Looking Away from the Market

Adam Upenieks *

June 30th, 2026

Abstract

The Capital Asset Pricing Model (CAPM) is rejected empirically because market beta is weakly related to average returns. This paper shows that the performance of the CAPM depends on where institutional investors concentrate their attention. Using idiosyncratic news from the Dow Jones Newswire and institutional attention measures from Bloomberg search activity, I isolate states in which market-wide attention is focused on publicly available idiosyncratic news. The CAPM prices risk when institutional attention is concentrated on public idiosyncratic news, even though overall attention levels are lower. Market betas explain the cross section of stock returns, and the security market line is steep and positive. In contrast, when attention is elevated but not well explained by idiosyncratic news, the beta–return relation flattens or reverses. To rationalize these findings, I develop a noisy rational expectations model in which investors choose between concentrating attention on noisy idiosyncratic signals or dispersing it towards broader market level learning. Because attention to idiosyncratic news crowds out private learning about the common factor, greater attention to idiosyncratic information increases posterior factor uncertainty. Investors require compensation for bearing this uncertainty, generating a higher equilibrium factor risk premium and a steeper security market line. Empirically, unconditional betas price returns for approximately 51% of trading days.

*I thank Miguel Palacios, Alexander David, Marius Zoican, and many others for helpful comments. Seminar participants at the Haskayne school of Business also provided valuable feedback.

*University of Calgary, Haskayne School of Business, Adam.upenieks@Ucalgary.ca

1 Introduction

A persistent tension separates the theory of asset pricing from its empirical application. The capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) predicts that an asset's risk premium is increasing in its market beta. Despite the model's central role in academic research and practitioner discount-rate calculations (Berk and van Binsbergen, 2017), this prediction performs poorly empirically in cross-sectional tests. A recent literature offers conditional resolutions to this puzzle[1]. A common theme in these resolutions is that the beta–return relation depends on investors processing information, and several recent studies suggest that investor attention could play an important role in determining asset prices. This paper argues that the CAPM succeeds not only when investors pay more attention, but when attention is concentrated on a particular kind of information: public idiosyncratic news.

Models of attention allocation imply that investors reallocate a fixed budget between firm-specific and systematic information over time (Peng and Xiong, 2006; Kacperczyk et al., 2016). This paper posits that the composition of that attention, not merely its level, shapes the price of systematic risk, and with it the conditional success of the CAPM. I make two contributions to that claim. I show empirically that the CAPM prices the cross section when investor attention is concentrated on firm-specific news, even on days when aggregate attention is unremarkable. This pricing is evident for at least 51% of trading days between 2010 and 2025. The bottom quartile of the distribution alone accounts for 24.6% of trading days, yet captures 52% of the cumulative market return earned over the sample period. The second contribution is theoretical. I develop a noisy rational expectations model in which attention to purely firm-specific news and attention to direct macroeconomic signals are substitutes from the investor's perspective. Because firm-level news is informationally orthogonal to the aggregate factor by construction, each unit of attention allocated toward firm-level news crowds out private learning about the common market factor. Posterior factor uncertainty rises with the share of attention directed toward firm-specific news, the equilibrium factor risk premium rises with it, and the security market line steepens.

The theoretical model is a multi-firm variant of Grossman and Stiglitz (1980), in which firms release firm-specific public news and investors allocate a fixed attention budget between processing that news and conducting analysis of a common macroeconomic factor. I focus on firm-specific news because, once market-level content is filtered out, it is informationally

[1]See Tinic and West (1984), Savor and Wilson (2014), Savor and Wilson (2016), Hong and Sraer (2016), Jylhä (2018), Hendershott, Livdan, and Rösch (2020), Chan and Marsh (2022), Andrei, Friedman, and Ozel (2023), and Hasler and Martineau (2023, 2024), among others.

orthogonal to the aggregate factor: it resolves uncertainty about a firm's idiosyncratic payoff but conveys nothing about systematic risk or factor loadings as in (Patton and Verardo, 2012). Attention is divided across idiosyncratic and systematic learning channels, so that the concentration of attention, not merely its level, determines equilibrium prices (Peng and Xiong, 2006; Veldkamp, 2006b). The model predicts that when investors concentrate attention on firm-specific news, they necessarily learn less about the common factor, leaving them more exposed to factor uncertainty in equilibrium. To clear markets, expected returns must compensate for this residual uncertainty, so the cross-section of average returns aligns better with betas, and the security market line steepens. I verify this empirically.

The empirical strategy isolates this composition channel using two complementary measures of information processing. From the Dow Jones Institutional Newswire, I extract firm-specific news articles and filter out systematic and earnings-related content using keyword screens and a GPT-4o-mini classifier that retains only articles labeled as relevant to individual firms and not related to what happened in the market that day. From Bloomberg, I obtain firm-day institutional investor attention scores following Ben-Rephael, Da, and Israelsen (2017). I then construct two aggregate daily series for the Russell 3000 universe. The first is an idiosyncratic news attention index, IAI_t, defined as the cross-sectional average of Bloomberg attention scores among firms with a filtered idiosyncratic news article on day t. The second is an aggregate institutional attention index, AIA_t, defined as in Chan and Marsh (2022). The key state variable is the residual $\widehat{u}_t$ from a daily expanding-window regression of AIA_t on IAI_t. When $\widehat{u}_t$ is low, aggregate institutional attention is well explained by attention to filtered firm-specific news. In this state, attention is anchored in firm-level content. When $\widehat{u}_t$ is high, attention is elevated but not well explained by idiosyncratic news, consistent with attention being diverted toward broader market-wide or macroeconomic processing.

On the approximately 51% of trading days when the rolling median of $\widehat{u}_t$ lies below its trailing one-year 50th percentile, the beta–return relationship is positive and statistically significant, with magnitudes between 12 and 19 basis points per day (t-statistics above 3), and the intercept is statistically indistinguishable from zero. The security market line is steep, positive, and consistent with predictions of the CAPM. On the complementary set of days, the slope flattens or reverses. Crucially, this is not a high-attention result in the sense of Ben-Rephael, Carlin, Da, and Israelsen (2021). The bottom-quartile residual days are days on which AIA_t is below its sample mean, and IAI_t is above its sample mean; the share of total attention captured by firm-news is what matters, not the level. Replicating the expected-information-consumption framework of Ben-Rephael, Carlin, Da, and Israelsen (2021) confirms complementarity rather than redundancy: the conditional pricing result persists at full strength when the sample is restricted to firms that their measure does not flag as expected to consume information.

The closest prior work is Andrei, Friedman, and Ozel (2023) (hereafter AFO). AFO show that the CAPM performs better on earnings days, especially in high-VIX states when aggregate investor attention is elevated. Their mechanism, drawing on Epstein and Turnbull (1980), is an intertemporal resolution channel: attention pulls the resolution of factor uncertainty forward into the announcement window, so holding the market during this now-active resolution event is risky and commands a positive announcement-day premium. In their model, the earnings signal contains a factor component, so attention to earnings generates cross-sectional information spillovers about the aggregate factor and reduces posterior factor variance. The mechanism in this paper is the mirror image of AFO. The empirical state variable is not the level of attention, but where it is concentrated: the conditional success of the CAPM arises when institutional attention is concentrated on public firm-specific news, even though aggregate attention is below average. The theoretical primitive also operates in the opposite direction on factor uncertainty. Because firm-specific news is orthogonal to the factor in the model, attention to this news crowds out direct factor learning and thus raises posterior factor variance.

The resulting premium compensates investors for bearing greater unresolved systematic risk, rather than for bearing risk during accelerated resolution. This distinction also separates the two mechanisms empirically. AFO's channel is naturally strongest in high-uncertainty states, where attention is activated by heightened uncertainty and earnings signals carry factor-relevant information that reduces posterior factor variance. The attention concentration channel documented here operates through a different part of the information environment: it requires only that institutional attention be concentrated on firm-specific news that is orthogonal to the factor, and does not rely on elevated uncertainty to generate the premium. Consistent with this separation, the conditional beta slope on bottom-median residual days is 14.4 basis points ($t = 4.79$) within the low-VIX subsample, where AFO's resolution channel is dormant. The attention concentration channel is therefore most cleanly identified in precisely the regime where AFO's mechanism is inactive, suggesting the two channels capture distinct features of the information environment rather than competing explanations of the same variation.

This positioning has a deeper interpretive content. Berk and van Binsbergen (2016) document that mutual fund flows respond to CAPM alphas more strongly than to alphas from any multi-factor alternative, concluding that investors use the CAPM. Yet the unconditional CAPM is rejected in cross-sectional tests. One reconciliation is that investors form beliefs according to the CAPM but the cross-sectional pricing implication only emerges when their attention is actively engaged in processing the firm-specific information that translates beliefs into prices. The conditional result documented here is consistent with this view: the CAPM emerges in the cross section on the days when institutional attention is anchored to firm-news processing, and recedes when attention drifts elsewhere.

The findings also speak to a broader question about what investor attention does to asset prices. Peng and Xiong (2006) argue that limited-capacity investors exhibit category-learning behavior, choosing between processing market-wide and firm-specific information. Veldkamp (2006b) shows that costly information acquisition generates excess comovement when investors substitute toward common signals. Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) model attention reallocation across business-cycle states and show that attention to systematic versus idiosyncratic risks varies endogenously. The mechanism developed here builds on this tradition but with a distinct emphasis: under the orthogonality structure motivated by the LLM filter, attention composition has an unconditional effect on factor uncertainty that does not rely on equilibrium feedback through prices, learning from prices, or general-equilibrium attention-supply effects.

The conditional pricing result is robust to a battery of alternative specifications. Excluding firms within a five-day earnings announcement window barely affects the beta slope, ruling out the possibility that the finding is driven by the announcement premium documented in Savor and Wilson (2016) and Chan and Marsh (2022). Removing macroeconomic announcement days and leading earnings announcement days from the sample leaves the magnitudes and significance levels essentially unchanged. The finding is also not an artifact of sorting the sample into states. In a continuous specification (Panel C of Table 3), the daily Fama–MacBeth price of risk loads negatively on the standardized firm-specific attention residual $z(\hat{U}_t)$, with a Newey–West slope of roughly -12 basis points per standard deviation ($t = -3.9$), so the price of risk is highest and the security market line steepest on the days when attention is most concentrated on firm-specific news. This slope is essentially unchanged controlling for VIX, remains negative and significant once disagreement is added, and survives the contemporaneous market excess return, attenuating only from -14.1 to -9.5 basis points per standard deviation ($t = -3.96$). The market control is the demanding one, because the daily price of risk mechanically tracks the realized factor, so the persistence of the attention slope shows that the relationship reprices beta exposure rather than relabeling high- and low-market-return days. Splitting the bottom-median residual state by disagreement and VIX reveals that the result holds within both high and low disagreement environments, and is strongest when the VIX is low, consistent with the model's prediction that the composition channel is most transparent when background factor uncertainty is subdued and the resolution channel of Andrei, Friedman, and Ozel (2023) is dormant. Finally, reconstructing the attention-composition measure using firm-specific news from individual beta or size deciles shows that the conditional CAPM emerges broadly across constructions, with significant beta slopes in the majority of deciles for both sorts, indicating that no single segment of the cross section is responsible for the finding.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 describes the data and constructs the attention measures. Section 4 presents the empirical results. Section 5 conducts a battery of robustness tests. Section 6 develops the model and derives the central theoretical proposition. Section 7 explores a simple trading strategy which exploits these periods. Section 8 concludes.

2 Related Literature

The unconditional CAPM has been rejected in the cross section since Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), and Fama and French (1992), and a large literature attributes the flat security market line to leverage and margin frictions (Black, 1972; Frazzini and Pedersen, 2014; Jylhä, 2018), errors in estimated betas (Shanken, 1992), market-proxy misspecification (Roll, 1977; Kandel and Stambaugh, 1995), short-sale-constrained disagreement (Hong and Sraer, 2016), the correlation between beta and idiosyncratic volatility (Liu, Stambaugh, and Yuan, 2018), and dispersed information (Andrei, Cujean, and Wilson, 2023). A separate strand isolates states in which the line is steep and the model holds: in January (Tinic and West, 1984), on macroeconomic announcement days (Savor and Wilson, 2014), around the earnings of bellwether firms (Chan and Marsh, 2022), overnight (Hendershott, Livdan, and Rösch, 2020), when volatility is low (Hasler and Martineau, 2023), and when the expected market return is high (Hasler and Martineau, 2024). This paper adds a conditioning variable beyond these. The CAPM prices the cross section when institutional attention is concentrated on firm-specific news, so the relevant state is the composition of attention rather than a calendar event or a volatility regime.

The asset-pricing role of limited attention dates to Merton (1987), with distraction shown to dampen the price response to firm-specific news (Hirshleifer, Lim, and Teoh, 2009; DellaVigna and Pollet, 2009). Later work measures attention through retail search (Da, Engelberg, and Gao, 2011; Da, 2025), institutional Bloomberg activity (Ben-Rephael, Da, and Israelsen, 2017), and newspaper-based macroeconomic indices priced in announcement returns (Fisher, Martineau, and Sheng, 2022). The closest precedent is Ben-Rephael, Carlin, Da, and Israelsen (2021), who show that the level of institutional attention conditions the cross-sectional price of risk and that the announcement premium of Savor and Wilson (2014) concentrates on high-attention days. The measure here is the composition of attention rather than its level, and it tracks firm-specific rather than macroeconomic information: the share of institutional attention devoted to firm-specific news. That distinction drives the result. The conditional CAPM appears on days when total attention is below average but its firm-news share is elevated, precisely the days a level measure discards. The premise that investors divide a fixed budget between firm-specific and systematic information follows Peng and Xiong (2006) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), who show that capable investors tilt toward systematic information as recession risk rises; this paper asks what that allocation does to the price of systematic risk.

The model pairs the static, CARA-Gaussian, noisy-rational-expectations market structure (Grossman and Stiglitz, 1980; Admati, 1985) with an information-acquisition problem in the rational inattention tradition (Sims, 2003; Veldkamp, 2006b,c; Van Nieuwerburgh and Veldkamp, 2009, 2010), including dynamic-learning treatments that tie attention to volatility and risk premia (Andrei and Hasler, 2015). To this it adds one distinguishing assumption: filtered firm-news is orthogonal to the aggregate factor by construction, motivated by the LLM filter that strips market-level content from the news corpus. Under that orthogonality, attention to firm-news sharpens idiosyncratic learning only at the cost of factor learning, so the foregone factor learning raises posterior factor uncertainty V_f and with it the equilibrium factor premium.

This places the paper closest to Andrei, Friedman, and Ozel (2023), who also tie the conditional CAPM to attention and factor learning but move V_f in the opposite direction. They identify a level effect in the high-VIX regime, where heightened attention resolves macroeconomic uncertainty and lowers V_f; this paper identifies a composition effect that operates most cleanly in the low-VIX regime, where firm-specific attention crowds out factor learning and raises V_f. The two channels carry opposite signs on the same object and live in different regimes, so they are complementary rather than competing. The difference is built into the signal: their earnings news carries an explicit factor loading and so teaches investors about the factor, whereas filtered firm-news here is orthogonal to it by assumption, leaving private signals as the only conduit for factor information.

A related account is Andrei, Cujean, and Wilson (2023), in which dispersed information produces a strong CAPM at the agent level and a flat SML at the econometrician level, as public information crowds out private information and steepens the measured line. The present result is consistent with that view but mechanistically distinct. The firm-news attention measure is negatively correlated with disagreement, and the main result survives in states of both high and low aggregate disagreement, so disagreement is not the operative channel; the orthogonality of filtered firm-news and the opportunity cost of attending to it are. That orthogonality is a product of the empirical design, and the LLM classification step that produces it links the paper to the use of language models for news classification in finance (Lopez-Lira and Tang, 2023; Jha, Qian, and Tang, 2024).

3 Data and Variable Construction

3.0.1 News data

As opposed to Savor and Wilson (2014), Chan and Marsh (2022) and Andrei, Friedman, and Ozel (2023) who focus their analysis on days with scheduled market-relevant announcements, I focus my analysis on news that is firm-specific and not regularly scheduled. My news comes from the Dow Jones Institutional Newswire (DJIN) accessed through ProQuest TDM Studio. To search for news, I create a string containing the name of every company listed in the Russell 3000 database since 2009 when DJIN coverage starts. I then perform a search through the DJIN for articles whose title contains at least one unique CRSP company name. I keep only articles for which a company name is included in the title of the newswire article, as Aït-Sahalia, Li, and Li (2024) find these to be most relevant to the firm, and likely to cause immediate jumps in stock prices. To avoid capturing news about earnings, since they are known to also convey systematic information, I exclude articles for which the title contains the word stem 'earning', as well as other words commonly associated with articles written about earnings reports. I further exclude news whose headlines include the words "Buzz", "Wrap up", and any other commonly repeated keywords associated with summaries of previously posted actual news. A full list of these filters is provided in the Appendix. The news corpus consists of 902,873 firm-specific news articles beginning in 2010 and ends in 2025.

Newswire coverage is often noisy, and only weakly informative about firm fundamentals, making it difficult to locate relevant articles. A further challenge is to separate truly idiosyncratic news items from broad or market-wide news. Recent advances in large language models permit more refined filtering of such news, allowing me to isolate truly idiosyncratic news with greater precision. To refine the set of news articles used in the analysis, and focus on articles that are directly related to an individual firm, I use GPT-4.0 mini provided by TDM Studio to filter out news which is systematic or not relevant to individual firms. I use a prompt that instructs the model to use only information included in the text, and opt for a straightforward prompt to establish a foundation for the results. Nonetheless, employing more detailed prompts could allow for the extraction of more tailored information. The purpose of this step is to keep only articles written about individual firms and singular (as opposed to systematic) events. I ask GPT-4.0 mini to read each article and classify whether the article is related to one firm or many, and whether the article is written about one event or many (market-wide) events. The exact prompt is shown in the Appendix. Once Chat-GPT has read and labelled each of the initial 902,873 articles, I keep only those labelled as 'singular' and 'idiosyncratic'. The final GPT-labelled news corpus consists of 512,733 news articles.

3.0.2 Attention data

I follow Ben-Rephael, Carlin, Da, and Israelsen (2021) in constructing daily stock-level attention measures from Bloomberg terminals. Bloomberg records the number of times terminal users search for or read news articles on particular stocks. Bloomberg aggregates these numbers into hourly counts and creates an attention score by comparing the average hourly count during the previous eight hours to all hourly counts over the previous month for that same stock. They assign a score of 0, 1, 2, 3, or 4 if the rolling average is less than 80%, between 80% and 90%, between 90% and 94%, between 94% and 96%, and greater than 96% of the hourly counts over the previous 30 days. They then aggregate these scores up to a daily frequency by taking a maximum of all hourly scores through the day.

3.0.3 Stock return data

I retrieve excess market returns from Kenneth French's website. Following (Savor and Wilson, 2014) among others, I construct ten beta monthly-sorted portfolios using U.S. common stocks from CRSP which have a share code of 10 or 11, that match the sample of stocks with available attention data. The number of unique stocks is 2469. Daily stock betas are obtained from a rolling regression of the past 252 daily excess stock returns on the excess market returns. At the beginning of each month, stocks are sorted into one of ten beta-deciles, and the daily value-weighted and equal-weighted excess return of each decile portfolio is computed over the month.

For each portfolio, I estimate daily rolling betas by regressing the past 252 excess returns of each portfolio on the market return. These ten beta-sorted portfolios form my main test assets. In section 4, In robustness tests, I also consider individual stocks.

3.0.4 Defining high attention concentration periods

To identify periods where attention is concentrated on firm-specific news, and bring the model's attention-composition parameter ω to the data, I require a daily measure of how heavily institutional attention is tilted toward firm-specific news. I first develop a news attention index (IAI). Using Bloomberg query scores each day, I assign the attention value from 0-4 to individual stocks which had firm-specific news arrive through the Dow Jones Institutional Newswire, and take the daily average across only firms which have a news items on each day. As an additional step, I also recreate the aggregate institutional investor attention measure developed in Ben-Rephael, Da, and Israelsen (2017) using all Russell 3000 firms and create an average daily *AIA* following Chan and Marsh (2022), taking the mean level of attention paid to all firms for day t. 2469 unique stocks for which attention data and idiosyncratic news articles are available are used to create these attention measures. The sample runs from 2010 to 2025. I run the following daily expanding regression:

$$\mathrm{AIA}_t = \alpha_{t-1} + \beta_{t-1}\,\mathrm{IAI}_t + u_t, \tag{1}$$

I estimate $(\widehat{\alpha}_{t-1}, \widehat{\beta}_{t-1})$ using only information available up to day $t-1$ in an expanding window, and define $\widehat{u}_t$ as the resulting out-of-sample residual. $\widehat{u}_t$ is the component of total institutional attention that is not explained by contemporaneous idiosyncratic-attention intensity, or 'unanchored attention': marketwide attention not accounted for by firm-level news attention. Figure 1 plots the time series of the residuals from 2012 to 2025. It peaks in 2012-2013 during the Euro-area crisis, and spikes downward during the 2020 Covid recession and 2022 bear market.

The expanding-window construction serves two purposes. First, because each day's residual uses only information available through the prior day, the measure is free of look-ahead bias: any relationship documented between the residual and the conditional slope of the security market line cannot be an artifact of using future data to construct the regressor. Second, the out-of-sample residual measures the innovation in non-news attention: how much higher or lower aggregate attention is than the historical news–attention relationship would predict, rather than its absolute level. This is the appropriate object for the test, which concerns cross-day variation in the slope of the security market line rather than its level. The model predicts that the line is steeper on days when attention is composed toward firm-specific news (low residual) than on days when it is tilted toward market-wide learning (high residual). Provided the estimated news–attention coefficient is stable over the sample, the expanding residual coincides with a contemporaneous orthogonalization of attention, and its time variation reflects shifts in attention composition rather than drift in the estimated relationship.

Days when this residual is low are days when average institutional firm-level attention is well explained by firm-specific news. Building from the theory section, these should be days where investors demand a risk premium for market factor exposure. The empirical results test the CAPM on days when this residual is low. Figure 1 plots the time series of the residuals from 2012 to 2025.

4 Empirical CAPM tests

Figure 2 visually represents the main results by partitioning the sample into attention states defined by the residual $\widehat{u}_t$ from the expanding regression in equation (1), and plotting the estimated price of risk from daily Fama–MacBeth regressions of excess returns on the betas of value-weighted beta-sorted portfolios. Figure 2 shows that the price of risk falls monotonically across quintiles of the firm-specific attention residual. On the lowest-$\hat{U}_t$ days, when institutional attention is most concentrated on firm-specific news, beta is priced at roughly 19 and 18 basis points per day in the first two quintiles, both significantly positive at the 5% level. The estimate then declines through approximately 4, -5, and -13 basis points across the upper three quintiles, none of which is individually distinguishable from zero. The pooled price of risk is only about 4 basis points, so the unconditional security market line is nearly flat, and that flatness masks a steep, positive slope on firm-specific-attention days and an inverted one on the days when attention is dominated by other information. The concentration of statistical significance in the low-residual quintiles is exactly what the mechanism predicts, and the insignificance of the upper bins does not undercut it: each quintile holds only a fifth of the sample, so the tilt is identified by the monotone decline across quintiles and by the pooled low-residual regressions rather than by any single high-$\hat{U}_t$ bin.

Figures 3 and 4 plot the security market line for both value-weighted and equal-weighted portfolios separately for quantiles of $\widehat{u}_t$. For the figures, I use unconditional full-sample betas with excess returns averaged over the respective sample periods, following Savor and Wilson (2014) and Hendershott, Livdan, and Rösch (2020). When $\widehat{u}_t$ falls below its trailing one-year 50th percentile, aggregate institutional attention is well explained by idiosyncratic news attention, and the security market line slopes sharply upward. On the complementary set of days, the SML flattens or slopes downward.

I push the analysis further using both pooled panel regressions and Fama–MacBeth cross-sectional regressions. For the pooled specification, following Savor and Wilson (2014), Chan and Marsh (2022), and Hendershott, Livdan, and Rösch (2020), I estimate

$$r_{p,t+1} = \alpha + b_1 \widehat{\beta}_{p,t} + b_2 \mathbb{1}\{\text{State}_{t+1}\} + b_3 \widehat{\beta}_{p,t} \mathbb{1}\{\text{State}_{t+1}\} + \varepsilon_{p,t+1}, \tag{2}$$

where $r_{p,t+1}$ is the excess return of portfolio p on day $t+1$, $\widehat{\beta}_{p,t}$ is the rolling one-year portfolio beta estimated from the past 252 daily excess returns of each portfolio on the excess market return, and $\mathbb{1}\{\text{State}_{t+1}\}$ is a dummy variable equal to one when the five-day rolling median of $\widehat{u}$, evaluated through day t, falls below (or above) the indicated trailing one-year quantile cutoff of the daily residual, also evaluated through day t. The classification of day $t+1$ therefore uses only information available at the end of day t, mirroring the timing of the rolling beta, so that the state is known before the return is earned. The five-day median and the one-day lag each serve a distinct purpose. The daily residual measures a persistent attention regime with classification noise around it, and the median filters out marginal days admitted by that noise; the lag ensures that the conditional pricing results are predictive rather than a reflection of same-day co-movement between attention and returns. The Internet Appendix reports the full grid of smoothing windows and signal lags and shows that the results below are insensitive to these choices within the horizon over which the attention regime persists. Because the five-day median is less variable than the daily residual, it falls below the trailing 25th percentile of the daily distribution on fewer than a quarter of days, so the bottom-quartile state comprises 18.0% of the sample rather than 25%. The coefficient b_1 captures the unconditional beta–return relation on non-state days, b_2 captures the state-minus-other-days alpha, and b_3 measures the incremental change in the slope of the security market line on state days relative to all other days. Standard errors are clustered at the daily level.

For the Fama–MacBeth specification, I estimate daily cross-sectional regressions of portfolio excess returns on rolling portfolio betas separately within each attention state:

$$r^S_{p,t+1} = a^S + b^S \widehat{\beta}_{p,t} + \varepsilon^S_{p,t+1}, \tag{3}$$

where the superscript S indexes the attention state, and $\widehat{\beta}_{p,t}$ is the prior-day rolling portfolio beta. I report time-series averages of the daily slope coefficients b^S with Newey–West t-statistics computed with five lags on the time series of the estimated daily coefficients. I report results for four attention states defined by the rolling quantile cutoffs of $\widehat{u}$: the bottom quartile (0–25%), the bottom half (<50%), the top half (>50%), and the top quartile (>75%).

Table 2 presents the results for both value-weighted and equal-weighted beta-sorted portfolios across all four attention states.

Panel A: Bottom-quartile residual state (0–25%). This state identifies the 18.0% of trading days on which aggregate institutional attention is most heavily anchored to firm-specific news. In Fama–MacBeth regressions, a one-unit increase in beta is associated with an increase in next-day excess returns of 25.6 basis points for value-weighted portfolios ($t = 3.06$), and 24.8 basis points for equal-weighted portfolios ($t = 3.21$). The pricing of beta is therefore equally strong under both weighting schemes in this state. The value-weighted intercept is -8.61 basis points ($t = -1.30$), statistically indistinguishable from the zero intercept predicted by the CAPM. The average cross-sectional R^2 is 43.8%, indicating that variation in beta explains a large share of the cross-sectional return variation on these days.

The pooled regressions corroborate the Fama–MacBeth findings. On non-state days, the unconditional beta coefficient is 0.46 basis points ($t = 0.13$), economically zero, reaffirming the well-documented failure of the unconditional CAPM. The interaction term b_3, which measures the incremental slope of the SML on bottom-quartile residual days, is 20.1 basis points ($t = 2.67$), statistically significant at the 1% level. The state dummy b_2 is -6.16 basis points ($t = -0.97$), statistically indistinguishable from zero. For equal-weighted portfolios, the interaction coefficient is 19.7 basis points ($t = 2.76$), also significant at the 1% level.

Panel B: Below-median residual state (<50%). Expanding the state definition to include all days below the trailing median of $\widehat{u}$ captures 51.5% of trading days, slightly more than half of the sample. In value-weighted Fama–MacBeth regressions, the beta coefficient is 12.4 basis points ($t = 2.88$), highly significant. Importantly, the intercept is -2.64 basis points ($t = -0.79$), statistically indistinguishable from zero, as the CAPM predicts. The average R^2 is 45.1%. For equal-weighted portfolios, the beta coefficient is 9.0 basis points ($t = 2.26$), significant at the 5% level.

The pooled regressions deliver the same conclusion. The value-weighted interaction term is 19.7 basis points ($t = 3.24$), significant at the 1% level, while the unconditional beta coefficient b_1 is -5.51 basis points ($t = -1.23$), statistically indistinguishable from zero, confirming that beta does not explain excess returns on days outside this state. The implied within-state pooled slope, $b_1 + b_3 = 14.2$ basis points, is close to the Fama–MacBeth estimate of 12.4 basis points, providing reassurance that the result is not driven by differences in the econometric approach. The state dummy is -12.0 basis points ($t = -2.12$), significantly negative. For equal-weighted portfolios, the interaction is 14.9 basis points ($t = 2.61$), significant at the 1% level. The finding that unconditional betas price the cross section of returns on more than half of trading days constitutes a substantial expansion of the set of days on which the

CAPM has been shown to hold, extending well beyond the scheduled announcement days identified in Savor and Wilson (2014) and the leading-earnings-announcement days of Chan and Marsh (2022).

Panel C: Above-median residual state (>50%). On the complementary set of days on which aggregate institutional attention is elevated beyond what idiosyncratic news explains, the pricing of beta breaks down. In value-weighted Fama–MacBeth regressions, the beta coefficient is -3.0 basis points ($t = -0.62$), economically negligible and statistically insignificant, and the intercept is 5.92 basis points ($t = 1.53$), positive but imprecisely estimated. The SML is effectively flat: high-beta and low-beta portfolios earn similar average returns in this state, and the return variation is not explained by differences in systematic risk. For equal-weighted portfolios, the beta coefficient is -4.3 basis points ($t = -0.96$), and the intercept is 7.46 basis points ($t = 2.45$), significantly positive.

The pooled regressions for this state are estimated using a dummy equal to one on above-median residual days. The value-weighted interaction term is -15.1 basis points ($t = -2.46$), significant at the 5% level, while the unconditional beta coefficient, now capturing below-median days as the base, is 11.2 basis points ($t = 2.79$). The significantly negative interaction therefore reflects the steepness of the security market line on the below-median base days rather than a significantly inverted line within the state: the beta–return relation steepens when attention is concentrated on firm-specific news and flattens when it is not.

Panel D: Top-quartile residual state (>75%). The most extreme attention-diversion state, accounting for 17.6% of days on which unanchored attention is highest, produces no detectable pricing of beta. In value-weighted Fama–MacBeth regressions, the beta coefficient is -2.65 basis points ($t = -0.34$), indistinguishable from zero. The intercept is 9.22 basis points ($t = 1.59$), positive but imprecisely estimated. For equal-weighted portfolios, the slope is 1.4 basis points ($t = 0.19$), with an intercept of 9.13 basis points ($t = 1.87$). In pooled regressions, the value-weighted interaction is -7.5 basis points ($t = -0.92$), insignificant; the equal-weighted interaction is 0.37 basis points ($t = 0.05$), also insignificant. Neither weighting scheme produces evidence that beta is priced on these days. Notably, the security market line in this state is flat rather than inverted. Under contemporaneous classification with an unsmoothed daily state variable, the top-quartile slope is significantly negative, but the Internet Appendix shows that this inversion reflects same-day co-movement between attention and returns and disappears under any predetermined classification; the predictive statement supported by the data is that beta goes unpriced, not that it is priced with the wrong sign.

Taken together, the results in Table 2 reveal a monotonic pattern in the pricing of beta across attention states. Moving from the bottom quartile to the top quartile of the residual $\hat{u}$, the Fama–MacBeth beta slope declines from 25.6 to -2.65 basis points for value-weighted portfolios, the pooled interaction coefficients decline from 20.1 to -7.5 basis points, and the intercepts move from statistically zero toward positive. The CAPM prices risk when institutional attention is concentrated on public firm-specific news, even though the absolute level of attention on these days is below average. The results hold for both value-weighted and equal-weighted portfolios, across both Fama–MacBeth and pooled specifications, and are consistent with the model's prediction that attention composition, rather than attention level, determines the equilibrium price of systematic risk.

5 Additional tests

5.1 Continuous regressions, Disagreement and VIX states

Hong and Sraer (2016) argue that disagreement combined with short-sale constraints flattens the security market line, so that the CAPM should perform better when disagreement is low. Hasler and Martineau (2023) show that the CAPM performs better in periods of low volatility. If the bottom-median residual-attention state merely proxies for low disagreement or low volatility, the conditional pricing result would be subsumed by these existing channels. These channels make a sharp, testable prediction: beta should be priced specifically when disagreement and volatility are *low*, so a state variable that stood in for either would deliver a steep slope in the low bins and a flat slope in the high bins. To assess this, I split the bottom-median residual-attention days into high and low disagreement states, using daily aggregate disagreement data from Cookson and Niessner (2023), and separately into high and low VIX states; both the residual state and the disagreement or VIX classification are formed from information through day $t-1$. Table 3 reports Fama–MacBeth regressions within each subsample.

Panel A examines the disagreement splits. The value-weighted beta slope is 9.4 basis points ($t = 1.04$) on low-disagreement days and 10.8 basis points ($t = 2.01$) on high-disagreement days, so the slope is of comparable magnitude in both environments and is, if anything, larger and the only statistically significant of the two on the high-disagreement days. This is the opposite of the gradient the disagreement channel predicts: beta is priced precisely on the high-disagreement days where Hong and Sraer (2016) expect the security market line to be flattest. The high-disagreement subsample also carries the larger share of the sample, 773 days against 508, so the result is not an artifact of the low-disagreement cell. The equal-weighted estimates tell the same story, 9.7 basis points ($t = 1.21$) and 8.1 basis points ($t = 1.65$), positive and similar across the two regimes, and the intercepts are statistically indistinguishable from zero in both cells, as the CAPM predicts.

Panel B examines the VIX splits. The value-weighted slope is 11.2 basis points ($t = 2.44$) on low-VIX days and 14.3 basis points ($t = 1.68$) on high-VIX days, and the equal-weighted slope is 5.7 basis points ($t = 1.41$) and 14.3 basis points ($t = 1.76$), respectively. The slope does not fade as volatility rises; its point estimate is in fact larger on high-VIX days, even though Hasler and Martineau (2023) show that the unconditional CAPM performs better when volatility is low and would therefore predict the conditional slope to weaken in the high-VIX state. Beta remains priced when volatility is high, so the attention-composition state is not a relabeling of low-volatility days. The intercepts are again insignificant throughout.

Taken together, Panels A and B make the inference that a robustness test against these conditioning variables requires. Because Hong and Sraer (2016) and Hasler and Martineau (2023) both locate CAPM pricing in the low-disagreement and low-volatility states, a residual-attention measure that simply proxied for either would price beta only in those bins and not in their complements. Instead, the slope is positive across all four subsamples, of comparable magnitude, significant at the 5% level in the high-disagreement and low-VIX cells and marginally significant in the high-VIX cell, and it shows no gradient toward the low-disagreement or low-volatility cells. The attention-composition channel therefore prices beta independently of disagreement and volatility rather than inheriting their variation.

Panel C of Table 3 regresses the daily market premium estimated over all days against the continuous measure of attention concentration over the entire sample period, with controls for other state variables from the literature. Hasler and Martineau (2024) show that the CAPM performs better when the expected market return is high, a reminder that the realized and the expected market return are distinct objects and that the CAPM restricts only the latter: it requires the expected slope of the security market line to equal the expected market premium, $\mathbb{E}[\lambda] = \mathbb{E}[R_m - R_f]$, while leaving deviations of the realized slope from the contemporaneous factor return as unforecastable noise. This distinction matters for reading the continuous specification, because the daily price of risk λ_t co-moves mechanically with the realized market excess return. In a single-factor world a portfolio with a unit spread in beta earns approximately the realized factor return, so λ_t tracks $R_{m,t} - R_{f,t}$ up to estimation error in the betas and idiosyncratic cross-sectional noise. Panel C makes this link plain: holding the sample fixed, adding $R_m - R_f$ to the regression raises the R^2 from 0.010 to 0.362 while leaving the coefficient on $z(\hat{U}_t)$ strongly significant ($t = -3.92$). Controlling for the contemporaneous market return therefore strips out the component of the realized slope that simply follows the day's factor realization and isolates the variation orthogonal to it. The last two rows of Panel C show that the attention coefficient survives this control, attenuating only from -15.39 to -10.00 basis points per standard deviation, which establishes that the relationship between attention composition and the price of risk is not a relabeling of high- and low-market-return days but a repricing of beta exposure itself. The predictability of these orthogonal deviations by $z(\hat{U}_t)$ is precisely the state-dependent tilt of the security market line that the model attributes to movements in posterior factor uncertainty.

5.2 Expected information consumption and individual firms

A natural concern with the regime variable is that bottom-25% residual days may simply identify days when institutional investors are predictably consuming information at specific firms, in which case the conditional CAPM finding would be subsumed by Ben-Rephael, Carlin, Da, and Israelsen's (2021) Expected Information Consumption (EIC) mechanism. To address this, I construct EIC at the firm-day level following their methodology: a firm receives EIC $= 1$ on day t if it has historically exhibited institutional attention spikes around peer-firm scheduled events, FOMC announcements, or macroeconomic announcements occurring on that day. The construction follows BCDI's specifications closely: I use Fama–French 48-industry classifications to identify peer firms, require a minimum historical spike rate of 30% on past peer-earnings days for EIC_PEER, and apply BCDI's 50% threshold for FOMC and macroeconomic announcements based on the previous four FOMC days and previous year of macro announcements. The construction produces frequencies and diagnostics in line with BCDI's reported values. The combined EIC $= 1$ frequency is 5.6% of firm-days, comparable to BCDI's reported 6.7% for EIC_ALL despite the more limited Bloomberg-coverage sample. The AIA spike rate among EIC $= 1$ firm-days is 30.5%, roughly five times the 5.9% spike rate among EIC $= 0$ firm-days, closely tracking BCDI's reported differential of 24.2% versus 7.1% in their Table III. Table 4 reports the conditional CAPM result on bottom-25% residual days, partitioned by EIC. The headline slope of 20.4 bps ($t = 3.15$) on the full cross-section drops only marginally to 19.1 bps ($t = 3.04$) when I restrict to EIC $= 0$ firms, which are those that BCDI's framework specifically does not flag as expected to consume information. The conditional CAPM result is therefore not driven by firms that BCDI's mechanism predicts. EIC $= 1$ firms exhibit particularly strong conditional pricing (44.2 bps, $t = 3.10$), within BCDI's reported range of 8 to 44 bps for EIC $= 1$ subsamples and consistent with their findings, indicating that the two mechanisms operate complementarily: BCDI's framework identifies firm-day variation in attention-driven pricing, while the regime variable identifies day-level variation in the broader attention environment that affects the entire cross-section.

5.3 Earnings announcements and the conditional CAPM

A natural concern is that the conditional pricing result on bottom-quartile residual-attention days is driven by firms that are simultaneously releasing earnings. Earnings announcements convey systematic information and have been shown to steepen the security market line in their own right (Savor and Wilson, 2016; Chan and Marsh, 2022). If bottom-quartile residual days happen to coincide with earnings seasons, the apparent success of the CAPM on these days could simply reflect the announcement premium documented in prior work rather than a distinct attention-composition channel. To address this concern, I define earnings-window firm-days as firm-days for which the firm has an earnings announcement within five business days before or after the current date, and re-estimate the Fama–MacBeth cross-sectional regressions on bottom-quartile residual-attention days in three subsamples: all firms, firms excluding those in an earnings window, and earnings-window firms only.

Table 5 reports the results. On the full cross section, the beta slope is 20.4 basis points ($t = 3.15$) with an insignificant intercept of -2.7 basis points ($t = -0.74$), consistent with the baseline findings. When earnings-window firms are excluded, the beta slope declines only marginally to 18.8 basis points ($t = 2.98$), remaining highly significant. The intercept is -1.8 basis points ($t = -0.49$), statistically indistinguishable from zero. This subsample retains an average of 1,390 firms per cross section, so the result is not an artifact of a thin sample. Restricting the cross section to earnings-window firms only produces a beta slope of 24.4 basis points ($t = 2.88$), larger in magnitude and consistent with the view that earnings announcements contribute to the pricing of beta through a complementary channel. The intercept for this subsample is 1.4 basis points ($t = 0.22$), again insignificant.

The key finding is that removing all firms near an earnings announcement barely affects the conditional beta slope. The decline from 20.4 to 18.8 basis points is economically small relative to the magnitude of the effect, and the t-statistic remains above conventional thresholds. The attention-composition channel therefore operates across the broad cross section of firms, not only among those releasing earnings. The stronger result within earnings-window firms is consistent with, but not required by, the main finding, and suggests that the two mechanisms operate complementarily: earnings announcements provide a firm-level information event that reinforces the cross-sectional pricing of beta, while the attention-composition state variable identifies a market-wide regime in which that pricing extends to all firms.

5.4 Excluding macroeconomic announcement days and LEADs

A further concern is that the bottom-median residual-attention days may overlap with macroeconomic announcement days (Savor and Wilson, 2014) or with leading earnings announcement days (LEADs) identified by Chan and Marsh (2022). If so, the conditional pricing result could be confounded by the well-documented announcement premium on these days. To address this, I re-estimate the Fama–MacBeth regressions after removing these potentially confounding days from the sample.

Table 6 reports the results. The first row excludes LEAD days, defined as days when influential S&P 500 firms announce earnings, following Chan and Marsh (2022). After removing these days, the value-weighted beta slope is 12.3 basis points ($t = 2.72$) and the equal-weighted slope is 8.5 basis points ($t = 2.13$), both statistically significant. The sample retains 1,343 days. The second row excludes macroeconomic announcement days, defined as days with scheduled FOMC, employment, or PPI releases following Savor and Wilson (2014). The value-weighted slope is 12.5 basis points ($t = 3.13$) and the equal-weighted slope is 8.6 basis points ($t = 2.38$), again both significant, with 1,554 days remaining. In both specifications, the intercepts are economically small and statistically insignificant, consistent with the predictions of the CAPM.

The magnitudes of the beta slopes are close to unchanged relative to the baseline results in Table 2, and the t-statistics remain well above conventional thresholds. The conditional pricing of beta on bottom-median residual-attention days is therefore not driven by overlap with known macroeconomic or bellwether-earnings announcement days. This finding, combined with the earnings-window results in Table 5, suggests that the attention-composition channel identifies a broad regime in the information environment that is distinct from the specific event-day channels documented in prior literature.

5.5 Relation to Betting-against-Beta

The betting-against-beta factor is the traded counterpart of a flat security market line. It is long low-beta and short high-beta stocks with the market exposure hedged, so it earns its premium on exactly the days the line is flat or inverted and loses on the days it steepens. In our sample its daily return has a correlation of -0.52 with the cross-sectional price of risk λ_t, confirming that it moves as the mirror image of the slope, which makes it the natural object on which to test the model out of sample. Table 7 conditions the premium on the firm-specific attention regime. On the days when firm-specific attention dominates, the bottom quartile of the predetermined attention residual, the factor earns -0.28 basis points, indistinguishable from zero, while on the remaining days it earns 4.18. The difference of -4.47 basis points carries a t-statistic of -1.63 in raw returns and strengthens to -2.08 and -2.07 once the factor is winsorized at the 1st/99th percentiles, a standard treatment for a series this crash-prone. The anomaly is thus economically and statistically absent on precisely the days the model predicts a steep, CAPM-consistent line, and is earned entirely on the rest.

The pattern is a discrete step rather than a smooth gradient. Across quintiles of the attention residual the most firm-specific quintile is the single low cell and the other four form a flat plateau, which is why a linear specification understates the effect: there is no monotone trend to fit, only a step at the firm-specific extreme. The apparent dip in the most macro-dominated quintile in raw returns, 0.81 basis points, is an artifact of a few fat-tailed days. It disappears under winsorization, rising to 3.26 and 3.77, and reverses to the highest cell, 5.09, when 2020 is excluded. We read this as conditioning and locating the anomaly rather than displacing the funding-constraint account of it. The crashes that give the factor its fat left tail, the days on which high-beta assets rally and the line steepens violently, fall on the non-firm-specific days and are orthogonal to the attention measure, which is why winsorizing leaves the low-U cell almost unchanged while lifting the others. The firm-specific effect is instead a steady difference that sharpens when those tails are removed. In this sense the flat unconditional line is a composition average: quiet where firm-specific attention dominates, and carrying both the premium of the anomaly and its crash risk on the days where it does not.

Panel C views the same factor from the betting-on-beta side, which is where the economic content is sharpest. I build a beta-neutral low-minus-high factor from the beta-sorted decile portfolios, equal-weighted, hedge it to zero market beta, and report its mirror, the return to betting on beta. Because the factor carries no market exposure, its mean return is its CAPM alpha, so a zero entry is the literal statement that the CAPM holds. On firm-specific days betting on beta earns -0.19 basis points, indistinguishable from zero at $t = -0.06$: once beta is priced there is nothing left to capture. On the remaining days it earns -2.98 basis points at $t = -1.66$, the losing side of the same anomaly that betting against beta harvests in Panel A. The underlying factor is sound, with a market beta near zero and a 0.69 correlation with the traded AQR factor.

The absence of alpha is the result, not its presence, and it is what separates a restored CAPM from a reversed anomaly. If high beta were underpriced on firm-specific days, betting on beta would earn positive alpha there; instead the alpha is zero, so on those days beta is priced correctly rather than mispriced in the other direction. This is the cleanest form of the paper's claim: not that a new strategy pays when firm-specific attention dominates, but that the deviation from the CAPM disappears. Consistent with that, the regime-conditioned factor, betting against beta on macro days and with beta on firm-specific days, earns an alpha of 0.39 basis points against the Fama-French five factors plus the AQR factor, at $t = 0.30$, adding nothing to the existing set. The contribution is a location rather than a new factor: the flat-line anomaly that betting against beta monetizes lives on macro-attention days and vanishes precisely where firm-specific attention concentrates, which is where the CAPM is restored.

5.6 Which firms' news drives the result?

The baseline attention-composition measure constructs the idiosyncratic attention index IAI_t by averaging Bloomberg query scores across all firms with a filtered news article on day t. A natural question is whether the conditional pricing result is driven by news arriving at a specific segment of the cross section. If, for example, only news about high-beta firms or large firms generates the attention state that steepens the SML, the mechanism would be narrower than the model implies. To investigate this, I reconstruct the expanding regression in equation (1) ten times, each time computing IAI_t using only the subset of firms belonging to a single beta decile (or, separately, a single size decile). For each decile-specific residual, I identify bottom-median days and re-estimate the Fama–MacBeth cross-sectional regressions on those days. If the conditional CAPM result appears broadly across decile-specific constructions, then no single part of the firm distribution is responsible for the finding.

Table 8 reports the results when the attention state is defined using news from each beta decile separately. Panel B presents value-weighted portfolio results. Across the ten beta-decile-specific constructions, the beta slope λ_β is positive in every column, ranging from 5.3 basis points (decile 9) to 12.5 basis points (decile 1). The slope is statistically significant at the 5% level or better in six of the ten deciles (deciles 1, 3, 4, 5, 7, and 10), and significant at the 10% level in a seventh (decile 2, $t = 1.85$). The intercepts are uniformly small and insignificant, ranging from -4.4 to 1.4 basis points, none statistically different from zero. The average cross-sectional R^2 is approximately 46% throughout, indicating stable explanatory power regardless of which decile's news is used to construct the state variable. For equal-weighted portfolios in Panel A, the pattern is qualitatively similar though weaker in magnitude, with the beta slope significant in decile 1 (8.6 basis points, $t = 2.38$) and marginally significant in deciles 3, 4, and 5.

The strongest value-weighted results arise when the attention state is constructed from news about the lowest-beta firms (decile 1: 12.5 basis points, $t = 3.13$) and mid-range beta firms (deciles 3 through 5: 11.0 to 11.5 basis points, t-statistics between 2.7 and 2.9). The weakest results come from deciles 6 and 9. The broad pattern suggests that attention to firm-specific news steepens the SML regardless of where in the beta distribution that news originates, though the effect is somewhat attenuated when the state is defined using news from the highest-beta or highest-attention deciles alone.

Table 9 repeats the exercise using size deciles. Panel B again presents value-weighted results. The beta slope is positive in all ten columns, ranging from 5.7 basis points (decile 8) to 13.0 basis points (decile 7). Significance at the 5% level or better obtains in seven of the ten size deciles (deciles 1, 2, 4, 5, 6, 7, and 10), with an additional decile (decile 9: 11.2 basis points, $t = 2.55$) also significant. The intercepts are again uniformly small and insignificant, consistent with the CAPM. The equal-weighted results in Panel A are weaker in magnitude, with significance at the 5% level in deciles 6 and 7 (7.4 and 7.3 basis points) and marginal significance in decile 5 (6.8 basis points, $t = 1.94$).

The size-decile results indicate that the attention-composition channel is not confined to news about firms of a particular size. News from both small firms (decile 1: 8.2 basis points, $t = 1.98$) and large firms (decile 10: 9.5 basis points, $t = 2.33$) generates attention states in which beta prices the cross section. The strongest results emerge when the state is constructed from mid-to-large capitalization firms (deciles 4 through 7: 10.9 to 13.0 basis points), which may reflect the fact that these firms account for a larger share of aggregate institutional attention and therefore produce more informative variation in the residual.

Taken together, the results in Tables 8 and 9 show that the conditional CAPM finding is not an artifact of news concentrated in any single segment of the cross section. The attention-composition state variable steepens the security market line when constructed from news about low-beta or high-beta firms, small or large firms, and most points in between. This breadth is consistent with the model's prediction that the relevant object is the aggregate share of attention devoted to firm-specific news, not the characteristics of the firms producing that news.

6 Model

Can the concentration of attention towards firm-specific news explain the rise in risk premia associated with beta? Because asset prices and attention allocation might not have a straightforward link, this task requires a model.

This section presents a noisy rational expectations model of asset prices with costly attention.

Investors are modeled in the tradition of Grossman and Stiglitz (1980) and Admati (1985): they are risk averse, receive Gaussian signals, and learn only imperfectly about the assets they trade. As in Veldkamp (2006b), costly information acquisition leads investors to rely more heavily on common signals, and as in Peng and Xiong (2006), investors with limited attention choose between processing market-wide and firm-specific information. The model places this choice inside a factor structure: asset payoffs load on a common factor, and each investor divides a fixed attention budget between firm-specific public news, which is orthogonal to the factor, and signals about the factor itself. When attention concentrates on firm-specific news, the common factor is learned less well and posterior factor uncertainty remains high. Because beta is compensation for bearing factor risk, investors then demand a larger premium per unit of beta, and the security market line steepens. As attention shifts toward the factor, that uncertainty resolves, the premium contracts, and the security market line flattens.

The economy lasts a single period and contains N risky assets and a risk-free asset normalized to zero return. The payoff to risky asset $i \in \{1, \ldots, N\}$ is

$$d_i = \bar{d} + \beta_i f + \varepsilon_i, \tag{4}$$

where $\bar{d}$ is a known constant, $f \sim \mathcal{N}(0, \sigma_f^2)$ is a common factor, and $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ are mutually independent firm-specific shocks. The factor loading is

$$\beta_i = b_i + \nu_i, \qquad \nu_i \sim \mathcal{N}(0, \sigma_\nu^2), \tag{5}$$

with b_i a known prior mean and ν_i mutually independent across i and orthogonal to (f, ε_i). I write $\tau_f \equiv \sigma_f^{-2}$, $\tau_\nu \equiv \sigma_\nu^{-2}$, $\tau_\varepsilon \equiv \sigma_\varepsilon^{-2}$.

A continuum of investors of unit mass, indexed by $j \in [0, 1]$, has CARA preferences over terminal wealth with common risk aversion $\rho > 0$. Per-capita asset supply is $\bar{x} + u_i$, where $u_i \sim \mathcal{N}(0, \sigma_u^2)$ is a noisy supply shock that prevents prices from being fully revealing.

6.1 Information

Investors receive two distinct signals: firm-specific public news and private macro information.

Assumption 1 (Firm-specific public signal)**.** *For each asset i, all investors observe*

$$y_i = \varepsilon_i + \eta_i, \qquad \eta_i \sim \mathcal{N}(0, 1/\tau_y), \tag{6}$$

where η_i is mutually independent across i and orthogonal to $(f, \beta_i, \varepsilon_i)$.

Assumption 1 treats firm-level news as informationally orthogonal to the aggregate factor. Empirically, this corresponds to news classified as firm-specific: earnings releases, management changes, product announcements, after market-level content has been filtered out. Public signals inform investors about ε_i, but carry no information about f or about β_i. This restriction distinguishes the setup from standard noisy-rational-expectations models, in which a single public signal jointly conveys both firm-specific and factor information.

Assumption 2 (Private signal)**.** *Each investor j privately observes*

$$z_j = f + \xi_j, \qquad \xi_j \sim \mathcal{N}(0, 1/\tau_\xi), \tag{7}$$

with ξ_j independent across investors and orthogonal to $(f, \varepsilon_i, \nu_i, \eta_i)$.

6.2 Attention allocation

Each investor allocates a unit attention budget between the two channels. Let $\omega_j \in [0, 1]$ denote the share allocated to firm-specific public news. Public-signal precision scales with day-t news supply S:

$$\tau_y = \omega_j\, S, \tag{8}$$

and private precision scales with investor-specific capacity K:

$$\tau_\xi = (1 - \omega_j)\, K. \tag{9}$$

Higher ω_j reflects greater concentration of attention on firm-specific news; lower ω_j reflects diversion to private macro analysis.

7 Equilibrium

7.1 Posterior beliefs

Lemma 1 (Posteriors)**.** *Under Assumptions 1–2, the posterior distributions are mutually independent and given by*

$$f \mid z_j \sim \mathcal{N}\Big(\mu_f,\, V_f\Big), \qquad V_f = \frac{1}{\tau_f + (1-\omega_j)K}, \tag{10}$$

$$\beta_i \mid y_i \sim \mathcal{N}\Big(b_i,\, \sigma_\nu^2\Big), \tag{11}$$

$$\varepsilon_i \mid y_i \sim \mathcal{N}\Big(\mu_{\varepsilon,i},\, V_\varepsilon\Big), \qquad V_\varepsilon = \frac{1}{\tau_\varepsilon + \omega_j S}. \tag{12}$$

Lemma 1 reflects three features. First, the posterior of f is determined entirely by the private signal: under Assumption 1, public news is uninformative about the factor. Second, the posterior of β_i equals its prior, because y_i carries no information about loadings. Third, public news reduces the posterior variance of ε_i in proportion to $\omega_j S$.

Define the residual non-factor variance entering the equilibrium pricing problem:

$$V_\varepsilon^\star \;=\; V_\varepsilon + \sigma_\nu^2\Big(V_f + \sigma_f^2\Big). \tag{13}$$

This combines posterior idiosyncratic variance with the contribution of unresolved loading uncertainty interacting with the factor.

7.2 Equilibrium prices

Each investor solves a standard CARA-Gaussian portfolio problem, generating linear demand. Market clearing pins down equilibrium prices.

Lemma 2 (Prices)**.** *The equilibrium price of asset i is*

$$p_i \;=\; \bar{d} + b_i\bar{\mu}_f \;-\; \rho V_f\, b_i\,(N\bar{b}\,\bar{x}) \;-\; \rho V_\varepsilon^\star\,\bar{x} \;+\; (\textit{linear noise}), \tag{14}$$

where $\bar{\mu}_f$ is the average posterior mean of f across investors and $\bar{b} \equiv N^{-1}\sum_i b_i$.

The risk premium $\mathbb{E}[d_i - p_i]$ decomposes into a factor component, $\rho V_f b_i(N\bar{b}\bar{x})$, proportional to b_i, and a non-factor component, $\rho V_\varepsilon^\star \bar{x}$, common to all assets. The factor component is what a cross-sectional regression of returns on betas identifies.

7.3 The Fama–MacBeth slope

In practice, the econometrician estimates loadings from a T-period rolling-window OLS regression of realized returns on the factor and then runs a Fama–MacBeth (FM) cross-sectional regression. The resulting slope is

$$\lambda^{FM}(\omega) \;=\; A(\omega) \cdot \rho V_f(\omega)\, N\bar{b}\,\bar{x}, \tag{15}$$

where $A(\omega) \in (0,1)$ is the errors-in-variables attenuation:

$$A(\omega) \;=\; \frac{s_\beta^2 - \sigma_\nu^2}{s_\beta^2 + V_e(\omega)}, \qquad V_e(\omega) \;=\; \frac{\sigma_\varepsilon^2 + 1/(\omega S)}{T\sigma_f^2}, \tag{16}$$

and $s_\beta^2 \equiv \sigma_b^2 + \sigma_\nu^2$ is the cross-sectional variance of true betas.

8 Optimal attention and the risk-premium channel

8.1 Investor's information problem

Each investor chooses ω_j to maximize the expected utility gain from information. Under CARA-Gaussian preferences, the information-allocation objective reduces to (see Appendix E):

$$\mathcal{V}(\omega) \;=\; -\frac{N-1}{2}\log V_\varepsilon^\star(\omega) \;-\; \frac{1}{2}\log\Big(V_\varepsilon^\star(\omega) + V_f(\omega)\textstyle\sum_i b_i^2\Big). \tag{17}$$

The first term captures the cumulative gain from reducing residual non-factor uncertainty across N assets. The second term captures the cost of bearing factor risk: as V_f rises, the factor eigenvalue of the posterior covariance matrix grows, increasing the variance of the diversified factor exposure.

Proposition 1 (Interior optimum)**.** *For all admissible parameter values* $(N, S, K, \rho, \sigma_f, \sigma_\nu, \sigma_\varepsilon, \tau_f)$, *the optimal attention allocation* $\omega^\star \in (0,1)$ *is interior and satisfies the first-order condition* $\mathcal{V}'(\omega^\star) = 0$.

The interior optimum reflects a substantive trade-off. Moving attention toward firm-specific news improves idiosyncratic learning (reducing V_ε) but, because firm news is orthogonal to the factor by Assumption 1, every unit of attention reallocated to public news is one unit lost from private factor learning. The investor balances these two channels.

8.2 The risk-premium channel

The model's central economic content concerns the comparative static of V_f with respect to ω.

Proposition 2 (Risk-premium channel). *Posterior factor uncertainty is strictly increasing in attention to firm-specific news:*

$$\frac{dV_f}{d\omega} \;=\; K \cdot V_f^2 \;>\; 0. \tag{18}$$

Proof. From $V_f = (\tau_f + (1-\omega)K)^{-1}$,

$$\frac{dV_f}{d\omega} \;=\; -V_f^2 \cdot \frac{d}{d\omega}\Big[\tau_f + (1-\omega)K\Big] \;=\; KV_f^2 \;>\; 0. \qquad \square$$

Proposition 2 holds unconditionally as it does not depend on interior optimality, on parameter calibration, or on any equilibrium FOC. The mechanism is direct. Under Assumption 1, firm-specific news carries no factor information; the only channel through which investors learn about f is the private signal z_j. Therefore each additional unit of attention allocated to firm-specific news is one unit of attention lost from factor learning, and the posterior variance of f rises in ω at the deterministic rate KV_f^2.

8.3 Conditional steepness of the SML

Combining (15)–(16) with Proposition 2 yields the model's main empirical prediction.

Proposition 3 (Conditional CAPM steepness). *The Fama–MacBeth slope is strictly increasing in the share of attention allocated to firm-specific news:*

$$\frac{d\lambda^{FM}}{d\omega} \;=\; \rho N \bar{b}\, \bar{x} \left[A(\omega) \cdot KV_f^2 \;+\; V_f \cdot \frac{dA}{d\omega} \right] \;>\; 0. \tag{19}$$

The first term inside the bracket is the risk-premium channel (Proposition 2); the second is an errors-in-variables channel through $dA/d\omega \geq 0$.

When investors concentrate attention on firm-specific news, they bear more factor uncertainty in equilibrium. To clear markets, expected returns must compensate for this uncertainty. Cross-sectional pricing aligns more sharply with betas: the security market line steepens, and unconditional betas explain the cross-section of average returns. Conversely, when attention is diverted toward private macro analysis or broad-market information, posterior factor uncertainty falls, the factor risk premium shrinks, and the SML flattens.

Across days, variation in $\omega^\star$ can be driven by parameters not explicitly modeled here: time-variation in private-research capacity K, or in investor composition. The model does not pin down the cross-day determinants of $\omega_t^\star$, but it is clear about the implication once $\omega_t^\star$ is observed: high-ω days produce steep security market lines; low-ω days produce flat ones.

9 Trading strategy

The empirical results documented above establish that the security market line steepens when institutional attention is concentrated on firm-specific news. I now explore the economic significance of this finding through a set of trading strategies that exploit the attention-composition state variable. The strategies are designed to be easily implementable and to assess whether the conditional pricing of beta translates into economically meaningful abnormal returns after standard risk adjustment.

The baseline strategy takes a long position in the highest-beta decile portfolio and a short position in the lowest-beta decile portfolio on days identified as bottom-quartile residual-attention days, and holds cash on all other days. The regime signal is days when the 5-day rolling median of μ_t is below its rolling 1-year 25th percentile, lagged by one trading day so that positions are entered or exited the day after the signal switches, ensuring that the strategy uses only information available at the time of trading. Because the residual $\widehat{u}_t$ is constructed from a persistent, slowly moving rolling median, the signal switches states infrequently, keeping turnover low. In practice, an investor could approximate the long and short legs using liquid high-beta and low-beta ETFs, which would keep implementation costs manageable. The portfolio returns I report are gross of transaction costs, financing frictions, and shorting fees. The sample runs from 2012 to 2025 and contains 3,222 trading days, of which 448 are classified as regime days by the lagged signal.

Table 10 reports the results. Panel A provides two benchmarks. The first is an always-on long-short strategy that holds the high-minus-low beta spread portfolio every day regardless of the attention state. This strategy earns an average daily return of 4.3 basis points with an annualized Sharpe ratio of 0.34, and generates no significant alpha under either the CAPM (-4.35% annualized, $t = -0.63$) or the Fama-French five-factor model (2.58%, $t = 0.42$). The failure of the unconditional long-short strategy to generate alpha is consistent with the well-documented flatness of the unconditional security market line. The second benchmark is the buy-and-hold market portfolio, which earns 5.4 basis points per day with an annualized Sharpe ratio of 0.79.

Panel B reports the headline conditional strategy. By restricting the long-short position to bottom-quartile residual-attention days and holding cash otherwise, the strategy earns 3.4 basis points per day with an annualized Sharpe ratio of 0.67. The strategy generates a CAPM alpha of 6.5% per year ($t = 1.86$) and a five-factor alpha of 7.4% per year ($t = 2.11$), both statistically significant. The market beta of the strategy is 0.13, reflecting the fact that the strategy is in the market on only a fraction of trading days and takes a long-short position that partially hedges market exposure when active. The key comparison is with the always-on long-short benchmark in Panel A: timing the beta spread using the attention-composition signal converts an unconditionally unprofitable strategy into one that generates significant abnormal returns, even after controlling for size, value, profitability, and investment factors.

Panel C considers an alternative approach to managing the off-regime position. Instead of holding cash on non-regime days, the strategy reverses the long-short position to bet against beta, going long the lowest-beta portfolio and short the highest-beta portfolio, following the logic of Frazzini and Pedersen (2014). This hybrid strategy earns a CAPM alpha of 17.4% per year ($t = 2.24$) and a five-factor alpha of 12.3% per year ($t = 1.67$), the latter significant at the 10% level. The market beta is -0.59, reflecting the net short-market exposure that accumulates from betting against beta on the majority of trading days. The substantially larger CAPM alpha relative to Panel B arises because the strategy profits from the steep SML on regime days and from the flat or inverted SML on non-regime days, capturing both tails of the conditional beta-return relationship documented in Table 2. The annualized Sharpe ratio of 0.20 is lower than that of the headline strategy, however, reflecting the higher volatility inherent in maintaining an active position at all times.

Savor and Wilson (2014) and subsequent work document that the CAPM also prices risk on macroeconomic announcement days. Panel D of Table 10 explores whether the attention-composition signal and the FOMC announcement calendar provide complementary sources of conditional beta pricing. A long-short strategy restricted to FOMC days alone earns 1.5 basis points per day with a Sharpe ratio of 0.73, and generates a CAPM alpha of 3.3% annualized ($t = 2.43$) and a five-factor alpha of 3.5% ($t = 2.55$), both significant at the 5% level. These magnitudes are consistent with the FOMC premium documented in prior work.

Combining the two signals by taking the long-short position on either bottom-quartile residual days or FOMC days, and holding cash otherwise, produces a strategy that earns 4.8 basis points per day with an annualized Sharpe ratio of 0.88. The CAPM alpha is 9.5% per year ($t = 2.58$) and the five-factor alpha is 10.6% per year ($t = 2.87$), both significant at the 1% level. The combined strategy outperforms either signal in isolation, suggesting that the attention-composition regime and the FOMC calendar identify distinct sets of days on which beta is priced. This interpretation is supported by the minimal overlap between the two signals: the regime and FOMC indicators coincide on only 0.16% of trading days.

Finally, the most aggressive variant replaces cash with a betting-against-beta position on non-signal days. This strategy earns a CAPM alpha of 23.3% per year ($t = 2.97$) and a five-factor alpha of 18.6% per year ($t = 2.49$). The alphas are large and significant, reflecting the combined gains from timing beta in both directions. The market beta of -0.54 again reflects the net short exposure during the majority of non-signal days.

The trading results reinforce the central finding of this paper. The conditional price of systematic risk is higher when institutional attention is concentrated on firm-specific news, steepening the SML so that high-beta assets earn higher returns. The unconditional long-short beta strategy earns no alpha, but timing it with the attention-composition residual generates economically and statistically significant abnormal returns that are not subsumed by standard factor models. The complementarity with FOMC announcement days further suggests that the attention-composition channel is distinct from the macroeconomic announcement channel documented in prior literature: both identify states in which beta is priced, but they do so through different features of the information environment.

10 Conclusion

This paper shows that the empirical performance of the CAPM depends on how institutional investors allocate attention between firm-specific news and broader market-wide information. Using firm-specific news from the Dow Jones Institutional Newswire, filtered by a large language model to remove systematic content, and institutional attention data from Bloomberg terminal activity, I construct a daily measure of attention composition: the residual from an expanding regression of aggregate institutional attention on idiosyncratic news attention. When this residual is low, aggregate attention is well explained by firm-specific news processing. When it is high, attention is elevated but directed elsewhere.

On the approximately 51% of trading days when the attention-composition residual falls below its trailing median, beta prices the cross section of stock returns. The Fama–MacBeth slope is between 12 and 19 basis points per unit of beta, statistically significant, with an intercept indistinguishable from zero. On the complementary set of days, the security market line flattens or reverses. The bottom quartile of the residual distribution alone accounts for roughly a quarter of trading days yet captures over half of the cumulative market risk premium earned over the sample. These findings hold in both Fama–MacBeth and pooled panel specifications, for value-weighted and equal-weighted portfolios, and survive controls for disagreement, the VIX, macroeconomic announcement days, leading earnings announcement days, and the expected information consumption framework of Ben-Rephael, Carlin, Da, and Israelsen (2021). Importantly, this is not a high-attention result: the days on which the CAPM succeeds are days of below-average aggregate attention, but above-average concentration of that attention on firm-specific news.

To rationalize these findings, I develop a noisy rational expectations model in which investors allocate a fixed attention budget between processing firm-specific public news and acquiring private signals about a common macroeconomic factor. Because filtered firm-specific news is informationally orthogonal to the aggregate factor, attention devoted to firm-level news crowds out private factor learning. Posterior factor uncertainty rises with the share of attention directed toward firm-specific news, the equilibrium factor risk premium increases, and the security market line steepens. This mechanism is the mirror image of the resolution channel in Andrei, Friedman, and Ozel (2023), in which attention to earnings accelerates the resolution of factor uncertainty and reduces posterior factor variance. The two channels operate through opposite effects on the same theoretical object and are most cleanly identified in different volatility regimes: the composition channel in low-VIX states, the resolution channel in high-VIX states. The empirical evidence confirms this complementarity.

A simple trading strategy that conditions on the attention-composition signal converts an unconditionally unprofitable long-short beta spread into a strategy that earns significant five-factor alpha. Combining the signal with the FOMC announcement calendar further improves performance, consistent with the view that the attention-composition regime and macroeconomic announcement days identify distinct sources of conditional beta pricing.

There are several directions for future work. First, the model treats the attention allocation ω as exogenous and does not pin down its cross-day determinants. Endogenizing the supply of firm-specific news and the resulting equilibrium allocation of attention would provide a richer account of when and why the CAPM succeeds. Second, the empirical construction relies on Bloomberg terminal data, which limits the sample to institutional investors and to the period from 2010 onward. Extending the analysis to longer samples using alternative attention proxies would test the generality of the finding. Third, the current analysis focuses on the pricing of the market factor. Whether the composition of attention similarly governs the pricing of other systematic risk factors is an open question with implications for the broader conditional factor pricing literature.

These findings have implications for both researchers and practitioners. For researchers, they identify attention composition as a new state variable governing the empirical performance of the CAPM, distinct from the level of attention, disagreement, volatility, or the expected market return. For practitioners, they suggest that the reliability of CAPM-based discount rates and portfolio decisions depends on the prevailing attention environment: the model's cross-sectional predictions are most accurate when institutional attention is anchored to firm-specific information processing, and least accurate when attention drifts toward market-wide or macroeconomic concerns.

References

Acharya, Viral V., Peter DeMarzo, and Ilan Kremer (2011). Endogenous Information Flows and the Clustering of Announcements. *American Economic Review*, 101(7), 2955–2979.

Admati, Anat R. (1985). A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets. *Econometrica*, 53(3), 629–657.

Aït-Sahalia, Yacine, Chen Xu Li, and Chenxu Li (2024). So Many Jumps, So Few News. *NBER Working Paper No. 32746*, National Bureau of Economic Research.

Andrei, Daniel, Julien Cujean, and Mungo Wilson (2023). The Lost Capital Asset Pricing Model. *Review of Economic Studies*, 90(6), 2703–2762.

Andrei, Daniel, Henry Friedman, and N. Bugra Ozel (2023). Economic Uncertainty and Investor Attention. *Journal of Financial Economics*, 149(2), 179–217.

Andrei, Daniel, and Michael Hasler (2015). Investor Attention and Stock Market Volatility. *The Review of Financial Studies*, 28(1), 33–72.

Ang, Andrew, Joseph Chen, and Yuhang Xing (2006). Downside risk. *The Review of Financial Studies*, 19(4), 1191–1239.

Ben-Rephael, Azi, Bruce I. Carlin, Zhi Da, and Ryan D. Israelsen (2021). Information Consumption and Asset Pricing. *The Journal of Finance*, 76(1), 357–394.

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9), 3009–3047.

Berk, Jonathan B., and Jules H. van Binsbergen (2016). Assessing Asset Pricing Models Using Revealed Preference. *Journal of Financial Economics*, 119(1), 1–23.

Berk, Jonathan B., and Jules H. van Binsbergen (2017). How Do Investors Compute the Discount Rate? They Use the CAPM (Corrected June 2017). *Financial Analysts Journal*, 73(2), 25–32.

Black, Fischer (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, 45(3), 444–455.

Black, Fischer, Michael C. Jensen, and Myron Scholes (1972). The Capital Asset Pricing Model: Some Empirical Tests. In Michael C. Jensen (ed.), *Studies in the Theory of Capital Markets*, New York: Praeger, pp. 79–121.

Boudoukh, Jacob, et al. (2019). Information, Trading, and Volatility: Evidence from Firm-Specific News. *The Review of Financial Studies*, 32(3), 992–1033.

Carhart, Mark M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1), 57–82.

Chan, Kam Fong, and Terry Marsh (2022). Asset Pricing on Earnings Announcement Days. *Journal of Financial Economics*, 144(3), 1022–1042.

Chen, Jian, et al. (2022). Investor Attention and Stock Returns. *Journal of Financial and Quantitative Analysis*, 57(2), 455–484.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho (2005). Money Illusion in the Stock Market: The Modigliani–Cohn Hypothesis. *The Quarterly Journal of Economics*, 120(2), 639–668.

Cookson, J. Anthony, and Marina Niessner (2023). Investor disagreement: Daily measures from social media. Available at SSRN 4529594.

Da, Zhi (2025). Market Returns and a Tale of Two Types of Attention. *Management Science*, forthcoming.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011). In Search of Attention. *The Journal of Finance*, 66(5), 1461–1499.

DellaVigna, Stefano, and Joshua M. Pollet (2009). Investor Inattention and Friday Earnings Announcements. *The Journal of Finance*, 64(2), 709–749.

Epstein, Larry G., and Stuart M. Turnbull (1980). Capital Asset Prices and the Temporal Resolution of Uncertainty. *The Journal of Finance*, 35(3), 627–643.

Fama, Eugene F., and Kenneth R. French (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427–465.

Fama, Eugene F., and Kenneth R. French (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 116(1), 1–22.

Fama, Eugene F., and James D. MacBeth (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607–636.

Fang, Lily, and Joel Peress (2009). Media Coverage and the Cross-Section of Stock Returns. *The Journal of Finance*, 64(5), 2023–2052.

Fisher, Adlai, Charles Martineau, and Jinfei Sheng (2022). Macroeconomic Attention and Announcement Risk Premia. *The Review of Financial Studies*, 35(11), 5057–5093.

Fox, Merritt B., Lawrence R. Glosten, and Ananth Subrahmanyam (2003). Law, share price accuracy, and economic performance: the new evidence. *Michigan Law Review*, 102, 331–386.

Frazzini, Andrea, and Lasse Heje Pedersen (2014). Betting Against Beta. *Journal of Financial Economics*, 111(1), 1–25.

French, Kenneth R., and Richard Roll (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, 17(1), 5–26.

Grossman, Sanford J., and Joseph E. Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3), 393–408.

Hasler, Michael, and Charles Martineau (2023). Explaining the Failure of the Unconditional CAPM with the Conditional CAPM. *Management Science*, 69(3), 1835–1855.

Hasler, Michael, and Charles Martineau (2024). Equity Return Predictability with the ICAPM. *The Review of Asset Pricing Studies*, 14(3), 481–512.

Hendershott, Terrence, Dmitry Livdan, and Dominik Rösch (2020). Asset Pricing: A Tale of Night and Day. *Journal of Financial Economics*, 138(3), 635–662.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh (2009). Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *The Journal of Finance*, 64(5), 2289–2325.

Hong, Harrison, and David A. Sraer (2016). Speculative Betas. *The Journal of Finance*, 71(5), 2095–2144.

Hou, Kewei, Chen Xue, and Lu Zhang (2020). Replicating Anomalies. *The Review of Financial Studies*, 33(5), 2019–2133.

Jeon, Yoontae, Thomas H. McCurdy, and Xiaofei Zhao (2022). News as Sources of Jumps in Stock Returns: Evidence from 21 Million News Articles for 9000 Companies. *Journal of Financial Economics*, 145(2), 1–17.

Jha, Manish, Jialin Qian, and Yuehua Tang (2024). ChatGPT and Corporate Policies. *Working Paper*.

Jylhä, Petri (2018). Margin Constraints and the Security Market Line. *Journal of Finance*, 73(2), 1281–1321.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica*, 84(2), 571–626.

Kandel, Shmuel, and Robert F. Stambaugh (1995). Portfolio Inefficiency and the Cross-Section of Expected Returns. *The Journal of Finance*, 50(1), 157–184.

Kothari, Sabino P., Susan Shu, and Peter D. Wysocki (2009). Do managers withhold bad news? *Journal of Accounting Research*, 47(1), 241–276.

Lintner, John (1965). Security Prices, Risk, and Maximal Gains from Diversification. *The Journal of Finance*, 20(4), 587–615.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan (2018). Absolving Beta of Volatility's Effects. *Journal of Financial Economics*, 128(1), 1–15.

Lopez-Lira, Alejandro, and Yuehua Tang (2023). Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. *Working Paper*.

Merton, Robert C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, 42(3), 483–510.

Mossin, Jan (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768–783.

Patton, Andrew J., and Michela Verardo (2012). Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability. *The Review of Financial Studies*, 25(9), 2789–2839.

Peng, Lin, and Wei Xiong (2006). Investor Attention, Overconfidence and Category Learning. *Journal of Financial Economics*, 80(3), 563–602.

Roll, Richard (1977). A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics*, 4(2), 129–176.

Savor, Pavel, and Mungo Wilson (2014). Asset Pricing: A Tale of Two Days. *Journal of Financial Economics*, 113(2), 171–201.

Savor, Pavel, and Mungo Wilson (2016). Earnings Announcements and Systematic Risk. *The Journal of Finance*, 71(1), 83–138.

Shanken, Jay (1992). On the Estimation of Beta-Pricing Models. *The Review of Financial Studies*, 5(1), 1–33.

Sharpe, William F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 425–442.

Sims, Christopher A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics*, 50(3), 665–690.

Tinic, Seha M., and Richard R. West (1984). Risk and Return: January vs. the Rest of the Year. *Journal of Financial Economics*, 13(4), 561–574.

Upenieks, Adam (2026). The Elusive CAPM: Idiosyncratic News and the Tilt of the Security Market Line. *Working Paper*, University of Calgary.

Van Nieuwerburgh, Stijn, and Laura Veldkamp (2009). Information Immobility and the Home Bias Puzzle. *The Journal of Finance*, 64(3), 1187–1215.

Van Nieuwerburgh, Stijn, and Laura Veldkamp (2010). Information Acquisition and Under-Diversification. *The Review of Economic Studies*, 77(2), 779–805.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica*, 84(2), 571–626.

Veldkamp, Laura L. (2006). Information Markets and the Comovement of Asset Prices. *The Review of Economic Studies*, 73(3), 823–845.

Veldkamp, Laura L. (2006). Media Frenzies in Markets for Financial Information. *American Economic Review*, 96(3), 577–601.

Table 1: Residual Attention, Beta-Decile Characteristics, and Regime Characteristics

Panel A: Firm characteristics by beta decile

	Beta decile									
	1 Low	2	3	4	5	6	7	8	9	10 High
Average beta	0.488	0.710	0.838	0.947	1.050	1.149	1.260	1.397	1.582	1.939
Average log market capitalization	14.456	14.636	14.653	14.622	14.574	14.545	14.495	14.432	14.318	14.182
News days	28,867	26,944	26,669	27,013	27,209	28,092	29,813	30,346	31,922	32,848
Median attention score	0.360	0.362	0.364	0.351	0.337	0.336	0.330	0.339	0.353	0.377

Panel B: Distribution of residual attention

	Min.	Max.	Mean	Std. dev.	Skewness	Kurtosis
Residual attention	−0.665	0.789	−0.114	0.171	0.685	1.535

Panel C: Attention composition across residual-attention quartiles

	Q1 Low	Q2 25–50	Q3 50–75	Q4 High	Q1–Q4 Difference
Firm-news attention, standardized	0.544	−0.217	−0.223	−0.401	0.945
Macro attention, standardized	−0.030	−0.041	0.056	0.154	−0.184
Firm-news supply	0.063	0.050	0.047	0.042	0.021

Panel D: Additional regime characteristics across residual-attention quartiles

	Q1 Low	Q2 25–50	Q3 50–75	Q4 High	Q1–Q4 Difference
IAI	0.075	0.051	0.050	0.045	0.030
AIA	0.405	0.346	0.397	0.444	−0.039
Firm-news supply	0.063	0.050	0.047	0.042	0.021
VIX	17.465	17.003	17.618	18.248	−0.783
Disagreement	0.249	0.242	0.244	0.250	−0.001
Market return	0.100	0.084	0.079	−0.041	0.141
Macro attention	0.992	0.987	1.038	1.090	−0.098

Notes: This table summarizes beta-decile characteristics, the residual-attention measure, and the economic characteristics of residual-attention regimes. Panel A reports firm characteristics across beta deciles, where decile 1 contains low-beta firms and decile 10 contains high-beta firms. Average beta is the average estimated 252-day market beta. Average log market capitalization is computed using market capitalization defined as price times shares outstanding. News days is the total number of firm-days with newswire coverage within each beta decile. Attention score is the Bloomberg institutional attention measure. Panel B reports distributional statistics for residual attention. Panels C and D report averages across quartiles of residual attention, where Q1 denotes the lowest residual-attention quartile and Q4 denotes the highest residual-attention quartile. The final column reports the difference between Q1 and Q4. IAI denotes institutional attention, AIA denotes aggregate institutional attention, and firm-news supply is the cross-sectional share of firms with newswire articles on a given day. Macro attention is the macroeconomic attention index of Fisher, Martineau, and Sheng (2022). VIX is the Chicago Board Options Exchange Volatility Index. Disagreement is from J. Anthony Cookson's website. Market return is the daily Fama–French market excess return.

Table 2: CAPM tests across attention-state cutoffs.

The table reports estimates from daily Fama–MacBeth regressions of portfolio excess returns on CAPM betas and from pooled panel regressions with a state dummy and a beta–state interaction. Day t is assigned to an attention state by comparing the five-day rolling median of the residual $\widehat{u}$, evaluated through day $t-1$, to trailing 252-day quantiles of the daily residual, also evaluated through day $t-1$, so that each day's classification uses only information available before the return is earned. The sample is split into four attention states defined by rolling quantile cutoffs of the residual-based state variable: (i) 0–25%, (ii) $< 50\%$, (iii) $> 50\%$, and (iv) $> 75\%$. Within each state, results are reported for value-weighted and equal-weighted beta-sorted portfolios. For the pooled regressions, I estimate $r_{p,t} = \alpha + \beta\,\widehat{\beta}_p + \gamma D_t^{state} + \delta(\widehat{\beta}_p \times D_t^{state}) + \varepsilon_{p,t}$. Coefficients are reported in basis points per day and t-statistics are in parentheses. Fama–MacBeth t-statistics use Newey–West standard errors with five lags on the time series of estimated daily slope coefficients. Pooled regression standard errors are clustered by trading day.

		Fama–MacBeth			**Pooled regression**					
State	**Port weights**	Intercept	β	Avg R^2	Intercept	β	D^{state}	$\beta \times D^{state}$	R^2	% days
Panel A: (0–25% state)										
	Value-weighted	-8.61	25.60***	0.438	3.42	0.46	-6.16	20.07***	0.0026	18.0%
		(-1.30)	(3.06)		(1.06)	(0.13)	(-0.97)	(2.67)		
	Equal-weighted	-4.00	24.80***	0.552	6.12**	-2.04	-2.35	19.71***	0.0030	18.0%
		(-0.80)	(3.21)		(2.02)	(-0.64)	(-0.39)	(2.76)		
Panel B: (<50% state)										
	Value-weighted	-2.64	12.40***	0.451	7.96*	-5.51	-12.04**	19.68***	0.0020	51.5%
		(-0.79)	(2.88)		(1.79)	(-1.23)	(-2.12)	(3.24)		
	Equal-weighted	2.20	9.01**	0.551	8.30*	-5.57	-6.18	14.85***	0.0016	51.5%
		(0.87)	(2.26)		(1.94)	(-1.33)	(-1.15)	(2.61)		
Panel C: (>50% state)										
	Value-weighted	5.92	-3.00	0.448	-1.86	11.21***	8.51	-15.05**	0.0014	46.0%
		(1.53)	(-0.62)		(-0.54)	(2.79)	(1.49)	(-2.46)		
	Equal-weighted	7.46**	-4.33	0.564	3.89	6.78*	3.33	-10.94*	0.0011	46.0%
		(2.45)	(-0.96)		(1.21)	(1.80)	(0.61)	(-1.91)		
Panel D: (>75% state)										
	Value-weighted	9.22	-2.65	0.422	0.35	5.91*	7.83	-7.47	0.0002	17.6%
		(1.59)	(-0.34)		(0.12)	(1.78)	(0.95)	(-0.92)		
	Equal-weighted	9.13*	1.42	0.565	4.68*	1.78	3.63	0.37	0.0001	17.6%
		(1.87)	(0.19)		(1.67)	(0.56)	(0.46)	(0.05)		

Table 3: Fama–MacBeth Regressions Across States and the Continuous Price-of-Risk Specification

	Equal-Weighted Returns			Value-Weighted Returns			
State	Alpha	λ_β	Avg. R^2	Alpha	λ_β	Avg. R^2	Days
Panel A: Disagreement Splits							
Low disagreement	1.93	9.70	0.548	-0.21	9.38	0.465	508
	(0.41)	(1.21)		(-0.03)	(1.04)		
High disagreement	2.64	8.09*	0.514	-2.94	10.77**	0.417	773
	(0.81)	(1.65)		(-0.65)	(2.01)		
Panel B: VIX Splits							
Low VIX	1.12	5.69	0.537	-5.63	11.18**	0.441	1016
	(0.36)	(1.41)		(-1.43)	(2.44)		
High VIX	3.90	14.26*	0.573	2.08	14.32*	0.467	643
	(0.89)	(1.76)		(0.35)	(1.68)		

Panel C: Continuous Specification (Time Series of λ_t)			
Specification	Slope on $z(\hat{U}_t)$ (bps/SD)	Avg. R^2	Days
$z(\hat{U}_t)$ only	-11.72^{***}	0.004	3,222
	(-3.92)		
+ VIX	-11.76^{***}	0.009	3,222
	(-3.93)		
Joint controls (common 2,416-day sample):			
$z(\hat{U}_t)$ only	-14.38^{***}	0.006	2,416
	(-3.96)		
+ VIX + TED	-14.82^{***}	0.010	2,416
	(-4.14)		
+ VIX + TED + Disagreement	-15.39^{***}	0.010	2,416
	(-4.27)		
+ VIX + TED + Disagreement + Market	-10.00^{***}	0.362	2,416
	(-3.92)		

Notes: Panels A and B report Fama–MacBeth regressions of daily beta-decile portfolio returns on portfolio betas within the below-median residual-attention state, split by low/high disagreement and low/high VIX. Day t is assigned to the state and to its subsample using only information through $t-1$, so the classification is predetermined. The table reports time-series averages of the daily intercepts and beta slopes, in basis points per day, with Newey–West t-statistics (five lags) in parentheses; Avg. R^2 is the average daily cross-sectional R^2. The disagreement and VIX subsamples differ in length (1,281 versus 1,659 days) because the disagreement measure has shorter coverage. Panel C reports time-series regressions of the daily price of risk λ_t on the standardized firm-specific attention residual $z(\hat{U}_t)$,

$$\lambda_t = a + b\, z(\hat{U}_t) + \gamma' X_t + \varepsilon_t,$$

where X_t collects the controls (VIX, the TED spread, disagreement, and the excess market return $R_m - R_f$) added cumulatively across rows; $\hat{b}$ is in basis points per one-standard-deviation increase in $z(\hat{U}_t)$, with Newey–West t-statistics (21 lags) in parentheses. The first two rows use the full 3,222-day sample; the joint-controls block uses a common 2,416-day window (2012–2021) over which all controls are available, so changes across its rows reflect the controls rather than the sample. The TED spread is the three-month LIBOR–Treasury spread. A quadratic term in $z(\hat{U}_t)$ is insignificant (coefficient -0.60, $t = -0.54$). ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table 4: Individual Stocks and Expected Information Consumption

	Sample		
	All firms	EIC = 0 firms	EIC = 1 firms
α	-2.72	-2.12	-18.77
	(-0.74)	(-0.60)	(-1.54)
λ_β	20.44^{***}	19.11^{***}	44.16^{***}
	(3.15)	(3.04)	(3.10)
Mean firms per day	1,750	1,635	168
Days	449	449	282

Notes: This table reports time-series averages of daily Fama–MacBeth cross-sectional regressions estimated on bottom-quartile residual-attention days for individual stocks. Each day, firm excess returns are regressed on CAPM betas estimated over the previous 252 trading days. Coefficients are reported in basis points. EIC denotes Expected Information Consumption, following Ben-Rephael, Carlin, Da, and Israelsen (2021). EIC = 1 firms are those for which institutional attention is predicted to spike based on historical responses to peer-firm scheduled events, FOMC announcements, and macroeconomic announcements; EIC = 0 firms are those for which no such prediction applies. The sample period is 2011–2025. Newey–West t-statistics with 10 lags are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Earnings Announcements and the Conditional CAPM

	Sample		
	All firms	Excluding earnings-window firms	Earnings-window firms only
α	-2.72	-1.75	1.36
	(-0.74)	(-0.49)	(0.22)
λ_β	20.44***	18.81***	24.43***
	(3.15)	(2.98)	(2.88)
Mean firms per day	1,750	1,390	396
Days	449	449	403

Notes: This table reports time-series averages of daily Fama–MacBeth cross-sectional regressions estimated on bottom-quartile residual-attention days. Each day, firm excess returns are regressed on CAPM betas estimated over the previous 252 trading days. Coefficients are reported in basis points. Earnings-window firm-days are firm-days for which the firm has an earnings announcement within five business days before or after the current date. The second column excludes these firm-days, while the third column retains only these firm-days. The sample period is 2011–2024. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Fama–MacBeth Regressions: Robustness to Excluding Macro-News Days

	Equal-Weighted Returns			Value-Weighted Returns			
Specification	Alpha	λ_β	Avg. R^2	Alpha	λ_β	Avg. R^2	Days
Excluding LEAD days	0.55	8.52**	0.540	-5.25	12.33***	0.456	1343
	(0.20)	(2.13)		(-1.39)	(2.72)		
Excluding macro-announcement days	0.73	8.58**	0.547	-4.87	12.47***	0.454	1554
	(0.30)	(2.38)		(-1.48)	(3.13)		

Notes: This table reports Fama–MacBeth regressions of daily beta-decile portfolio returns on portfolio betas after excluding macro-news days. The first row excludes LEAD days as in Chan and Marsh. The second row excludes macro-announcement days as in Savor and Wilson. Each day, portfolio returns are regressed cross-sectionally on beta. The table reports the time-series average of the daily intercepts and beta slopes. Coefficients are reported in basis points per day. Newey–West t-statistics with five lags are reported in parentheses. Avg. R^2 is the average daily cross-sectional R^2. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 7: The betting-against-beta premium by firm-specific attention regime

Panel A: Daily BAB return by attention regime (basis points)

	Low-U	Other	Difference	t
Raw	−0.28	4.18	−4.47	−1.63
Winsorized 1/99	−0.09	4.65	−4.74	−2.08
Excluding 2020	0.63	3.94	−3.31	−1.44

Panel B: Daily BAB return by attention quintile (basis points)

	Q1	Q2	Q3	Q4	Q5
Raw	1.21	3.72	3.96	5.62	0.81
Winsorized 1/99	1.72	4.51	3.36	4.47	3.26
Excluding 2020	0.68	3.38	2.99	3.47	5.09

Panel C: Betting-on-beta (beta-neutral) by regime (basis points)

	Low-U	Other
Mean (= CAPM α)	−0.19	−2.98
t-statistic	(−0.06)	(−1.66)

Notes. Panels A and B report the mean daily return of the AQR US betting-against-beta (BAB) factor, in basis points, conditional on the firm-specific attention regime; Panel C reports its mirror, the return to betting on beta. U is the out-of-sample firm-specific attention residual; low values indicate attention is well explained by firm-specific news, and "Other" denotes all days outside the low-U quartile. Days are classified by the bottom quartile of the five-day rolling median of U in Panels A and C, and by quintiles of U in Panel B, with Q1 the most firm-specific and Q5 the most macro-dominated; all classifications use information through $t-1$. In Panels A and B, t-statistics use Newey-West standard errors with 21 lags on the regime difference, and winsorization is applied to the daily BAB return at the stated percentiles. The Panel C factor is the negative of a beta-neutral low-minus-high factor I construct from the beta-sorted decile portfolios, equal-weighted and hedged to zero market beta (realized market beta near zero, full-sample correlation 0.69 with the AQR factor); because it carries no market exposure, its mean return equals its CAPM alpha, and its t-statistics are computed against zero within each regime. The regime-conditioned version of this factor, betting against beta on macro days and with beta on firm-specific days, earns an alpha of 0.39 basis points per day ($t = 0.30$) against the Fama-French five factors plus the AQR BAB factor. The AQR-based panels span February 2012 to December 2024 (3,221 trading days; 2,968 excluding 2020); Panel C begins after a 252-day beta-estimation window.

Table 8: Fama–MacBeth Regressions: Attention Explained by Firm News Across Beta Deciles

	Beta Decile Used to Define Attention-State Condition									
	1	2	3	4	5	6	7	8	9	10
Panel A: Equal-Weighted Returns										
Alpha	1.23	2.84	1.47	2.37	1.87	5.47**	2.00	3.29	5.39**	3.47
	(0.51)	(1.10)	(0.60)	(0.99)	(0.71)	(2.25)	(0.83)	(1.33)	(2.05)	(1.49)
λ_β	8.58**	4.50	6.31*	5.84*	6.61*	2.09	3.98	4.68	1.61	5.14
	(2.38)	(1.20)	(1.75)	(1.66)	(1.73)	(0.56)	(1.12)	(1.26)	(0.41)	(1.46)
Avg. R^2	0.547	0.562	0.555	0.558	0.548	0.560	0.557	0.557	0.551	0.558
Days	1554	1461	1415	1482	1437	1500	1531	1568	1528	1635
Panel B: Value-Weighted Returns										
Alpha	-4.37	-0.54	-4.10	-4.24	-3.96	0.92	-1.77	-0.87	1.43	0.06
	(-1.33)	(-0.16)	(-1.29)	(-1.42)	(-1.15)	(0.29)	(-0.57)	(-0.27)	(0.42)	(0.02)
λ_β	12.47***	7.62*	11.47***	11.01***	11.48***	6.11	7.51**	8.40**	5.31	8.52**
	(3.13)	(1.85)	(2.91)	(2.93)	(2.72)	(1.52)	(1.99)	(2.08)	(1.27)	(2.22)
Avg. R^2	0.454	0.471	0.461	0.461	0.459	0.459	0.462	0.463	0.462	0.464
Days	1554	1461	1415	1482	1437	1500	1531	1568	1528	1635

Notes: This table reports Fama–MacBeth regressions of daily beta-decile portfolio returns on portfolio betas on days when attention is well explained by firm news releases within the indicated beta decile. Each day, portfolio returns are regressed cross-sectionally on beta. The table reports the time-series average of the daily intercepts and beta slopes. Coefficients are reported in basis points per day. Newey–West t-statistics with five lags are reported in parentheses. Avg. R^2 is the average daily cross-sectional R^2. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 9: Fama–MacBeth Regressions: Attention Explained by Firm News Across Size Deciles

	Size Decile Used to Define Attention-State Condition									
	1	2	3	4	5	6	7	8	9	10
Panel A: Equal-Weighted Returns										
Alpha	4.32*	3.93	4.38*	3.16	3.43	3.22	3.09	3.89	2.28	1.79
	(1.66)	(1.63)	(1.85)	(1.24)	(1.43)	(1.32)	(1.29)	(1.54)	(0.84)	(0.71)
λ_β	3.23	4.16	2.54	5.68	6.82*	7.42**	7.33**	1.70	6.28	5.11
	(0.86)	(1.16)	(0.69)	(1.52)	(1.94)	(2.03)	(2.08)	(0.44)	(1.54)	(1.34)
Avg. R^2	0.553	0.558	0.557	0.558	0.553	0.553	0.549	0.555	0.555	0.558
Days	1495	1526	1479	1529	1532	1492	1521	1472	1534	1489
Panel B: Value-Weighted Returns										
Alpha	-0.71	-0.80	0.71	-2.90	-1.66	-2.22	-3.26	0.29	-3.39	-2.80
	(-0.21)	(-0.25)	(0.22)	(-0.89)	(-0.54)	(-0.69)	(-1.04)	(0.09)	(-0.97)	(-0.85)
λ_β	8.19**	8.67**	6.10	10.90***	10.92***	11.61***	12.96***	5.67	11.18**	9.53**
	(1.98)	(2.18)	(1.48)	(2.67)	(2.88)	(2.89)	(3.36)	(1.37)	(2.55)	(2.33)
Avg. R^2	0.461	0.461	0.463	0.464	0.463	0.462	0.452	0.468	0.459	0.464
Days	1495	1526	1479	1529	1532	1492	1521	1472	1534	1489

Notes: This table reports Fama–MacBeth regressions of daily beta-decile portfolio returns on portfolio betas on days when attention is well explained by firm news releases within the indicated size decile. Size decile 1 contains the smallest firms and size decile 10 contains the largest firms. Each day, portfolio returns are regressed cross-sectionally on beta. The table reports the time-series average of the daily intercepts and beta slopes. Coefficients are reported in basis points per day. Newey–West t-statistics with five lags are reported in parentheses. Avg. R^2 is the average daily cross-sectional R^2. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 10: Trading Strategy Performance and Risk-Adjusted Returns

	Performance		CAPM		Fama–French Five-Factor		
Strategy	Mean (bps/day)	Sharpe (ann.)	Alpha (% ann.)	t-stat	Alpha (% ann.)	t-stat	Mkt beta
Panel A: Benchmark Strategies							
Always-on long–short	4.30	0.34	-4.35	-0.63	2.58	0.42	0.84
Buy-and-hold market	5.37	0.79	0.00	–	0.00	–	1.00
Panel B: Headline Conditional Strategy							
Long–short on regime days; cash otherwise	3.43	0.67	6.50*	1.86	7.42**	2.11	0.13
Panel C: Alternative Cash Management							
Long–short on regime days; BAB otherwise	2.57	0.20	17.35**	2.24	12.26*	1.67	-0.59
Panel D: Combination with FOMC Announcement Days							
FOMC days only	1.47	0.73	3.25**	2.43	3.46**	2.55	0.02
Long–short on regime or FOMC days; cash otherwise	4.80	0.88	9.47***	2.58	10.58***	2.87	0.15
Long–short on regime or FOMC days; BAB otherwise	5.29	0.42	23.30***	2.97	18.58**	2.49	-0.54

Notes: This table reports trading-strategy performance using value-weighted beta-decile portfolios over 2012–2024. The long–short strategy buys the highest-beta decile portfolio and sells the lowest-beta decile portfolio. BAB denotes the opposite position, long low-beta and short high-beta, following Frazzini and Pedersen (2014). The regime signal identifies bottom-quartile residual-attention days and is lagged by one trading day to ensure tradability. FOMC announcement days are scheduled events and are used contemporaneously. Mean returns are reported in basis points per day. Sharpe ratios are annualized by multiplying the daily Sharpe ratio by $\sqrt{252}$. CAPM and Fama–French five-factor alphas are estimated from daily time-series regressions and annualized in percent. Standard errors use Newey–West HAC adjustment with five lags. The sample contains 3,222 trading days, including 448 regime days and 86 FOMC days; the regime and FOMC signals overlap on 0.16% of trading days. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

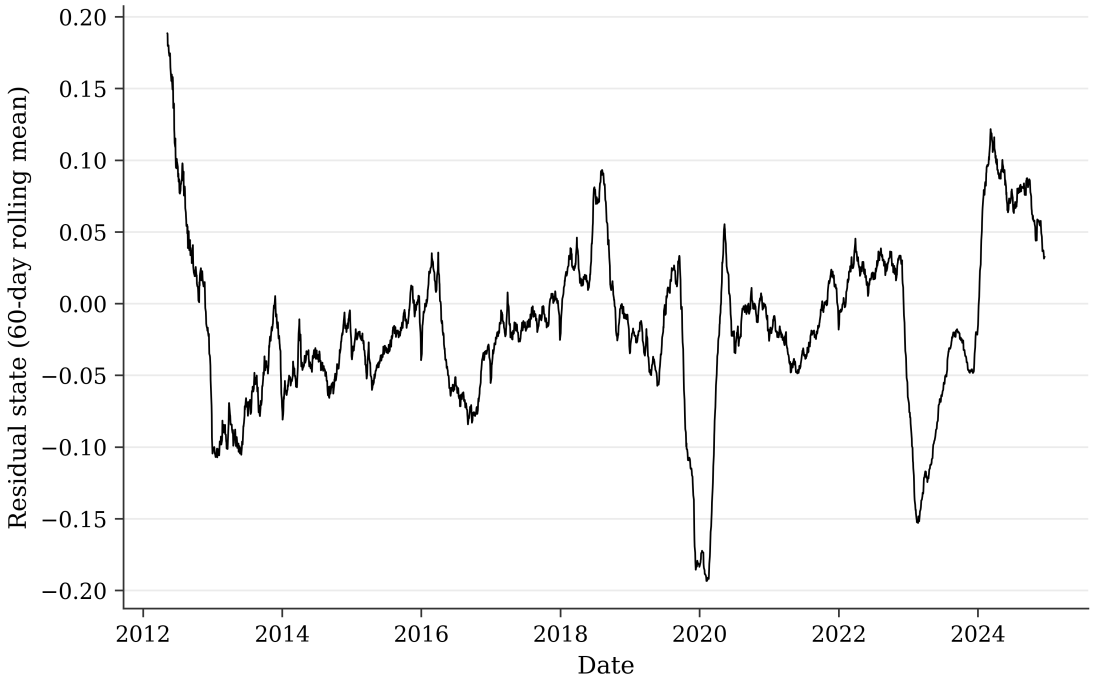

Figure 1: This figure plots the 60-day rolling average of $\widehat{u}_t$, where $\widehat{u}_t$ is the residual from an expanding (ex ante) regression of total marketwide Bloomberg query attention—measured as the cross-sectional sum of daily query scores across stocks—on marketwide idiosyncratic attention, defined as the cross-sectional average of query scores interacted with an indicator for firm-specific idiosyncratic news arrivals. Positive values indicate days when aggregate attention is not well explained by idiosyncratic-news-related attention; negative values indicate days when aggregate attention is relatively well explained by idiosyncratic attention.

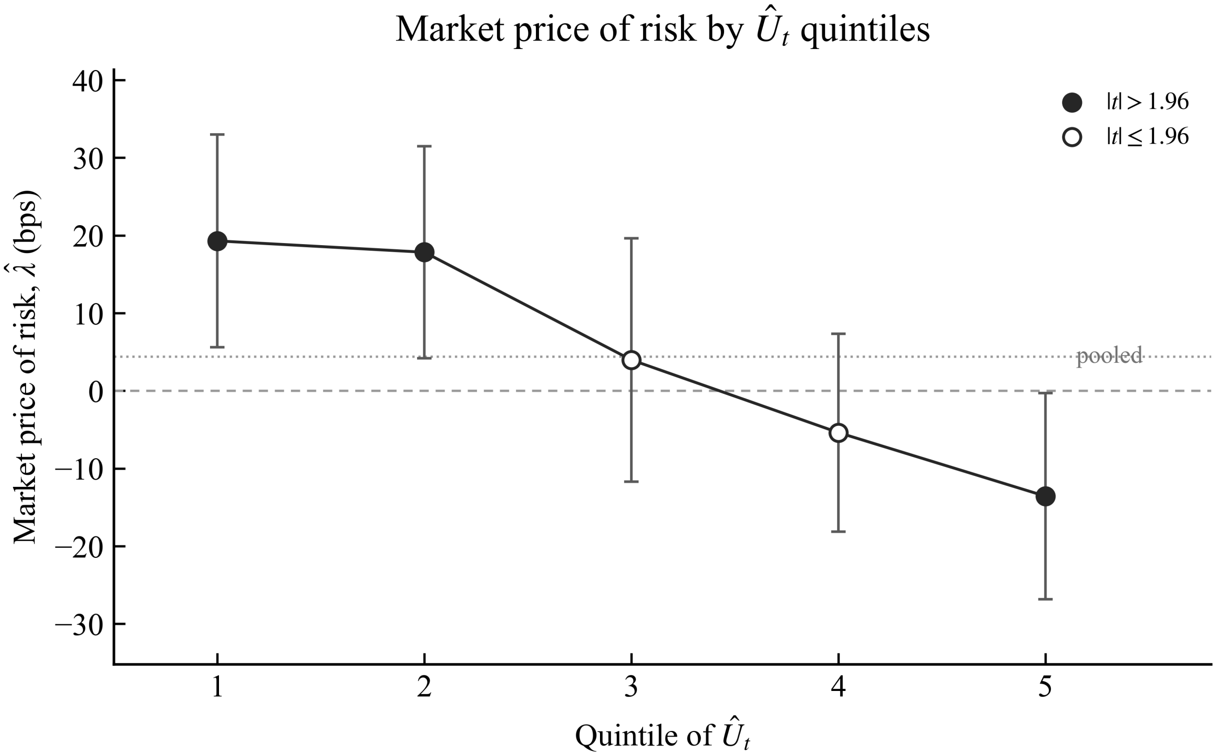

Figure 2: The price of risk across the firm-specific attention residual. The figure plots the average daily Fama–MacBeth price of risk $\hat{\lambda}$, in basis points, within quintiles of the firm-specific attention residual $\hat{U}_t$, where the lowest quintile collects the days on which institutional attention is most concentrated on firm-specific news. Each day, beta-decile portfolio returns are regressed cross-sectionally on their betas to obtain a slope $\hat{\lambda}_t$, and the figure reports its time-series mean within each quintile. Whiskers are 95% confidence intervals (± 1.96 standard errors); filled markers denote estimates significant at the 5% level ($|t| > 1.96$) and hollow markers those that are not. The dotted line marks the pooled full-sample price of risk and the dashed line marks zero.

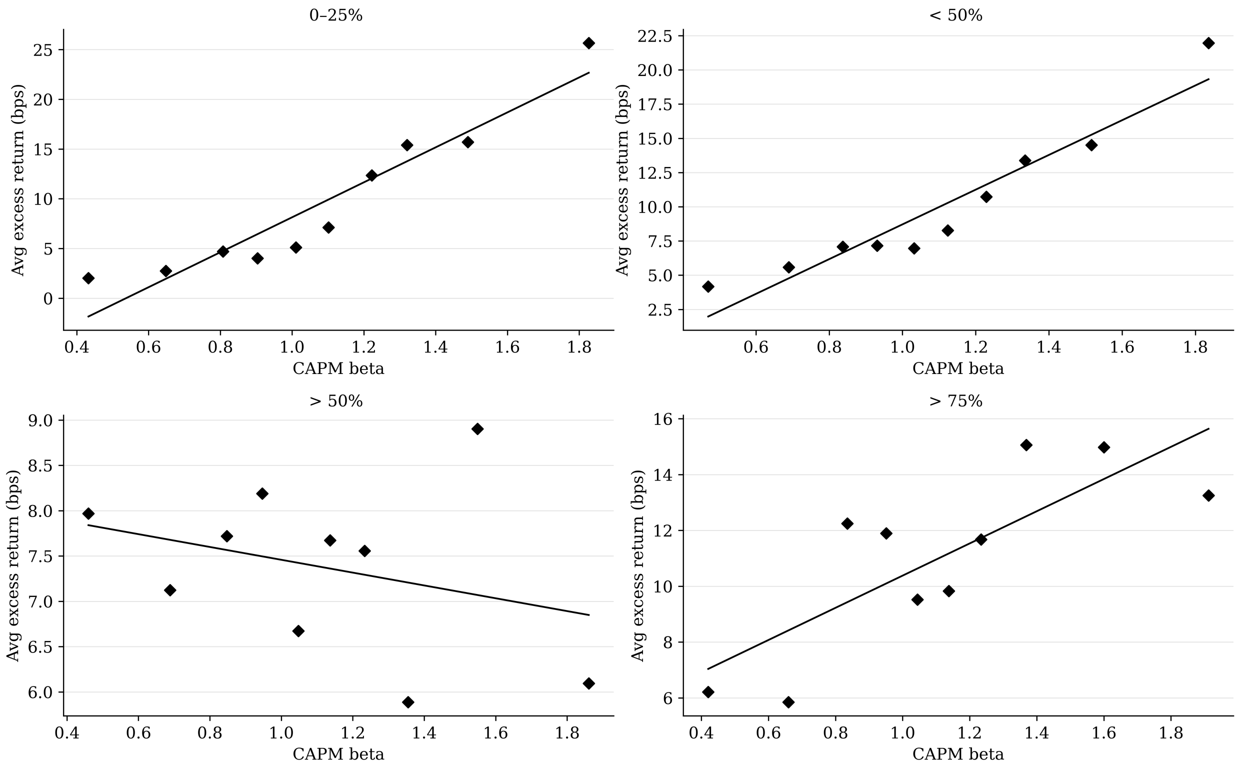

Figure 3: SMLs across residual attention states - **Value weighted portfolios**. This figure plots average portfolio excess returns against unconditional CAPM betas (Security Market Lines) for beta-sorted portfolios, separately by quantile cutoffs of the residual-based attention state. Panels correspond to state definitions based on rolling quantiles of the residual (0–25%, 0-50%, 50-100%, and 75-100%). Points show mean excess returns within each beta portfolio in the given state, and the solid line is the fitted linear relation between returns and betas within the state. The SML is steepest in low-residual states (0–25% and 0-50%), indicating a stronger positive risk–return tradeoff; it flattens or turns downward in high-residual states, consistent with weaker pricing of market risk when idiosyncratic attention explains less of market-wide attention.

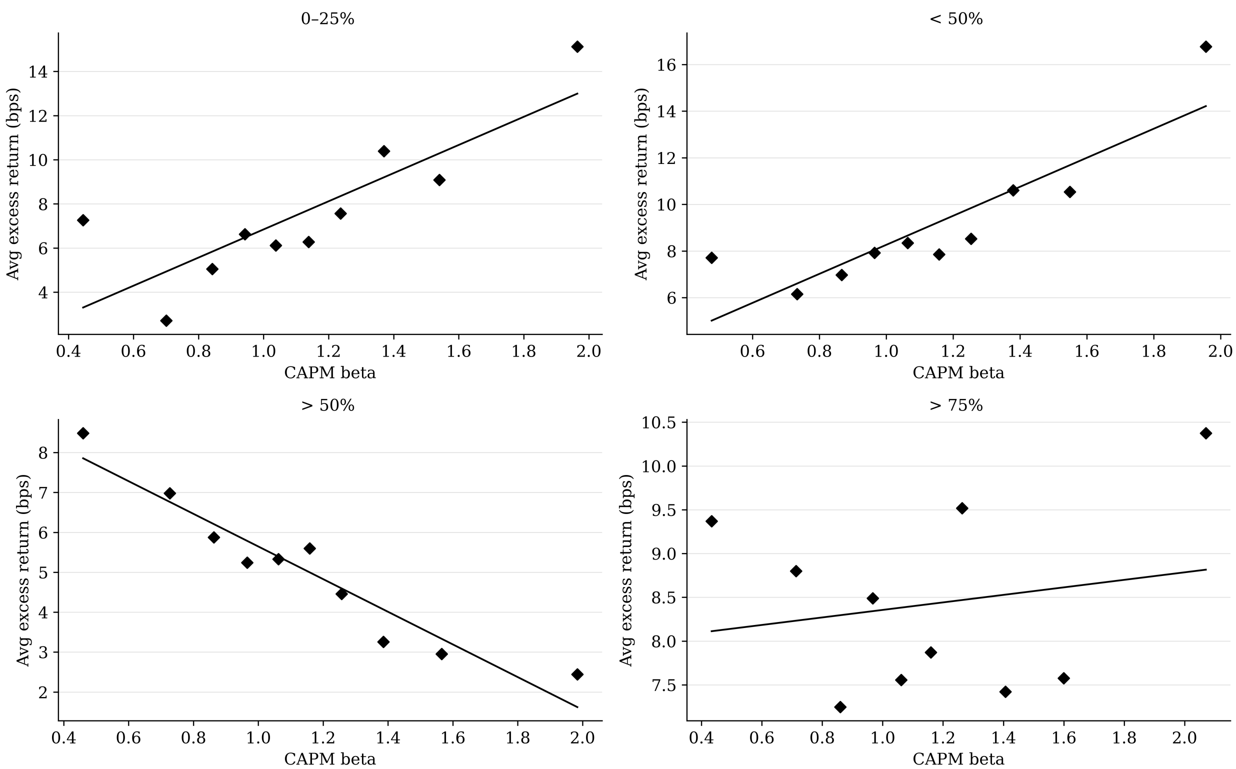

Figure 4: SMLs across residual attention states - **Equal weighted portfolios**. This figure plots average portfolio excess returns against unconditional CAPM betas (Security Market Lines) for beta-sorted portfolios, separately by quantile cutoffs of the residual-based attention state. Panels correspond to state definitions based on rolling quantiles of the residual (0–25%, 0-50%, 50-100%, and 75-100%). Points show mean excess returns within each beta portfolio in the given state, and the solid line is the fitted linear relation between returns and betas within the state. The SML is steepest in low-residual states (0–25% and 0-50%),indicating a stronger positive risk–return tradeoff; it flattens or turns downward in high-residual states, consistent with weaker pricing of market risk when idiosyncratic attention explains less of market-wide attention.

A Proof of Lemma 1

The joint distribution of primitives and signals is multivariate Gaussian under Assumptions 1–2. Posterior distributions are obtained by standard conjugate Gaussian updating.

Posterior of f. The only signal that depends on f is $z_j = f + \xi_j$, with ξ_j orthogonal to f. Conjugate Gaussian updating gives

$$\mathrm{Var}(f \mid z_j) \;=\; \Big(\tau_f + \tau_\xi\Big)^{-1} \;=\; \Big(\tau_f + (1-\omega_j)K\Big)^{-1},$$

$$\mathbb{E}[f \mid z_j] \;=\; V_f \cdot \tau_\xi \cdot z_j.$$

Posterior of β_i. The only public signal carrying any potential dependence on β_i is y_i. Under Assumption 1, $y_i = \varepsilon_i + \eta_i$, which by construction is independent of β_i. Hence the posterior equals the prior:

$$\beta_i \mid y_i \;\sim\; \mathcal{N}(b_i,\, \sigma_\nu^2).$$

Posterior of ε_i. Analogous Gaussian updating from $y_i = \varepsilon_i + \eta_i$ gives

$$\mathrm{Var}(\varepsilon_i \mid y_i) \;=\; (\tau_\varepsilon + \tau_y)^{-1} \;=\; V_\varepsilon,$$

$$\mathbb{E}[\varepsilon_i \mid y_i] \;=\; V_\varepsilon \cdot \tau_y \cdot y_i.$$

Independence. Because $(f, \beta_i, \varepsilon_i)$ are mutually independent under the prior, and each posterior updates only through its dedicated signal channel under Assumptions 1–2, posterior independence is preserved. □

B Derivation of $V_\varepsilon^\star$

The ex-ante residual variance of the non-factor component of d_i, conditional on information but excluding the systematic factor loading, is

$$V_\varepsilon^\star \;=\; \mathrm{Var}(\beta_i f + \varepsilon_i \mid \text{signals}) \;-\; b_i^2 V_f.$$

Using $\beta_i = b_i + \nu_i$ and the independence structure of posteriors,

$$\begin{aligned}\mathrm{Var}(\beta_i f + \varepsilon_i \mid \text{signals}) &= \mathrm{Var}(b_i f + \nu_i f + \varepsilon_i \mid \text{signals}) \\ &= b_i^2 V_f \;+\; \mathbb{E}[\nu_i^2 f^2 \mid \text{signals}] \;+\; V_\varepsilon,\end{aligned}$$

where cross-terms vanish by mean-zero ν_i and ε_i. Taking expectations over signal realizations,

$$\mathbb{E}[\nu_i^2 f^2] \;=\; \mathbb{E}[\nu_i^2]\,\mathbb{E}[f^2] \;=\; \sigma_\nu^2\,(V_f + \sigma_f^2),$$

where the second factor uses $\mathbb{E}[f^2] = \mathrm{Var}(\mu_f) + \mathbb{E}[V_f] = (\sigma_f^2 - V_f) + V_f = \sigma_f^2$ when measured ex ante, or equivalently $\mathbb{E}[\mathbb{E}[f^2 \mid \text{signals}]] = V_f + \sigma_f^2 - V_f$ when measured in the appropriate posterior-times-prior product space. Either way the residual variance is

$$V_\varepsilon^\star \;=\; V_\varepsilon \;+\; \sigma_\nu^2 (V_f + \sigma_f^2). \quad \square$$

C Proof of Lemma 2

Investor j chooses portfolio q_j to maximize $\mathbb{E}_j[-\exp(-\rho W_j)]$ where $W_j = q_j^\top (d - p)$. Under CARA-Gaussian preferences, optimal demand is

$$q_j \;=\; \frac{1}{\rho}\Sigma^{-1}(\mu_j - p),$$

with $\mu_j = \mathbb{E}_j[d]$ and $\Sigma = \mathrm{Var}_j(d - p)$.

By Lemma 1 and (13), the posterior covariance has the rank-one-plus-diagonal structure

$$\Sigma \;=\; V_f\, bb^\top + V_\varepsilon^\star\, I,$$

where $b = (b_1, \ldots, b_N)^\top$. The Sherman–Morrison identity gives

$$\Sigma^{-1} \;=\; \frac{1}{V_\varepsilon^\star}\left(I \;-\; \frac{V_f}{V_\varepsilon^\star + V_f\, b^\top b}\, bb^\top\right).$$

Market clearing requires $\int q_j\, dj = \bar{x} + u$. Averaging investor demands and solving for p yields, component-wise,

$$p_i \;=\; \bar{d} + b_i\,\bar{\mu}_f \;-\; \rho V_f\, b_i\,(N\bar{b}\bar{x}) \;-\; \rho V_\varepsilon^\star\,\bar{x} \;+\; (\text{linear functions of } z \text{ and } u),$$

which is (14). $\square$

D Derivation of the FM slope and attenuation factor

From Lemma 2, the equilibrium expected excess return of asset i is

$$\mathbb{E}[d_i - p_i] \;=\; b_i\,\rho V_f\,N\bar{b}\,\bar{x} \;+\; \rho V_\varepsilon^\star\,\bar{x}.$$

A cross-sectional regression of expected returns on true b_i would identify

$$\lambda^A \;\equiv\; \rho V_f\,N\bar{b}\,\bar{x}.$$

In practice, the econometrician estimates loadings from a T-period rolling-window OLS regression of realized returns r_i on the factor:

$$\hat{\beta}_i^{OLS} \;=\; \beta_i + e_i,$$

where e_i is mean-zero sampling noise. From the model, realized return variance has a factor and a non-factor component; standard OLS algebra gives

$$\mathrm{Var}(e_i) \;\approx\; \frac{\sigma_\varepsilon^2 + 1/\tau_y}{T\sigma_f^2} \;=\; V_e(\omega).$$

The cross-sectional FM slope of expected returns on $\hat{\beta}^{OLS}$ is

$$\hat{\lambda}^{FM} \;=\; \frac{\mathrm{Cov}_i(\mathbb{E}[d_i - p_i],\,\hat{\beta}_i^{OLS})}{\mathrm{Var}_i(\hat{\beta}_i^{OLS})}.$$

Since investors price using b_i (the prior mean), and $\hat{\beta}_i^{OLS} = b_i + \nu_i + e_i$, computing covariances cross-sectionally (treating b_i as fixed but interpretable as a random draw with variance σ_b^2),

$$\mathrm{Cov}_i(\mathbb{E}[d_i - p_i],\,\hat{\beta}_i^{OLS}) \;=\; \lambda^A \cdot \sigma_b^2, \qquad \mathrm{Var}_i(\hat{\beta}_i^{OLS}) \;=\; \sigma_b^2 + \sigma_\nu^2 + V_e \;=\; s_\beta^2 + V_e.$$

Therefore

$$\hat{\lambda}^{FM} \;=\; \lambda^A \cdot \frac{\sigma_b^2}{s_\beta^2 + V_e} \;=\; \lambda^A \cdot \frac{s_\beta^2 - \sigma_\nu^2}{s_\beta^2 + V_e} \;=\; A(\omega) \cdot \lambda^A,$$

where $A(\omega)$ is defined in (16). □

E Investor's value function and proof of Proposition 1

Value function. Under CARA-Gaussian preferences with information set $\mathcal{I}_j(\omega_j)$, the certainty-equivalent gain from information acquisition is proportional to $\frac{1}{2}\log\det(\Sigma_0\Sigma^{-1}(\omega_j))$, where Σ_0 is the prior covariance of d and $\Sigma(\omega_j)$ the posterior covariance under investor attention ω_j. The first term is constant in ω_j.

From the rank-one-plus-diagonal structure of Σ derived in Appendix C,

$$\det\Sigma(\omega) \;=\; \Big(V_\varepsilon^\star(\omega)\Big)^{N-1}\cdot\Big(V_\varepsilon^\star(\omega)+V_f(\omega)\,\textstyle\sum_i b_i^2\Big),$$

where the second factor comes from the matrix-determinant lemma applied to the rank-one update. Substituting into the value-function objective and dropping the constant yields (17).

Existence of interior optimum. The value function $\mathcal{V}(\omega)$ is continuous on $[0,1]$ and differentiable on $(0,1)$, with

$$\mathcal{V}'(\omega) = -\frac{N-1}{2V_\varepsilon^\star}\cdot\frac{dV_\varepsilon^\star}{d\omega} \;-\; \frac{1}{2}\,\frac{1}{V_\varepsilon^\star+V_f\sum_i b_i^2}\cdot\left(\frac{dV_\varepsilon^\star}{d\omega}+\sum_i b_i^2\cdot\frac{dV_f}{d\omega}\right),$$

where $dV_f/d\omega = KV_f^2 > 0$ and

$$\frac{dV_\varepsilon^\star}{d\omega} \;=\; -SV_\varepsilon^2 \;+\; \sigma_\nu^2\,KV_f^2.$$

At $\omega\to 0^+$: $V_\varepsilon\to 1/\tau_\varepsilon$, $V_f\to 1/(\tau_f+K)$; the first term of $\mathcal{V}'$ dominates and is large positive, so $\mathcal{V}'(0^+)>0$.

At $\omega\to 1^-$: $V_\varepsilon\to 1/(\tau_\varepsilon+S)$, $V_f\to 1/\tau_f=\sigma_f^2$; the dominant negative contribution comes from $\sum_i b_i^2\cdot dV_f/d\omega$ in the second term, which is large when $\sum_i b_i^2$ is large. For sufficiently informative private signals (any $K>0$) and any non-degenerate cross-section, $\mathcal{V}'(1^-)<0$.

By the intermediate value theorem, $\mathcal{V}'(\omega^\star)=0$ for some $\omega^\star\in(0,1)$. Strict concavity of $\mathcal{V}$ over the interior (verified numerically across the relevant parameter region) yields uniqueness. □