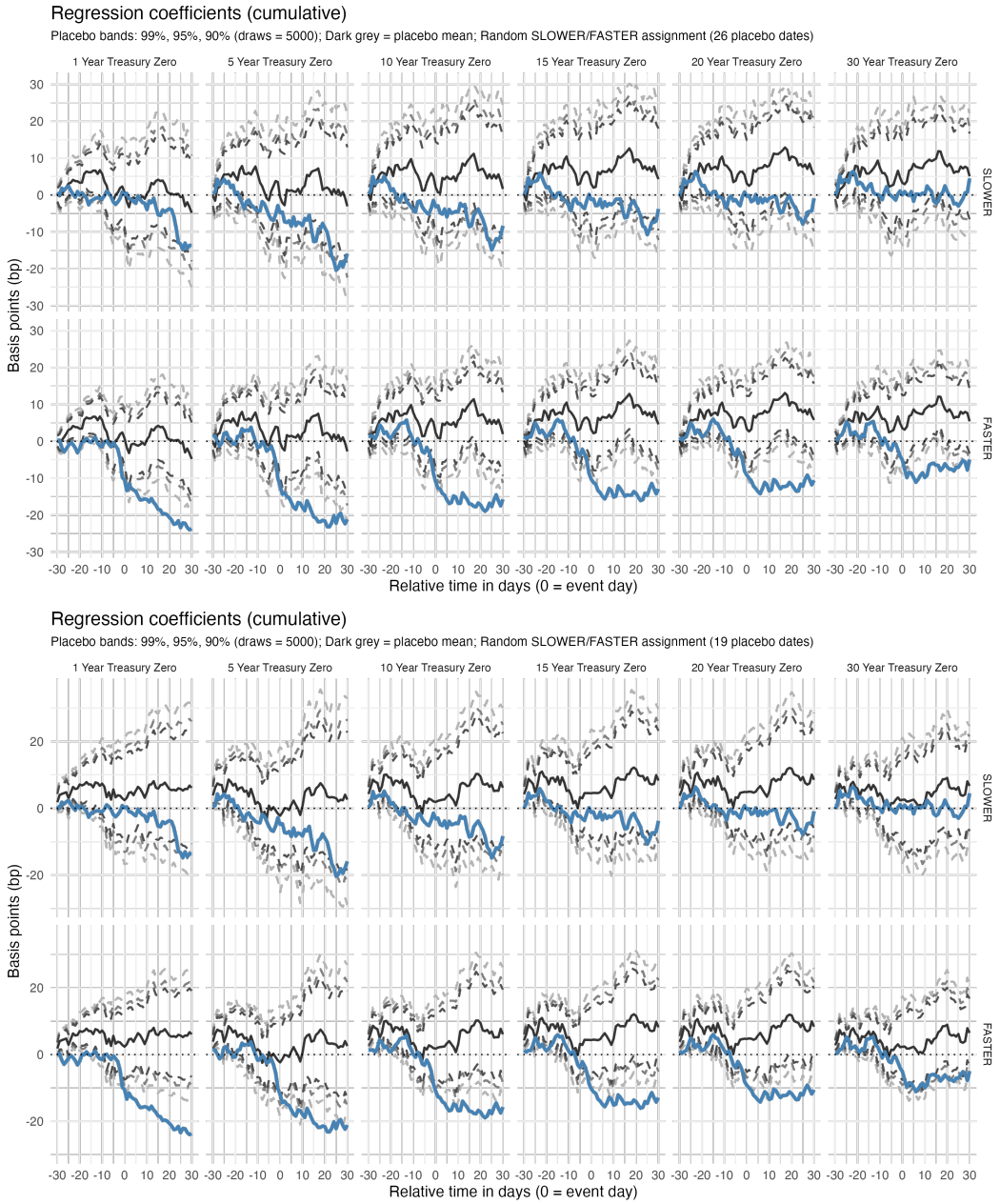

Figure 18: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, with placebo event dates drawn from FOMC meeting dates (top panel) and major tech-conference dates (bottom panel). Sample: November 2022 through December 2025.

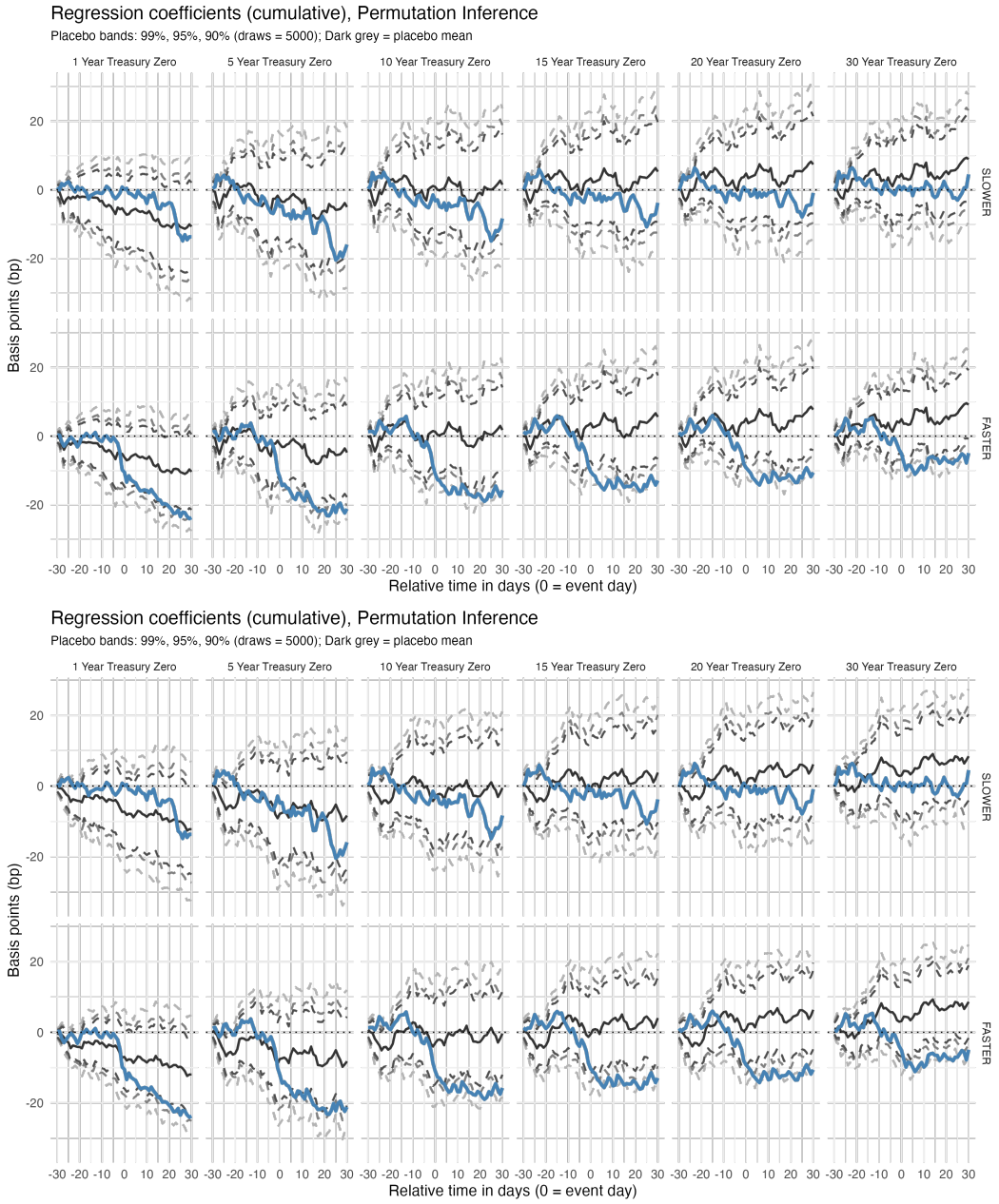

Figure 19: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, with placebo event dates drawn from CPI release dates (top panel) and BLS employment situation release dates (bottom panel). Sample: November 2022 through December 2025.

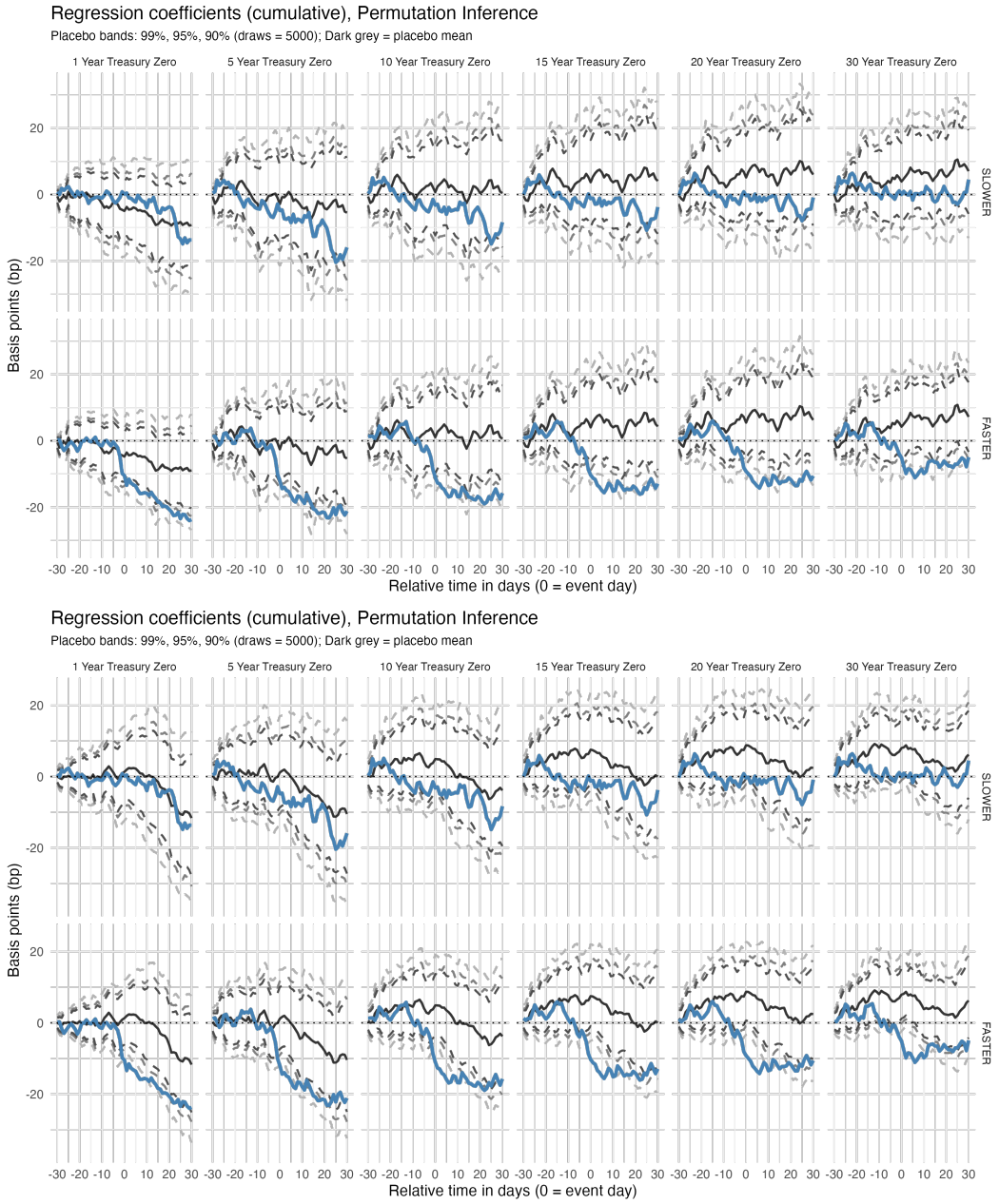

Figure 20: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, with placebo event dates drawn from retail sales release dates (top panel) and 10/20/30-year US Treasury auction dates (bottom panel). Sample: November 2022 through December 2025.

Table 8: Major Tech Conference Dates

Date	**Company**	**Conference**
2023 Conferences		
3/23/2023	Nvidia	Nvidia GTC
5/10/2023	Google	Google I/O
5/25/2023	Microsoft	Microsoft Build
6/9/2023	Apple	Worldwide Developers Conference
9/28/2023	Meta	Meta Connect
12/1/2023	Amazon	AWS re:Invent
2024 Conferences		
3/21/2024	Nvidia	Nvidia GTC
5/14/2024	Google	Google I/O
5/23/2024	Microsoft	Microsoft Build
6/14/2024	Apple	Worldwide Developers Conference
9/26/2024	Meta	Meta Connect
12/6/2024	Amazon	AWS re:Invent
2025 Conferences		
3/21/2025	Nvidia	Nvidia GTC
5/21/2025	Google	Google I/O
5/22/2025	Microsoft	Microsoft Build
6/13/2025	Apple	Worldwide Developers Conference
9/18/2025	Meta	Meta Connect
12/5/2025	Amazon	AWS re:Invent

Notes: Major annual tech conferences from the "Magnificent Seven" firms used as one of the alternative placebo families in Appendix D.3. We omit Tesla, which, unlike the other six firms, does not hold a comparable recurring annual conference. For multi-day conferences we use the final day.

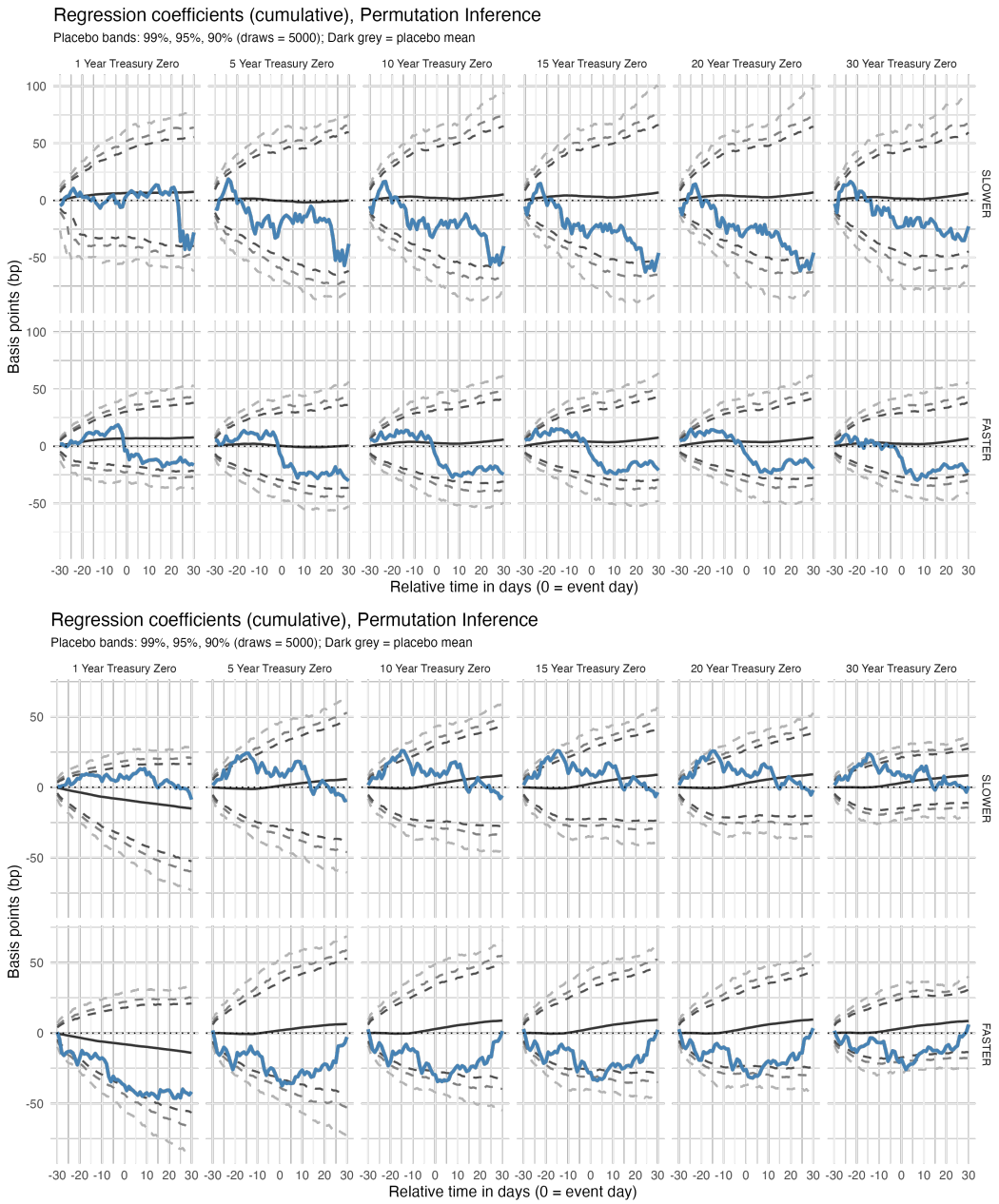

Figure 21: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, estimated on the November 2022–December 2023 subsample (top panel) and the January–December 2024 subsample (bottom panel).

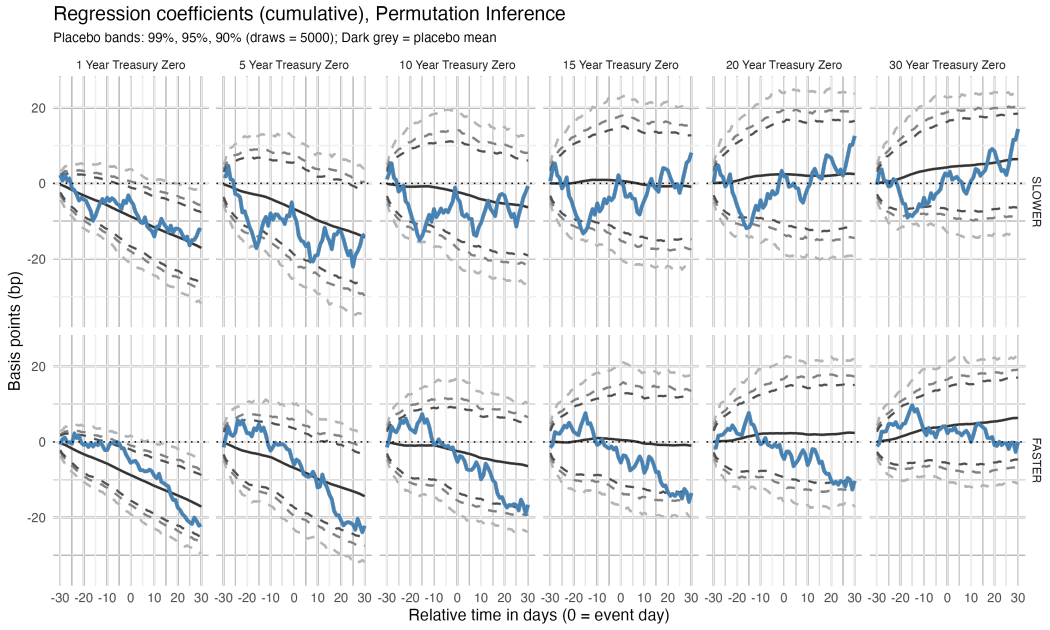

Figure 22: Regression-based event study for Treasury zero-coupon yields, split by Metaculus direction, estimated on the January–December 2025 subsample.