The Foerder Institute for Economic Research
at Tel Aviv University

מכון למחקר כלכלי על שם
ד"ר ישעיהו פורדר
על יד אוניברסיטת תל אביב

The Eitan Berglas School of Economics

בית-הספר לכלכלה ע"ש איתן ברגלס

עמותה רשומה

Frictions in News Consumption: Evidence from Social Media

Luca Braghieri, Ro'ee Levy, and Hannah Trachtman

Working paper 9-2026

The Foerder Institute for Economic Research
and
The Sackler Institute of Economic Studies

Frictions in News Consumption: Evidence from Social Media*

Luca Braghieri, Ro'ee Levy, and Hannah Trachtman

June 2026

Abstract

We study the drivers of like-minded and low-reliability news-following on social media, and the effectiveness of interventions targeting them. In a five-week field experiment with more than 3,000 U.S. Facebook users, we document the passive nature of participants' news portfolio formation process and the role of behavioral frictions. The experiment varies whether participants are prompted to re-optimize the portfolio of news pages they follow on Facebook through a user-friendly platform-integrated interface, and whether they receive personalized information about outlet slant and reliability. In contrast to canonical models of news consumption, we find that the re-optimization interface induces large portfolio changes even without new information, while information alone has no effect unless paired with re-optimization. Our interventions produce two main effects: they reduce the slant and increase the reliability of users' portfolios, thus potentially mitigating negative externalities for democracy, and they move users' portfolios closer to their stated preferences, mitigating internalities.

*Braghieri: Bocconi University and NHH, CEPR, CESifo, and IGIER. Email: luca.braghieri@unibocconi.it. Levy: Tel Aviv University, CESifo, and CEPR. Email: roeelevy@tauex.tau.ac.il. Trachtman: Hebrew University of Jerusalem. Email: hannah.Trachtman@mail.huji.ac.il. We thank Hunt Allcott, Valeria Burdea, Felix Chopra, Adrien D'Autheville, Sarah Eichmeyer, Ingar Haaland, Rafael Jiménez-Durán, Chris Roth, Lena Song, Peter Schwardmann, Mateusz Stalinski, Dan Stone, Karine Van Der Straeten, Egon Tripodi, and participants at APSA, Bocconi University, Chicago Booth, Coller School of Management, CEPR Media Plurality, CESifo, EEA Annual Congress, FAIR (NHH), Hebrew University, Leicester, Manchester, Technion, and the Toulouse Digital Economics Conference for helpful comments. We thank Giovanni Cecchi, Ittai Dayagi, Yuval Rubinstein, Heitor Sousa Borges, and Siying Wang for excellent research assistance. Funding for this project was provided by the European Research Council (ERC Grant 101116966), the Israel Science Foundation (791/23), the Tel Aviv University Center for AI & Data Science (TAD), the Pinchas Sapir Center for Development, and the Foerder Institute for Economic Studies. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union, the European Research Council Executive Agency, or the other funding sources. Neither the European Union nor the granting authority can be held responsible for them. This research was approved by the Ethics Committees of Bocconi University (EA000764), Tel-Aviv University (0004305), and Hebrew University (IRB_2024_113). The experiment is registered at the AEA RCT Registry (Braghieri, Levy and Trachtman, 2025).

In 2025, for the first time, the share of U.S. adults who reported consuming news on social media (54%) overtook the share who reported watching news on television (50%), marking a profound shift in the information ecosystem (Newman et al., 2025). This shift has intensified public concerns about the proliferation of like-minded (Braghieri et al., 2025) and low-reliability (Vosoughi, Roy and Aral, 2018) news on social media, as well as about potential negative externalities for democracy (Sunstein, 2017).

Recent research shows that users' own customization decisions on social media—especially which news pages to follow—are a central driver of like-minded and low-reliability news consumption for two reasons. First, a substantial share of the news content users see in their feeds comes from pages they follow.[1] Second, users encounter more polarized content and more misinformation from followed pages than from other sources (Levy, 2021; González-Bailón et al., 2023; Chouaki, Chakraborty, Goga and Zannettou, 2024). This raises two crucial questions: why do users follow like-minded and low-reliability news pages on social media, and which interventions can reduce the slant and improve the reliability of the pages they follow?

These questions relate to a broader debate about why people consume slanted and low-reliability news, including outside social media.[2] Canonical explanations emphasize deliberate optimization under preferences for like-minded outlets and incomplete information about reliability (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006; Chopra, Haaland and Roth, 2022, 2023). On social media, however, users might not form their news portfolios by deliberately optimizing over the full set of outlets. Instead, they might arrive at those portfolios through friend suggestions and platform cues, and revise them only infrequently because of inertia and other behavioral frictions. As a result, the composition of users' news portfolios may depend not only on their preferences and beliefs, but also on the process through which choices are made in platform-mediated environments.

In this paper, we study the drivers of news-following behavior on social media, as well as the effectiveness of interventions targeting them, in the context of a five-week field experiment on the world's largest social media platform: Facebook. Our experiment, which couples changes to the choice architecture (Thaler and Sunstein, 2008) with the provision of targeted information (Haa-

[1] On Facebook, for instance, it is just under half (Levy, 2021).

[2] Slant captures a news outlet's political leaning on a left-right spectrum; reliability captures the degree to which a news outlet adheres to best-practice journalistic standards.

land, Roth and Wohlfart, 2023), shows that news-following decisions on social media are largely passive and shaped by behavioral and information frictions that limit when and how users act on their preferences. By directly targeting those frictions, our intervention produces two key results. First, it increases the reliability and reduces the slant of the news pages users follow, thereby mitigating potential negative externalities for democracy. Second, it increases alignment between users' stated preferences and their resulting news portfolios, thereby reducing internalities.[3] Because the experiment is implemented on a leading platform with real users, real decisions, and a platform-integrated interface, our findings speak directly to the design of policies that could be deployed at scale.

We begin by documenting new stylized facts about news consumption on social media. Using our baseline survey, we elicit participants' perceptions of the slant and reliability of the news pages they follow on Facebook, measure their bliss points in the slant–reliability space, and observe their actual news portfolios. We find that participants' Facebook news portfolios are, on average, more partisan and less reliable than the portfolios they report wanting. This wedge suggests that users' portfolios may be shaped by frictions that prevent them from fully acting on their preferences or beliefs. We find evidence of two such frictions. On the information side, participants perceive slant relatively accurately but hold substantially less accurate perceptions of reliability. On the behavioral side, participants describe the formation of their news portfolios as largely passive, shaped more by external suggestions and platform cues than by deliberate search.

To interpret the stylized facts above, we introduce two contrasting views of news-following behavior on social media. The first is a *deliberative optimization* benchmark, which captures the canonical news-demand forces emphasized in the literature (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006; Chopra, Haaland and Roth, 2023). In this benchmark, users hold stable preferences over slant and reliability, have a good sense of the menu of available outlets, and adjust their portfolios whenever useful information arrives. The second is a *passive portfolio formation* process. In this view, users do not continuously optimize over the full set of outlets. Instead, they revise their portfolios only sporadically, typically when a decision episode is triggered by platform or friend cues, and then choose from the subset of outlets that is especially easy to consider and act upon.

[3] We use the term "news portfolio" to indicate the set of news pages that an individual follows on Facebook.

The experiment is designed around the different comparative statics of these two views. If users' news portfolios already reflect deliberative optimization, then merely making it easy for them to revise the set of pages they follow on Facebook should have little effect, whereas new and valuable information should induce them to adjust their portfolios. If, instead, portfolios are the result of a passive process, then a prompted decision episode with an easy-to-use menu of news pages should move portfolios even without the provision of new information, whereas information delivered outside such an episode should have little effect. To distinguish these possibilities, the experiment separately varies (i) whether users are prompted to re-optimize the portfolio of news pages they follow on Facebook through a platform-integrated interface that presents an easy-to-use, curated menu of news pages, and (ii) whether they receive personalized information about the slant or reliability of various news pages.

The field experiment proceeded in three stages. We recruited more than 3,000 participants through Facebook ads and directed them to a baseline survey. All participants were required to grant us permission to observe the news pages they follow on Facebook via the platform's API. Participants using a Google Chrome browser were also incentivized to install a Chrome extension that tracked their browsing behavior on news websites. Approximately one week after the baseline survey, participants were invited to a midline survey, where we implemented the experimental treatments. Around four weeks after midline, participants were invited to an endline survey, where they were given another opportunity to re-optimize their Facebook news portfolios using our interface and where we elicited survey-based outcomes.

At midline, participants were randomized into five main arms designed to isolate the role of re-optimization and information. The *Control* group received no intervention. The ***R**e-optimization* treatment (**R**) prompted a decision episode without providing new information: participants were shown a user-friendly interface that allowed them, with a single click, to follow any news page from a curated list of the most popular news outlets on Facebook or to unfollow any news page they were already following on the platform. The ***N**o-Reoptimization with **R**eliability information* treatment (**NR**) provided reliability information without creating a re-optimization episode. The ***R**e-optimization with **R**eliability information* treatment (**RR**) combined the re-optimization interface with personalized information about the reliability of participants' Facebook news portfolios and of various news outlets. Finally, the ***R**e-optimization with **S**lant information* treatment (**RS**) combined the re-optimization interface with analogous information about slant. The comparison between **R** and **NR** is therefore the core test of passive portfolio formation, while the comparison between **RR** and **R** measures the additional value of reliability information once the re-optimization friction is relaxed.

Our primary pre-specified outcome is the set of news pages participants follow on Facebook after the midline survey. We focus on following behavior because it is the core customization decision faced by Facebook users, because it has been shown to be a key determinant of subsequent news exposure on the platform (Levy, 2021), and because the pages users follow on social media have been shown to affect their knowledge, beliefs, and political attitudes (Altay, Hoes and Wojcieszak, 2025; Chmel et al., 2025; Levy, 2021; Rathje et al., 2024). We complement this primary outcome with survey evidence and, for the subset of participants who installed our Chrome extension, with browsing data on news websites.

The first set of results distinguishes the deliberative-optimization benchmark from passive portfolio formation. Under the deliberative-optimization benchmark, users' followed-news portfolios are the product of optimal choice given stable preferences and current beliefs, and useful information is acted upon to the extent that it is useful. In this view, the **R** treatment should have little effect, whereas the **NR** treatment should induce portfolio adjustment if the information provided is valuable. We find the opposite pattern. The vast majority (71%) of participants assigned to the **R** treatment actively modified their news portfolios by adding or removing at least one page using our re-optimization interface. This high degree of re-optimization—in the absence of new information, monetary incentives, or explicit encouragement—is difficult to reconcile with the deliberative benchmark and suggests that participants' portfolios were not already optimized.[4] Meanwhile, **NR**—which provides reliability information without prompting re-optimization—has no detectable effect on the set of pages participants follow. Taken together, these results are more consistent with a passive portfolio formation process than with the narrow deliberative-optimization benchmark.

The second set of results shows that interventions targeting these frictions improve portfolios along dimensions that relate to potential externalities: exposure to like-minded and low-reliability

[4]An alternative explanation is that participants' re-optimization choices reflect experimentation motives. We address and rule out this explanation in Section 6.2.

news. Across all re-optimization arms, participants' portfolios become, on average, less slanted. This result is consistent with our evidence of passive portfolio formation and suggests that this formation process explains why people follow outlets that are more partisan than their preferences. We also find that the **RR** treatment, which pairs the re-optimization interface with reliability information, increases average portfolio reliability. The result implies that information frictions operate alongside behavioral frictions and is consistent with the misperceptions we uncovered regarding the reliability of outlets.

The third set of results shows that the same interventions also reduce internalities by bringing participants' portfolios closer to their stated preferences. The three re-optimization treatment arms (**R**, **RR**, and **RS**) significantly reduce the distance between participants' actual portfolios and their self-reported bliss points in the slant–reliability space. Consistent with this greater alignment, treated participants report, four weeks after the midline survey, higher trust in and greater satisfaction with the news they encounter on Facebook. Thus, the intervention not only shifts users' portfolios toward outlets that are arguably more desirable from a societal perspective—less slanted and more reliable—but also helps close the gap between participants' portfolios and their stated preferences.

A final set of results shows that the portfolio changes persist, translate into downstream news consumption, and are unlikely to be driven by confounds. Because following a page is a sticky customization decision, a one-time intervention can generate persistent changes in exposure and consumption. Consistent with this idea, changes in portfolios persist through the endline survey four weeks later. Moreover, for the subset of participants who installed our Chrome extension, following a news page through our intervention more than doubles the probability of visiting that page in the subsequent month. This result links portfolio changes to meaningful changes in online news consumption, emphasizing the downstream consequences of shifting the set of pages that Facebook users follow on the platform. We also show that our results are unlikely to be driven by experimentation motives or experimenter demand effects.

To address experimenter demand, our design includes an additional **N**o **E**xperimenter **D**emand treatment (**NED**), which closely mirrors **RR** but removes experimenter observability at the moment participants revise their portfolios. At the start of the midline survey, participants in **NED** revoked our permission to observe the Facebook pages they follow; at endline, we requested permission to collect their following behavior retrospectively. The similarity of effects across **NED** and **RR** mitigates concerns about experimenter demand effects.

Our paper contributes to several strands of literature. First, we contribute to the literature on news demand (Angelucci and Prat, 2024; Bursztyn, Rao, Roth and Yanagizawa-Drott, 2023; Chopra, Haaland and Roth, 2022, 2023; Chopra et al., 2025; Thaler, 2024).[5] Much of the theoretical literature has emphasized two reasons why individuals consume news aligned with their existing ideological beliefs: they may perceive pro-attitudinal news as more reliable (Gentzkow and Shapiro, 2006), and they may derive greater utility from consuming like-minded content (Mullainathan and Shleifer, 2005). We introduce and study a complementary mechanism for like-minded and low-reliability news consumption that is likely to be especially important in platform-mediated environments: passive portfolio formation.[6]

Second, our paper adds to the growing economic literature on social media (see Aridor, Jiménez-Durán, Levy and Song (2024), Campante, Durante and Tesei (2023), and Zhuravskaya, Petrova and Enikolopov (2020) for reviews), and in particular to the growing body of work exploring interventions aimed at reducing the consumption and sharing of fake, low-reliability, or partisan news. The policy tools proposed so far—including content moderation (Jiménez-Durán, 2025; Beknazar-Yuzbashev, Jiménez-Durán, McCrosky and Stalinski, 2025), algorithmic re-engineering (Kalra, 2026; Stray et al., 2026), fact-checking (Chopra, Haaland and Roth, 2022; Pennycook and Rand, 2021), and accuracy labels (Aslett et al., 2022)—vary substantially in complexity, cost-effectiveness, scalability, and persistence. We introduce a simple and scalable intervention that produces persistent shifts toward less slanted and more reliable news portfolios. The intervention operates as a form of soft commitment: a choice made in a relatively deliberative moment subsequently shapes day-to-day news exposure without requiring repeated attention or ongoing effort. Because analogous platform-mediated choices arise outside social media as well—for example, when users decide which news sites to bookmark, which mailing lists to subscribe to, or which apps to install—a similar approach may be valuable in other digital settings.

Finally, our paper contributes to the literature on experimenter demand effects (de Quidt,

[5] See Cagé, Hengel, Hervé and Urvoy (2025), Caprini (2025), Djourelova, Durante and Martin (2025), Mastrorocco, Ornaghi, Pograxha and Wolton (2025), Mastrorocco, Ornaghi, Pograxha and Wolton (2025) for analyses of news supply.

[6] See Ursu, Simonov and An (2025) for related evidence on consumer passivity in the domain of online advertising.

Haushofer and Roth, 2018; Dhar, Jain and Jayachandran, 2022; Mummolo and Peterson, 2019; Zizzo, 2010) and to the behavioral economics literature on inertia, defaults, and choice architecture (Carroll et al., 2009; Chetty, Looney and Kroft, 2009; Johnson and Goldstein, 2003; Thaler and Sunstein, 2008). On the experimenter demand effects front, we introduce a methodology that leverages the possibility of collecting retrospective data to isolate the effect of experimenter demand. On the behavioral economics front, we demonstrate the critical role that behavioral frictions plays in platform-mediated news consumption and the potential for choice architecture to mitigate them.

1. Background

1.1 News Consumption on Social Media

Social media platforms are central players in the news ecosystem, with 54% of U.S. adults reporting they consumed news on social media within the past week (Newman et al., 2025). Among social media platforms, Facebook remains the dominant source of news in the United States. With approximately 190 million monthly active users domestically, Facebook's sheer scale positions it as a critical gateway to news consumption. Indeed, according to 2025 survey data from Reuters, 32% of U.S. adults regularly obtain news from Facebook—more than from any other social media platform and on par with the most popular U.S. news channel, Fox News (Newman et al., 2025). For comparison, only 23% and 12% of U.S. adults report regularly accessing news through Twitter (now "X") and TikTok, respectively (Newman et al., 2025).

The prevalence of partisan and low-reliability news on social media has generated widespread concerns (Sunstein, 2017). In line with these concerns, Braghieri et al. (2025) document a high degree of polarization in the news consumed on Facebook, showing that news accessed through Facebook is substantially more polarized than news accessed via other channels. For instance, it is 2.5 times more polarized than news accessed through search engines. In parallel, Vosoughi, Roy and Aral (2018) find that fake news spreads both more rapidly and more widely on social media platforms compared to truthful content.

Recent research shows that pro-attitudinal and low-reliability news consumption on social media is in large part driven by users' own feed customization decisions (Levy, 2021; González-Bailón et al., 2023; Chouaki, Chakraborty, Goga and Zannettou, 2024). For instance, Levy (2021) finds that, within a person's feed and according to the isolation index from Gentzkow and Shapiro (2011), content from followed pages is more than twice as ideologically segregated as content from friends.[7] Similarly, Chouaki, Chakraborty, Goga and Zannettou (2024) shows that social media users encounter more misinformation from pages they follow than from incidental exposure. These findings matter both directly—because 41% of the news content that users observe on their feeds and around half of Facebook-referred visits to news sites come from followed pages (Levy, 2021)—and indirectly—because the pages that users follow provide signals to the content ranking algorithm, which might amplify subsequent exposure. Moreover, multiple studies show that the accounts individuals follow on social media affect their knowledge, beliefs, and attitudes (Altay, Hoes and Wojcieszak, 2025; Chmel et al., 2025; Levy, 2021; Rathje et al., 2024).

The evidence above motivates our focus on users' news customization decisions. To connect these decisions to concerns about slanted and low-reliability news content, we need empirical measures of both concepts. The next subsection describes how we operationalize slant and reliability in our setting.

1.2 Slant and Reliability

As far as slant is concerned, we adopt the measure from Bakshy, Messing and Adamic (2015), which assigns ideology scores to hundreds of English-language news outlets based on the average political leaning of Facebook users who shared content from those outlets. This measure satisfies three desiderata: first, it is one of the most comprehensive measures of outlet-level slant available in the literature; second, it correlates strongly with other established measures of slant, helping alleviate concerns about validity (Braghieri et al., 2025); third, it is relatively easy to describe to participants.

Following Aslett et al. (2022), we operationalize the notion of reliability using NewsGuard ratings, which score outlets along nine criteria of journalistic quality, including passing fact-checks, appropriately sourcing material, clearly flagging paid content, distinguishing news from opinion,

[7] Throughout, we use "following a news page" to refer to both "following" and "liking" a page. These actions have similar effects on the content users see in their feeds. The main difference is that following—introduced more recently—offers greater control over notifications and privacy settings.

and having a transparent ownership structure. NewsGuard ratings correlate strongly with several other measures of reliability used in the literature (Lin et al., 2023) and also have high coverage. Appendix B provides additional details on the slant and reliability measures employed in this project.

Our final dataset of outlet slant and reliability comprises 262 news outlets for which we have both a slant score and a reliability score, and that we successfully matched to a corresponding Facebook page. These outlets account for the lion's share of online news consumption: according to Comscore data (Comscore, 2022), they capture 93% of visits to online news websites.[8] Appendix Figure A.1 presents a scatter plot of these 262 outlets in the slant–reliability space. As noted by various media watchdogs (e.g., AdFontes Media), the relationship between slant and reliability is inverse-U-shaped: outlets with extreme slant—particularly those on the right-end of the political spectrum—tend to have lower reliability scores.

2. Motivating Evidence

We begin by documenting new stylized facts about news consumption on Facebook using the baseline survey of our field experiment, which includes 3,353 Facebook users. In the baseline survey, we elicit participants' perceptions of outlet slant and reliability, measure stated bliss points in the slant–reliability space, and observe the news pages participants follow on Facebook. We first describe these elicitation procedures and then present the stylized facts.

2.1 Elicitation of Perceptions and Bliss Points

The baseline survey elicits three objects. First, after explaining our measures of slant and reliability, we asked participants to report their perceptions of the average slant and reliability of the news pages they followed on Facebook. Second, we elicited participants' "bliss points," which we define as the ideal combination of slant and reliability of a person's Facebook news portfolio. Participants reported these bliss points on a two-dimensional grid, with slant on the x-axis and reliability on the

[8] To compute the total number of visits to news websites in the 2022 Comscore Web Behavior Database Panel, we sum visits to all domains that NewsGuard classifies as U.S. English news sites focusing on "Local News," "General News," and/or "Political news or commentary." We manually exclude domains that clearly do not fit these categories (e.g., substack.com).

y-axis. Third, with participants' permission, we used Facebook's API to observe the set of pages they followed on the platform.[9]

Several features of the survey design help ensure that participants understood the elicitation task and had incentives to answer carefully. First, we provided a detailed explanation of the slant and reliability measures and asked participants comprehension questions about each measure. 87% and 67% of participants passed the reliability and slant comprehension questions on their first attempt, respectively. Participants who failed one of the questions were given another opportunity to read the explanations and provide an answer. We automatically screend out 11% of participants who failed either the slant or reliability questions twice; as a result, the analysis sample includes only participants who demonstrated understanding of our slant and reliability measures. Second, to mitigate expressive responding, we incentivized the perceptions elicitation.[10]

The data elicited from participants follow intuitive patterns, which further increases our confidence in the elicitation. For instance, the elicited bliss points recover the distinction between horizontal and vertical attributes that is central to canonical models of news consumption (Gentzkow and Shapiro, 2006; Mullainathan and Shleifer, 2005). In these models, slant is horizontal: more conservative consumers prefer consuming more conservative news. Reliability, in contrast, is vertical: consumers are assumed to generally prefer more reliable news to less reliable news. Our data display this pattern. Appendix Figure A.2a shows that slant bliss points vary systematically with ideology: conservatives report preferring more right-leaning portfolios, while liberals report preferring more left-leaning portfolios. Conversely, Appendix Figure A.2b shows that reliability bliss points have the footprint of a vertical characteristic: the vast majority of both liberals and conservatives report preferring high-reliability portfolios.[11] Besides following intuitive patterns, the elicited bliss points are also behaviorally meaningful. In Section 5.1, we show that, in a later part of the experiment, participants choose to follow outlets that move their portfolios toward the bliss points elicited in an earlier survey. This pattern suggests that the baseline bliss points match

[9] The full baseline survey instrument is provided here.

[10] The measures were incentivized using a binarized scoring rule, where the probability of winning a fixed prize, independently drawn for each participant, increases in the accuracy of the participant's forecasts (Hossain and Okui, 2013). Danz, Vesterlund and Wilson (2022) show that simplifying the instructions for the binarized scoring rule can improve the accuracy of belief elicitation. Following this insight, our survey instructions describe the incentives in non-technical terms and provide a link to the quantitative details for participants who want more information.

[11] Some participants express a preference for reliability below 100. This could reflect demand for entertainment (Chopra et al., 2025)—if respondents perceive a trade-off between reliability and entertainment—misunderstanding, or noise.

the direction in which participants are willing to adjust their actual Facebook news portfolios.

We now turn to the stylized facts. These facts describe how participants' current Facebook news portfolios compare with their stated bliss points, and they provide the empirical motivation for the interventions studied in the rest of the paper.

2.2 New Stylized Facts

Using our measures of perceptions and bliss point, we establish three stylized facts. To our knowledge, these facts have not been empirically documented before and contribute to the literature by providing new descriptive evidence on people's news perceptions, preferences, and portfolios.

Fact 1. *Participants hold news portfolios that, on average, are more partisan and less reliable than their stated bliss points.*

We first compare participants' actual portfolios to their self-reported preferences regarding slant and reliability. Figure 2 plots actual portfolio slant (reliability) against desired slant (reliability). The left panel shows that, although actual and desired slant are positively correlated, participants typically report wanting less extreme portfolios than they currently have: the fitted line is flatter than the 45-degree line. The average gap between actual and desired absolute average slant (the absolute value of the average slant of participants' portfolio) is 0.39, roughly the ideological distance between MSNBC and *The New York Times*.

The right panel shows little relationship between actual and desired reliability: most participants want high reliability regardless of their baseline portfolio, implying that desired reliability typically exceeds actual reliability. The average reliability gap is 15.1 points on the 0–100 scale, roughly the difference between *The Huffington Post* and *The New York Times*.

This fact provides meaningful context for previous results in the literature. Various papers have found that individuals consume like-minded news on social media (e.g., Peterson, Goel and Iyengar, 2021; Braghieri et al., 2025). It is unclear, however, whether these patterns reflect consumers' preferences or whether they are partly driven by other factors at play on social media. Our evidence suggests that, at least to some extent, individuals may be following content that is more ideologically extreme than they would prefer. As individuals have ample flexibility in determining their news portfolio, this raises an important question for both slant and reliability: why do people's choices not match their bliss points? The next two stylized facts suggest that both information frictions and behavioral frictions play a role.

Fact 2. *Information frictions: slant perceptions are relatively accurate; reliability perceptions are relatively inaccurate.*

The left panel of Figure 1 plots perceived portfolio slant (y-axis) against actual portfolio slant (x-axis). The figure shows that participants' slant perceptions are reasonably on target: observations cluster around the 45-degree line, and the correlation between perceived and actual average portfolio slant is 0.61. To assess accuracy categorically, we classify portfolios as left-leaning (slant < -0.2), centrist ($-0.2 \leq$ slant ≤ 0.2), or right-leaning (slant > 0.2). Appendix Table A.1 shows that only 7%-13% of participants substantially misclassify their portfolio's ideological orientation (e.g., perceiving it as right-leaning when it is left-leaning).

Reliability perceptions are markedly noisier and much less accurate than slant perceptions. The right panel of Figure 1 plots perceived portfolio reliability (y-axis) against actual portfolio reliability (x-axis). The correlation is only 0.19, indicating substantial misperceptions.[12] To characterize reliability errors categorically, we group portfolios into low reliability (score < 60), moderate reliability ($60 \leq$ score ≤ 80), and high reliability (score > 80). Appendix Table A.2 shows that a substantial fraction (31%) of participants with low-reliability portfolios incorrectly place them in the high-reliability category, thus suggesting that the provision of targeted information about reliability might help participants improve their news portfolios. Appendix F shows a similar pattern for major outlets: participants are more accurate about slant rankings than about reliability rankings.

Fact 3. *Behavioral frictions: participants describe their news-portfolio formation process as largely passive and driven by suggestions from the platform or from friends.*

Figure 3 shows that 53% of participants report encountering news online incidentally rather than intentionally—that is, while doing something else rather than actively seeking it out. The same figure shows that, when asked how they typically come to follow a news page on Facebook, only

[12] A potential concern is that, despite our explicit instructions and comprehension questions, respondents' interpretation of "reliability" is different from NewsGuard's. Appendix Figure A.3 shows robustness to alternative measures of slant and reliability: we compare perceived reliability to Media Bias Fact Check's factual reporting rating and perceived slant to the slant measure from Braghieri et al. (2025). The same qualitative pattern holds: slant perceptions are substantially more accurate than reliability perceptions.

28% said they had searched for it intentionally; 72% reported following a page because someone shared content from it or because Facebook recommended it. Moreover, few participants appear to think of the pages they follow as a coherent portfolio. When asked whether they consider how a page complements the ones they already follow, only 16% of them answered affirmatively. Finally, to examine whether people pay attention when deciding which outlets to follow, we asked participants if they remembered the last news pages they followed on Facebook, and asked those who said they remembered, to name the page. 70% of participants stated that they do not remember the last page they followed, consistent with passive portfolio formation. Even among those who said they remembered, only 8% entered the correct page.

Taken together, these facts point to the possibility that the curation of news portfolios on Facebook is subject to informational and behavioral frictions. First, users misperceive the reliability of their portfolios, and potentially as a result, hold portfolios that are less reliable than their stated bliss points. Second, despite having relatively accurate beliefs about slant, users' portfolios are also more extreme than their bliss points, potentially due to passivity in news portfolio formation process. The conceptual framework in the next section formalizes the role of frictions, and the experiment described in Section 4.1 tests for these frictions directly and evaluates whether interventions that address them can improve users' news diets on Facebook.

3. Motivating Framework

This section describes the contours of the framework that motivates our experimental design; Appendix C provides a more formal treatment. The goal of the framework is to contrast two broad views of how users arrive at the set of news pages they follow on social media: a view in which users' news portfolio choices on Facebook are highly deliberate and in which they are largely passive and driven by behavioral frictions.

The first view is a *deliberative optimization* benchmark. In that view, users hold reasonably stable preferences over slant and reliability, have a good sense of the menu of outlets available to them, and adjust their portfolios whenever useful information arrives. This benchmark captures the spirit of canonical models of news demand such as Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2006). It naturally implies pro-attitudinal portfolios and allows for low-reliability portfolios when users misperceive reliability. But it also implies that, conditional on their current beliefs and preferences, users' followed-page portfolios should be approximately optimized. As a result, a prompt to reconsider one's portfolio—without any new information—should have little effect. In contrast, useful information about outlet reliability should induce portfolio adjustment even if it is delivered outside a specially prompted choice moment.

The second view is a *passive portfolio formation* model. In that view, users' portfolios are shaped not only by preferences and beliefs, but also by passivity in the decision-making process. Users do not continuously optimize over the full set of possible outlets. Instead, portfolio revision tends to occur only when a decision episode is triggered, and it is then shaped by the subset of outlets that is especially easy to consider and act upon.

This second view fits naturally with the motivating evidence from Section 2. Participants report that they usually come to follow news pages because someone shared content from them or because Facebook recommended them, not because they actively searched for them. They also report wanting portfolios that are less extreme and more reliable than the portfolios they actually hold. Taken together, these patterns suggest that followed-page portfolios may often be inherited, incidental, and only sporadically revised, rather than continuously optimize through deliberate choice.

The distinction between the two views yields a simple set of empirical predictions. Consider first an intervention that prompts users to reconsider the set of news pages they follow and provides an especially easy way to follow or unfollow outlets, but does not provide new information. Under the deliberative optimization benchmark, this intervention should have little effect: if users are already close to or at their optimum, merely prompting reconsideration should not substantially change behavior. Under passive portfolio formation, the same intervention can have a large effect, because it triggers a decision episode, suggests a set of pages that is potentially different from the ones that users naturally encounter on the platforms, and lowers the practical barriers to acting on existing preferences.

Now consider an intervention that provides personalized information about outlet reliability, but does not prompt an active re-optimization moment and does not make portfolio changes especially easy to implement. Under the deliberative optimization benchmark, this intervention should affect behavior to the extent that users value reliability and were previously misinformed about it. Under passive portfolio formation, the effect may be small or zero: users may update beliefs without changing behavior if the information arrives outside a moment in which portfolio revision is naturally triggered.

These comparative statics map directly into our experimental design. The **R** treatment isolates the effect of a prompted re-optimization moment without new information. The **NR** treatment isolates the effect of reliability information without a re-optimization prompt. This structure allows us to test whether users' followed-news portfolios are approximately optimized given stable preferences and current beliefs, so that information alone should move behavior while a prompt without information should not. If, instead, we observe the opposite pattern—large responses to **R**, little response to **NR**—that would be difficult to reconcile with deliberative optimization and more consistent with a passive, frictional account of portfolio formation.

4. Experimental Design and Implementation

4.1 Experiment Overview

Figure 4 presents a schematic overview of the experimental design, which we summarize below.[13]

We recruited participants for a five-week experiment via Facebook ads. To avoid attracting a sample with specific views about news consumption on social media, the ads did not disclose the nature of the study. Upon clicking an ad, users were shown a consent form explaining that participation required logging into the study via Facebook and granting the research team permission to access the list of pages they follow on the platform. We collected the baseline news portfolio at the beginning of the survey and personalized some of the questions accordingly. Participants who accessed the baseline survey using the Google Chrome browser were offered a bonus to install a Chrome extension that tracked their visits to a large set of news websites, allowing us to observe off-platform online news consumption.

Participants who consented and logged in via Facebook were directed to a baseline survey. The survey collected demographic data, information about news consumption habits, and pre-treatment measures of various outcome variables.[14] The baseline survey also included detailed

[13] Screenshots of the experimental interface are available at the following links: baseline, midline, and endline.

[14] We excluded participants who met any of the following criteria: were not U.S. residents, accessed the survey

explanations of our slant and reliability measures, comprehension questions ensuring participants understood these measures, incentivized questions to elicit participants' perceptions, and questions on participants' bliss points, as discussed in Section 2.1.

Around one week after the baseline survey, participants were invited to complete a midline survey. During that survey, they were randomized into one of six study arms: a *Control* group, three *Re-optimization* treatment arms, a *Reliability Information* treatment (NR), and a *No Experimenter Demand* treatment (NED). All study arms other than the NED treatment are described in this section; the NED treatment—employed to check for potential experimenter demand effects—is described in Section 6. Randomization was stratified based on the average slant and reliability of participants' baseline news portfolios and based on whether they installed the Chrome extension.[15]

Participants in the control group received no intervention. We implemented three re-optimization treatments: Re-optimization (R), Re-optimization with Reliability Information (RR), and Re-optimization with Slant Information (RS). Participants assigned to any of the re-optimization treatment arms were shown a simple, user-friendly interface that allowed them to follow or unfollow news pages on Facebook with a single click. Specifically, participants in the re-optimization treatments were shown a list of news outlets and were informed that they could update their Facebook news portfolio by clicking the "follow" or "unfollow" button next to each outlet. The interface was integrated with the Facebook API, so that any changes made through the survey were immediately reflected in the participants' Facebook accounts.

The re-optimization interface was divided into two tables. The first table showed each participant 12 outlets they did not already follow and asked them whether they wanted to follow any of those outlets going forward. The table always included a mix of outlets that spanned the slant spectrum and had either medium or high reliability.[16] The second table showed each participant

via a VPN, did not follow any news pages on Facebook, reported never obtaining news from Facebook, did not grant us permission to observe the pages they follow on the platform or whose Facebook data we could not access (e.g., because the participant had a professional profile), failed comprehension checks twice regarding our definitions of slant and reliability, were flagged as potential scammers or duplicates, and requested data deletion. Appendix Figure A.4 provides more details on participants who were screened out before randomization.

[15] As specified in the pre-analysis plan, participants who installed the Chrome extension were randomized only into the control group or one of the three re-optimization treatments. The reason is that these participants were more expensive to recruit; thus, to maximize power, we assigned them only to the treatment arms that we expected to have an effect.

[16] We omitted low-reliability pages from the table because we did not want to spread misinformation. One might worry that doing so mechanically boosts reliability—for instance, because "random" choices from the re-optimization interface would then be drawn from a more reliable pool. Appendix Figure A.6 shows this is not the case; if anything,

the news outlets they were already following and provided an option to unfollow any of them going forward.[17] We did not provide an incentive to follow or unfollow pages, nor did we encourage users to make changes to their portfolios. Appendix D provides more information about how we selected the news pages displayed in the re-optimization interface.[18]

We randomized the provision of additional information across the treatment arms. Participants assigned to the Re-optimization (R) treatment received no additional information. In the Re-optimization Reliability Info (RR) treatment, participants received information about the reliability of each news outlet they followed on Facebook, the average reliability of their current news portfolio on the platform, and the reliability of each outlet shown in the re-optimization interface. Participants in the Re-optimization Slant Info (RS) treatment received analogous information regarding the slant of news outlets.

Finally, we deployed the Reliability Info (NR) treatment in order to study the consequences of mere information provision. Participants in the NR treatment were not shown the re-optimization interface, but were given the same reliability information as participants in the RR treatment—both for the outlets they followed and for those they would have seen in the re-optimization interface had they gone through it.

Approximately four weeks after the midline survey, participants were invited to complete an endline survey. This final survey elicited secondary outcomes, including participants' satisfaction with their Facebook news feed and trust in the news they see on Facebook. At the end of the endline survey, participants were given another opportunity to re-optimize their news portfolios using the same interface and the same set of news pages. This allowed us to assess whether participants remained satisfied with the portfolio choices they made earlier in the experiment.

4.2 Pre-analysis Plan and Procedures

Pre-analysis Plan. We pre-registered the study by submitting a pre-analysis plan (Braghieri, Levy and Trachtman, 2025) in which we specified the sample size, experimental design, outcome variables—including a distinction between primary and secondary outcomes, as well as the con-

random selection from the interface reduces portfolio reliability on average.

[17]Participants were shown up to 12 news pages they already followed. If a participant's Facebook news portfolio contained more than 12 pages, we randomly selected 12 pages to display.

[18]The midline survey with screenshots of the interface is available here.

struction of each outcome variable—and our main regression specifications. The only notable deviation from the pre-analysis plan is that we did not manage to recruit as large a sample as we had anticipated, due to higher-than-expected recruitment costs. Thus, the sample size in the experiment is approximately 70% of the one specified in the pre-analysis plan.

Procedures. We implemented the experiment in three separate but overlapping waves, beginning on January 14, 2025, and concluding in April 2025. Participants received \$5 for completing the baseline survey and \$8 for each of the midline and endline surveys. Those who installed the Chrome extension and kept it active for the full duration of the experiment received an additional \$20. Incentive payments for accurately answering the perception questions were disbursed at the end of the study. We also provided additional incentives for specific survey invitations to decrease attrition. Appendix E contains additional information about the implementation of the experiment.

4.3 Outcome Variables

Following our pre-analysis plan, we distinguish between primary and secondary outcome variables.

Primary Outcome Variables. We pre-specified the following as our primary outcomes:

- **Perceptions of slant and reliability.** Participants provided incentivized assessments of the average slant and reliability of the portfolio of news pages they follow on Facebook, as well as of various news outlets.
- **News-following behavior on Facebook.** Using data obtained directly from the Facebook API, we track the set of news pages each participant follows on Facebook at various points throughout the experiment. From this list, we compute the average reliability and the absolute average slant of each participant's news portfolio.
- **Distance to bliss points in slant–reliability space.** Participants reported their ideal combination of slant and reliability in a two-dimensional space. We calculate the distance between each participant's bliss point and the actual slant and reliability of her news portfolio. We construct both dimension-specific distances and a composite Euclidean distance after standardizing our measures of slant and reliability.

- **Off-platform news consumption.** For participants who installed the Chrome extension, we record the number of visits to a large set of news websites.
- **Self-reported news exposure.** We record participants' reports of the average slant and reliability of the news content they encountered on Facebook between the midline and endline surveys.

Secondary Outcome Variables. Our pre-specified secondary outcomes consist of survey-based measures of trust in news, satisfaction with one's news diet, and overall news consumption habits on Facebook.

We did not pre-specify downstream outcomes related to political attitudes or knowledge for two reasons. First, a core objective of this paper is to understand the drivers of news customization decisions in platform-mediated environments, where followed-page portfolios may reflect passive portfolio formation and frictions to active re-optimization. Doing so requires a multi-arm design that varies both choice architecture and information provision; as a result, the study was neither designed nor powered to detect changes in downstream political outcomes. Second, a large literature has already documented that media exposure (including social media) can affect downstream political outcomes (e.g. Adena et al., 2015; Allcott, Braghieri, Eichmeyer and Gentzkow, 2020; Altay, Hoes and Wojcieszak, 2025; Broockman and Kalla, 2025; Bursztyn, Rao, Roth and Yanagizawa-Drott, 2023; DellaVigna and Kaplan, 2007; Chen and Yang, 2019; Djourelova, 2023; Enikolopov, Petrova and Zhuravskaya, 2011; Fujiwara, Müller and Schwarz, 2024; Levy, 2021; Martin and Yurukoglu, 2017) and, as we show in Section 5.4, our intervention affects media exposure and consumption.

4.4 Summary Statistics

Table 1 summarizes the number of participants at each stage of the experiment. We also report the number of participants, at each stage, for whom we successfully collect the set of pages they follow on Facebook from the platform's API.[19]

[19] We do not observe the set of followed pages on Facebook for all participants at all points in the experiment because participants can revoke these permissions and because, in a small fraction of cases, our calls to the Facebook API fail. We provide more details in Appendix E.

Appendix Table A.5 compares the demographic characteristics of our sample to those of the broader Facebook and U.S. populations. Relative to the average Facebook user and the average U.S. adult, our sample is more likely to be female, white, older than fifty, and college-educated. Our sample is also more likely to identify as Democrat compared to the overall U.S. adult population.[20] Appendix Table A.6 shows randomization achieved balance along the expected number of baseline observable characteristics.

Appendix Table A.7 reports attrition rates across treatment arms. Attrition between midline, where randomization occurs, and endline was minor—approximately 15%. Only one out of our four main treatment arms, namely the R treatment, exhibits evidence of differential attrition at the endline survey. Such differential attrition is not a concern for our primary outcome, namely the set of pages participants follow on Facebook right after the midline survey, because that outcome is not subject to survey attrition.[21]

4.5 Main Empirical Strategy

To evaluate the effects of our interventions, we estimate the impact of our treatments on each outcome variable from Section 4.3. We estimate the following OLS regression:

$$Y_i = \alpha + \sum_{j \in J} \beta_j T_i^j + \rho \cdot Y_i^b + \eta \cdot \mathbf{X_i} + \mu_s + \varepsilon_i \tag{1}$$

where Y_i is an outcome variable, Y_i^b is the baseline value of Y_i (when available), $T_i^j \in \{0,1\}$ is an indicator for assignment to treatment arm $j \in J = \{R, RS, RR, NR\}$ (with the Control group as the reference group), μ_s denote recruitment and randomization strata fixed effects,[22] ε_i is an idiosyncratic error term, and $\mathbf{X_i}$ is a vector of pre-specified baseline covariates: political ideology, reliance on Facebook for news, interest in politics, recruitment wave fixed effects, and device fixed

[20] In Appendix Tables A.9-A.10 we show that our results are robust to weighting the sample to mimic the US population on age, gender, education, race, ethnicity, and party affiliation. There could, of course, remain unobservable differences between our experimental sample and the overall Facebook population.

[21] In Section 5.2 we show that the effects on our primary outcomes collected in the midline survey, where no attrition occurs, are remarkably similar to the primary outcomes collected in the endline survey, further allaying concerns about attrition at endline. Moreover, Section 5.2 also shows that all the re-optimization treatments had a similar effect on slant, and Appendix Table A.7 does not find evidence for differential attrition when pooling these treatments. Finally, Appendix Table A.11 shows that balance on observable characteristics is largely preserved at endline.

[22] As described in Section 4.1, these include average slant and reliability of baseline news portfolios and an indicator for installing the browser extension.

effects (desktop vs. mobile). When estimating Equation 1, we use robust standard errors.

5. Results

We present our experimental results in four subsections. First, we use the differential response to the **R** and **NR** treatments to assess the deliberative-optimization benchmark from Section 3 and to evaluate whether the data are more consistent with passive portfolio formation under behavioral frictions. Second, we examine whether the intervention reduces portfolio slant and increases average reliability. Third, we show that interventions reduce the gap between participants' Facebook news portfolios and their stated bliss points, thereby mitigating internalities. Fourth, leveraging participants who installed our Chrome extension, we document that changes in users' news portfolios affected downstream news consumption.

5.1 News Portfolio Choices and Evidence of Behavioral Frictions

Re-optimization (R) Treatment Under the narrow deliberative-optimization benchmark from Section 3, our re-optimization treatment—which prompts active reconsideration and makes portfolio adjustment especially easy—should have little effect. The intuition is straightforward: if users' followed-news portfolios are approximately optimal given stable preferences and current beliefs, then a prompt to reconsider one's portfolio, without new information, should not substantially change behavior.[23,24] As shown in the first panel of Appendix Figure A.5, the vast majority (71%) of participants assigned to the R treatment adjusted their portfolios by following or unfollowing at least one news page, despite receiving no new information.

If portfolios are so sensitive to the presence of the re-optimization interface, we might also expect them to be sensitive to the particular set of pages that the interface explicitly puts in front

[23]Following a page typically requires typing its name into the Facebook search bar and clicking "Follow." Unfollowing is even easier: users can click "Unfollow" directly from any post that appears in their feed.

[24]Strictly speaking, our experiment identifies the joint effect of behavioral frictions and the minimal transaction cost of following or unfollowing a news page. We interpret the results as evidence of behavioral frictions because the implied transaction costs are extremely small. We estimate that following a page takes around 3 seconds. Participants were willing to take a survey for a compensation equivalent to a $30/hour wage. Thus, an implied upper bound on their time cost of following a news page outside of the re-optimization interface is about $0.025. As shown in the next few sections, following a news page affects both what users see on Facebook and the news sites they visit over time. In order for transaction costs alone to explain our re-optimization results, the total value of these effects to participants would have to be less than 2.5 cents.

of users. Indeed, in the four weeks after the midline survey, only 6.3% of participants in the R treatment add to their portfolios news outlets that were not included in the re-optimization interface. In other words, nearly all participants in the R group who make changes to their portfolios do so from within the set of pages explicitly offered by the interface.

That said, participants' choices within the interface are systematic rather than random. To test this, we simulate counterfactual portfolios under random choice.[25] Appendix Figure A.6 panel (a) reports simulated effects on absolute average slant and panel (b) reports simulated effects on reliability. The gray bars show the distribution of average treatment effects obtained under random choice. In both panels, the actual average treatment effect of the R treatment lies far in the tails of this distribution, leading us to reject the hypothesis that participants simply choose randomly from the re-optimization interface. Panel (a) also reports a distribution generated by simulating random choice from among pro-attitudinal outlets; we reject this version of random choice as well.[26] Consistent with this interpretation, we show in Section 5.3 that re-optimization treatments reduce the distance between participants' portfolios and their stated bliss points, indicating that participants use the interface to move portfolios in directions consistent with their stated preferences.

No-Reoptimization Reliability Information (NR) Treatment Under the deliberative-optimization benchmark, providing participants with outlet- and portfolio-level information they value—such as reliability information—should induce re-optimization even without an explicit prompt. By the end of the experiment, participants assigned to the NR treatment added, on average, only 0.09 news pages to their Facebook news portfolios, compared to 1.90 in the R treatment. Put differently, a prompt that triggers re-optimization without providing any new information induces portfolio changes roughly twenty times larger than an intervention that provides targeted reliability information without a re-optimization interface. Consistent with this, and as shown in Section 5.2, the NR treatment does not increase the average reliability of participants' Facebook news portfolios relative to the control group.

A natural question is whether NR is ineffective because participants do not value the infor-

[25] For each participant in the R treatment, we randomly draw from the offered choice set the same number of outlets the participant actually followed via the interface. We then compute the resulting portfolio slant and reliability, estimate Equation 1, and repeat the procedure one thousand times.

[26] The fact that the R treatment differs from the pro-attitudinal random-choice benchmark also illustrates that the balanced menu in the re-optimization interface does not mechanically force moderation.

mation, or because they value it but fail to act on it in the absence of a prompt. We can rule out the former interpretation for two reasons. First, as documented in Section 2—and as assumed in canonical models of news demand (Gentzkow and Shapiro, 2006)—participants substantially misperceive outlet reliability, so information about reliability is plausibly useful. Second, and more importantly, when the same reliability information is paired with the re-optimization interface in the RR treatment, it meaningfully raises average portfolio reliability relative to the control group as shown in Section 5.2. Together, these findings indicate that information alone is not sufficient to change behavior: it becomes behaviorally relevant only when an active choice is prompted.

Overall, the pattern of results from the R and NR treatments is difficult to reconcile with a benchmark in which users' followed-news portfolios are approximately optimized given stable preferences and current beliefs, and in which useful information is acted upon even outside a prompted re-optimization moment. It is instead more consistent with passive portfolio formation under behavioral frictions.

5.2 Externalities

We next ask whether our interventions move participants' Facebook news portfolios along the two dimensions that are central to concerns about news consumption on social media producing negative externalities on democracy (Sunstein, 2017): slant and reliability.

Slant Figure 5 presents treatment effects on the absolute average slant of participants' Facebook news portfolios. (Appendix Table A.8 reports the full regression results along with comparisons between treatment arms).[27] Relative to the control group, all three re-optimization treatments significantly reduce portfolio slant. The effects appear immediately after the midline survey—when participants complete the re-optimization task—and persist through the endline survey roughly four weeks later (columns (1) and (2), respectively).[28] In contrast, the NR treatment has no detectable effect. Because the post-midline effect sizes are very similar across the three re-optimization arms,

[27] The treatment effect estimates in this section are obtained from Equation 1 and weight each observation equally. Appendix Tables A.9-A.10 show similar results when estimating population average treatment effects using weights that align the sample with the U.S. population on age, gender, race, education, ethnicity, and party affiliation. It is unsurprising that the weighted results are similar, as Appendix Figures A.7-A.8 find little heterogeneity based on age, gender, education, and party affiliation.

[28] Appendix Table A.12 runs the same analysis among participants from whom we have both midline and endline data, and confirms that the effects we find persist when we fix the sample.

we streamline the exposition by pooling them in what follows; the pooled regression appears in column (1) of Appendix Table A.13.

The magnitude of the slant effects is substantial. The re-optimization treatments decreased absolute average slant by 0.24 standard deviation units—equivalent to 51% of the baseline gap between self-identified independents and partisans in our sample. As an alternative benchmark, the estimate corresponds to more than half the distance between the slant of The Washington Post (a liberal outlet) and that of Bloomberg (a centrist outlet). Appendix Table A.14 shows that the effect applies to both Democrats and Republicans, as both groups reduce portfolio slant in response to our intervention. Appendix Figure A.7 expands the analysis and shows that the re-optimization treatment reduced absolute average slant across almost all demographic groups we studied. Overall, we do not find evidence for substantial heterogeneity across education, gender, age, and party affiliation. The consistent effects suggest that the experiment has high external validity and would have similar results in other populations, but of course, the effect may still differ along unobserved dimensions.

A reduction in absolute average slant could arise through two channels: (i) greater exposure to centrist outlets (low slant in absolute value), or (ii) greater exposure to counter-attitudinal outlets (outlets whose ideological orientation differs from the participant's own). Appendix Table A.13 points to the second channel. Column (2) shows no statistically significant decline in average absolute slant (as opposed to absolute average slant), suggesting that increased emphasis on centrist outlets is not the primary driver. Column (3), in contrast, shows that the share of counter-attitudinal pages in a participant's portfolio more than triples under the pooled re-optimization treatments. Therefore, treated participants do not simply shift toward more neutral content; they diversify their news portfolios by adding outlets from the other side of the ideological spectrum. Columns (3)–(6) of Appendix Table A.14 show that the diversification effects apply to both Democrats and Republicans.

Figure 6 provides a complementary view by focusing on participants in the pooled re-optimization treatments and splitting them by baseline portfolio type: left-leaning (average slant < -0.2), moderate ($-0.2 \leq$ slant ≤ 0.2), and right-leaning (average slant > 0.2). For each group, we plot desired portfolio slant, perceived portfolio slant, actual baseline slant, and post-treatment slant. Three patterns stand out. First, participants with partisan portfolios perceive them as more moderate than they are, indicating some mild slant misperceptions. Second, these participants nonetheless report wanting portfolios that are even more moderate than what they perceive, consistent with the presence of behavioral frictions. Third, the re-optimization interface moves partisan portfolios toward these more moderate bliss points, in line with the results shown in Section 5.3.

Reliability Figure 7 presents the treatment effects on the average reliability of participants' Facebook news portfolios, and Appendix Table A.15 reports the full regression results. The Re-optimization with Reliability Information (RR) treatment produces a large and statistically significant increase in portfolio reliability. As with slant, the effect emerges immediately at midline—when re-optimization occurs—and persists through endline (Appendix Table A.16 shows that the effects also persist when analyzing the same sample in both periods). In contrast, the Re-optimization (R) and Re-optimization with Slant Information (RS) treatments generate much smaller, statistically insignificant changes in reliability.

The comparison between RR and NR further underscores the importance of pairing information with an active-choice moment. Participants in the NR arm receive the same reliability information as those in RR, but without the re-optimization interface. By the end of the experiment, RR participants added, on average, 2.00 pages to their Facebook news portfolios, whereas NR participants added only 0.09. Consistent with this limited re-optimization, the last row of Figure 7 shows that the NR treatment does not significantly increase portfolio reliability relative to the control group.[29]

The magnitude of the RR effect is sizable. Relative to the control group, RR increases average portfolio reliability by 0.21 standard deviation units. This corresponds to closing 85% of the baseline reliability gap between college-educated and non-college-educated participants. As an additional benchmark, the average improvement is roughly comparable to adding The New York Times to a portfolio consisting of The Huffington Post, Jezebel, and Slate. Appendix Figure A.8 shows that the RR treatment increases reliability across demographic groups. Again, we do not find substantial heterogeneity, with the exception of a stronger effect among women.

[29]These findings echo Aslett et al. (2022), who study a browser-extension intervention that displayed NewsGuard ratings and find no discernible effect on news consumption. Our results reinforce this conclusion—information alone has limited power to shift behavior—and also suggest a solution: coupling the provision of information with a decision prompt.

Relation between Treatment Effects and Misperceptions The pattern of treatment effects mirrors the asymmetry in participants' baseline beliefs. As shown in Section 2, participants perceive outlet slant relatively accurately but substantially misperceive reliability. Consistent with this, re-optimization alone is sufficient to reduce portfolio slant, whereas improving reliability requires coupling re-optimization with targeted reliability information.

In Appendix Table A.18, we use our main specification to analyze the effects of the intervention on participants' perceptions of their Facebook feeds: perceived slant (column 1) and perceived reliability (column 2). The RR treatment increases perceived feed reliability, consistent both with belief updating among participants who initially underestimate reliability and with the observed shift toward higher-reliability portfolios under RR. In contrast, we find no significant effects on perceived slant. This may reflect limited sensitivity of our survey measure: the slant question is relatively coarse and may not capture changes in absolute average slant even when portfolio composition shifts in that direction.

5.3 Internalities

We have shown that our re-optimization interface induces the majority of participants to make adjustments that result in more balanced and, when paired with information, more reliable portfolios. We have also documented that, at baseline, there is a wedge between users' preferences over the slant of their portfolios and the actual slant of their portfolios. Do the portfolio changes induced by the interventions also help align participants' news portfolios with their stated preferences?

To quantify alignment, we construct distance measures between each participant's portfolio and her bliss point in the slant–reliability space.[30] Along each dimension, we compute the absolute gap between the average slant (reliability) of the news pages the participant follows on Facebook and her corresponding slant (reliability) bliss point. We also construct a composite measure: the Euclidean distance between the portfolio and the bliss point in two dimensions, after standardizing slant and reliability relative to the control-group distribution.

We find evidence that, indeed, our interventions reduce the gap between participants' Facebook news portfolios and their stated bliss points. Table 2 reports treatment effects using our main specification (Equation 1), with the slant distance, reliability distance, and Euclidean distance

[30] We elicited bliss-points in the baseline survey before the intervention.

as outcomes. All three re-optimization treatments significantly reduce distance to the slant bliss point. In addition, the R and RR treatments significantly reduce distance on the reliability dimension. Together, these results suggest that our intervention—by addressing behavioral frictions and, when applicable, providing targeted reliability information—helps participants bring their Facebook news portfolios closer to the portfolios they themselves report wanting.

To benchmark these results, we compare the change in participants' Facebook portfolios with how close participants could have moved toward their bliss point. We first assume that the number of followed or unfollowed outlets remains the same for each participant and check which outlets participants should have followed or unfollowed to get as close to their bliss point as possible. We then compare the change in the portfolio to the optimal change considering the choice set. We find that the effect of the re-optimization treatment (R) is 56.2% of the optimal adjustment in slant and the effect of the re-optimization and reliability information treatment is 55.5% of the optimal adjustment in reliability.

Appendix Table A.17 examines two pre-specified survey outcomes that also pertain to participant welfare: satisfaction with the news participants see on Facebook (column 1) and trust in that news (column 2). We find that the RR treatment significantly increases satisfaction with the news participants encounter on Facebook and so do the pooled re-optimization treatments, albeit with a p-value that straddles significance at the 5% level (p-value=0.064). Column (2) shows that the pooled re-optimization treatments significantly increase trust in the news participants see on Facebook.

5.4 Downstream News Consumption

We have shown that our intervention induces meaningful changes in the set of news pages participants follow on Facebook. We now ask whether these portfolio changes translate into changes in actual news consumption. To study whether following an outlet's Facebook page increases consumption of content from that outlet, we follow our pre-analysis plan and implement an instrumental variables (IV) strategy, using treatment assignment as an instrument for following behavior.

IV Specification We consider the following regression equation, where each observation is a participant-page:

$$Y_{i,k} = \phi + \tau \cdot F_{i,k} + \psi \cdot Y^b_{i,k} + \chi \cdot \mathbf{X_i} + \mu_s + \upsilon_{i,k} \tag{2}$$

where $Y_{i,k}$ is an outcome that relates to participant i and outlet k, $Y^b_{i,k}$ is the baseline value of $Y_{i,k}$ (when available), and $F_{i,k}$ is an indicator for whether participant i follows outlet k on Facebook. We instrument for $F_{i,k}$ using an indicator that takes value one if the participant was assigned to any of the three re-optimization treatments (R, RR, RS), and zero otherwise. We include the same set of controls $\mathbf{X_i}$, and μ_s, as in our main specification and cluster standard errors at the participant level.

To define the set of outlet–participant pairs (i,k) used in the IV specification, we construct, for each participant who completed the midline survey—including those in the control group—the set of outlets they would have been shown had they been assigned to one of the three re-optimization treatments. We refer to these as "potentially offered outlets." Importantly, these outlets exclude outlets that the participant already followed at baseline.[31]

IV Results As a first check, Appendix Table A.19 shows that participants who follow a news page as a result of treatment report seeing more posts from that page in their Facebook feeds at endline. This shows that following behavior has noticeable consequences for on-platform exposure.

To examine off-platform news consumption, we use browsing data from the subset of participants who installed our Chrome extension. The extension tracks browsing only on the device where it is installed (a desktop or laptop computer). Because much online news consumption occurs on mobile devices, the estimates in this section should therefore be interpreted as lower bounds on total changes in consumption.

Table 3 shows that following an outlet's Facebook page as a result of assignment to one of our re-optimization treatments increases visits to that outlet's website in the month after the midline survey. We estimate the causal effect of following on visits using the instrumental-variables specification in Equation 2, pooling participants assigned to the three re-optimization treatments. Column (2) shows that following a page increases the probability of a Facebook-originating visit

[31] One might be worried about possible violations of monotonicity and of the exclusion restriction. To assuage those concerns, we also present reduced-form effects.

by 4.02 percentage points, and column (4) shows that it more than doubles the overall probability of a visit (an increase of 10.55 percentage points). The effects operate on both the extensive margin and the intensive margin: total visits also increase, particularly visits originating from Facebook (columns (3) and (5)).[32]

Table A.20 reports reduced form effects of the pooled re-optimization treatments on visits. Unlike the IV specification, these estimates rely only on random assignment and do not require the monotonicity and exclusion-restriction assumptions. Taken together, the results from this section provide direct evidence that changes in Facebook following behavior generate meaningful shifts in off-platform news consumption.

6. Mechanisms and Robustness

6.1 Experimenter Demand

Because our interventions—especially RR—are relatively transparent, a natural concern is experimenter demand: participants may adjust their behavior to align with what they believe the research team wants to find. Our downstream consumption results make this explanation prima facie less compelling. In particular, it seems unlikely that participants would incur meaningful time costs over the course of a month to visit specific news sites simply to conform to perceived experimental aims.

To directly test for experimenter demand effects, we implemented an additional arm that closely mirrors the Re-optimization with Reliability Information (RR) treatment but removes perceived and actual experimenter oversight. In the No Experimenter Demand (NED) treatment, participants were explicitly told during the midline survey that we would no longer observe the pages they followed on Facebook. To reinforce this, we revoked the Facebook permissions participants had granted us at baseline. Participants could remove the permissions themselves using instructions we provided or authorize us to revoke them automatically. Participants' knowledge that we

[32] As pre-specified, we also examine the slant and reliability of the websites participants visited in the month following the intervention. Since the subset of participants for whom we observe browsing data is small, we have limited statistical power. Table A.21 presents these results and shows suggestive evidence that the re-optimization treatments increased visits to counter-attitudinal outlets; in light of the limited power, however, we refrain from drawing strong conclusions.

no longer had access to their page-following data from that point forward substantially reduces the scope for demand-driven behavior during the midline survey.

At endline, we surprised participants by re-requesting permission to observe their followed pages. Nearly all participants (99%) who took the endline survey consented. This design allows us to compare endline portfolios between the NED and RR groups and thereby gauge how much of the RR effect could plausibly be attributed to experimenter demand.

Appendix Table A.22 reports the results. Like RR, the NED treatment significantly increases the average reliability of participants' Facebook news portfolios. The point estimate under NED is somewhat smaller than under RR, but the difference is not statistically significant at conventional levels. Thus, we can be confident that our results are not primarily driven by experimenter demand effects.

6.2 Experimentation Motives

Throughout the paper, we interpret the high rate of re-optimization in the R treatment as evidence that behavioral frictions shape participants' news-following decisions on Facebook. A natural alternative explanation is experimentation: participants may have temporarily followed outlets out of curiosity, rather than to correct misalignment between their portfolios and their preferences.

Several pieces of evidence weigh against the experimentation hypothesis. First, a substantial share of participants (20%) unfollowed at least one news page during the midline survey. This behavior is difficult to reconcile with experimentation, which typically entails sampling new options rather than disengaging from familiar ones. Second, among participants who followed at least one new outlet at midline, 86% reported being familiar with at least one of the outlets they chose to follow, suggesting that many additions reflect deliberate choices among familiar options rather than exploratory clicks. Third, experimentation could in principle improve portfolio alignment with stated bliss points in the long run if participants "try out" pages and then keep only the ones they like. However, this mechanism does not predict the sharp, immediate reduction in distance to bliss points we observe right after midline, when portfolio changes occur (Section 5.3). Fourth, if midline changes were primarily trial runs, we would expect substantial churn by endline. In fact, we make it easy for participants in the endline survey to unfollow pages they started following at midline by offering them another opportunity to revise their Facebook news portfolios using the same interface and the same set of pages. Instead of substantial churn, we find that only 5% of pages followed at midline are unfollowed by endline. This low reversal rate is hard to square with a transient experimentation story.

Taken together, these patterns suggest that experimentation is unlikely to be the main driver of re-optimization in the R treatment and support our interpretation that the intervention operates primarily by reducing behavioral frictions.

7. Conclusion

In this paper, we showed that users' followed-news portfolios on social media are not simply the product of stable preferences and information. They are also shaped by passive portfolio formation and information frictions. A friction-reducing re-optimization intervention moves users' Facebook news portfolios closer to their stated bliss points in the slant–reliability space, reduces portfolio slant, and, when paired with reliability information, increases portfolio reliability.

These results speak directly to questions of external validity and policy relevance. The intervention is both cheap and scalable: it does not rely on algorithmic overhauls, regulatory enforcement, or heavy-handed content moderation. Instead, it mirrors the kind of nudges already deployed on platforms (e.g., privacy check-ups or account security prompts) and could be implemented widely with minimal cost. If platforms are unwilling to implement such interventions, NGOs can implement them instead, or regulators can promote them. The fact that our intervention produces measurable behavior change in a real-world setting offers strong grounds for optimism about external validity.

More broadly, our results highlight a general lesson about platform-mediated environments: users' followed-page portfolios are shaped not only by preferences and the options available, but also by when platforms prompt active reconsideration and by whether portfolio adjustment is made especially easy. Consistent with this mechanism, information that users value may have little effect when delivered on its own, whereas small changes to the choice architecture can produce large and persistent shifts in behavior.

References

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of the Nazis in Prewar Germany." *The Quarterly Journal of Economics*, 130(4): 1885–1939.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2020. "The Welfare Effects of Social Media." *American Economic Review*, 110(3): 629–676.

Allen, Jennifer, Antonio A Arechar, Gordon Pennycook, and David G Rand. 2021. "Scaling Up Fact-Checking Using the Wisdom of Crowds." *Science Advances*, 7(36): eabf4393.

Altay, Sacha, Emma Hoes, and Magdalena Wojcieszak. 2025. "Following news on social media boosts knowledge, belief accuracy and trust." *Nature Human Behaviour*.

American National Election Study. 2024. "American National Election Study."

Angelucci, Charles, and Andrea Prat. 2024. "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News." *American Economic Review*, 114(4).

Aridor, Guy, Rafael Jiménez-Durán, Ro'ee Levy, and Lena Song. 2024. "The Economics of Social Media." *Journal of Economic Literature*, 62(4): 1422–1474.

Aslett, Kevin, Andrew M Guess, Richard Bonneau, Jonathan Nagler, and Joshua A Tucker. 2022. "News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions." *Science Advances*, 8(18).

Bakshy, Eytan, Solomon Messing, and Lada A Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science*, 348(6239): 1130–1132.

Beknazar-Yuzbashev, George, Rafael Jiménez-Durán, Jesse McCrosky, and Mateusz Stalinski. 2025. "Toxic Content and User Engagement on Social Media: Evidence from a Field Experiment."

Bhadani, Saumya, Shun Yamaya, Alessandro Flammini, Filippo Menczer, Giovanni Luca Ciampaglia, and Brendan Nyhan. 2022. "Political Audience Diversity and News Reliability in Algorithmic Ranking." *Nature Human Behaviour*, 6(4): 495–505.

Braghieri, Luca, Ro'ee Levy, and Hannah Trachtman. 2025. "Demand for Online News, Inertia, and Misperceptions." AEA RCT Registry. https://doi.org/10.1257/rct.14392.

Braghieri, Luca, Sarah Eichmeyer, Ro'ee Levy, Markus Mobius, Jacob Steinhardt, and Ruiqi Zhong. 2025. "Article-level Slant and Polarization of News Consumption on Social Media." *Working Paper*.

Broockman, David E, and Joshua L Kalla. 2025. "Consuming cross-cutting media causes learning and moderates attitudes: A field experiment with Fox News viewers." *The Journal of Politics*, 87(1): 246–261.

Bursztyn, Leonardo, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott. 2023. "Opinions as facts." *The Review of Economic Studies*, 90(4): 1832–1864.

Cagé, Julia, Moritz Hengel, Nicolas Hervé, and Camille Urvoy. 2025. "Political Bias in the Media: Evidence from the Universe of French Broadcasts, 2002–2020."

Campante, Filipe R., Ruben Durante, and Andrea Tesei. 2023. *The Political Economy of Social Media.*

Caprini, Giulia. 2025. "Visual Bias."

Carroll, Gabriel D, James J Choi, David Laibson, Brigitte C Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *The Quarterly Journal of Economics*, 124(4): 1639–1674.

Chen, Yuyu, and David Y Yang. 2019. "The impact of media censorship: 1984 or brave new world?" *American economic review*, 109(6): 2294–2332.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99(4): 1145–77.

Chmel, Kirill, Eunji Kim, John Marshall, Tiffany Fisher-Love, and Nathaniel Lubin. 2025. "How social media creators shape mass politics: A field experiment during the 2024 US elections." *arXiv preprint arXiv:2512.15401*.

Chopra, Felix, Ingar Haaland, and Christopher Roth. 2022. "Do People Demand Fact-Checked News? Evidence from US Democrats." *Journal of Public Economics*, 205: 104549.

Chopra, Felix, Ingar Haaland, and Christopher Roth. 2023. "The Demand for News: Accuracy Concerns versus Belief Confirmation Motives." *CEPR Discussion Paper No. DP17169*.

Chopra, Felix, Ingar Haaland, Fabian Roeben, Christopher Roth, and Vanessa Sticher. 2025. "News Customization with AI." *Working Paper*.

Chouaki, Salim, Abhijnan Chakraborty, Oana Goga, and Savvas Zannettou. 2024. "What News Do People Get on Social Media? Analyzing Exposure and Consumption of News through Data Donations." *WWW '24*, 2371–2382. New York, NY, USA:Association for Computing Machinery.

Comscore. 2022. *Web Behavior Database Panel 2022.* Wharton Research Data Services, University of Pennsylvania.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2022. "Belief elicitation and behavioral incentive compatibility." *American Economic Review*, 112(9).

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News effect: Media bias and voting." *The Quarterly Journal of Economics*, 122(3): 1187–1234.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and

Bounding Experimenter Demand." *American Economic Review*, 108(11).

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2022. "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." *American Economic Review*, 112(3): 899–927.

Djourelova, Milena. 2023. "Persuasion through slanted language: Evidence from the media coverage of immigration." *American economic review*, 113(3): 800–835.

Djourelova, Milena, Ruben Durante, and Gregory J. Martin. 2025. "The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist." *The Review of Economic Studies*, 92(3).

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and political persuasion: Evidence from Russia." *American economic review*, 101(7): 3253–3285.

Fujiwara, Thomas, Karsten Müller, and Carlo Schwarz. 2024. "The effect of social media on elections: Evidence from the united states." *Journal of the European Economic Association*, 22(3): 1495–1539.

Gentzkow, Matthew, and Jesse M Shapiro. 2006. "Media Bias and Reputation." *Journal of Political Economy*, 114(2): 280–316.

Gentzkow, Matthew, and Jesse M Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." *Econometrica*, 78(1): 35–71.

Gentzkow, Matthew, and Jesse M Shapiro. 2011. "Ideological Segregation Online and Offline." *The Quarterly Journal of Economics*, 126(4): 1799–1839.

González-Bailón, Sandra, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M Guess, Shanto Iyengar, Young Mie Kim, Neil Malhotra, Devra Moehler, Brendan Nyhan, Jennifer Pan, Carlos Velasco Rivera, Jaime Settle, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A Tucker. 2023. "Asymmetric ideological segregation in exposure to political news on Facebook." *Science*, 381(6656): 392–398.

Guess, Andrew M. 2021. "(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets." *American Journal of Political Science*, 65(4): 1007–1022.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature*, 61(1).

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political analysis*, 20(1): 25–46.

Hossain, Tanjim, and Ryo Okui. 2013. "The Binarized Scoring Rule." *Review of Economic Studies*, 80(3).

Jiménez-Durán, Rafael. 2025. "The Economics of Content Moderation: Theory and Experimental Evidence from Hate Speech on Twitter."

Johnson, Eric J., and Daniel Goldstein. 2003. "Do Defaults Save Lives?" *Science*, 302(5649).

Kalra, Aarushi. 2026. "Algorithmic Drivers of Online Behavior: Evidence from a Large-Scale Experiment."

Levy, Ro'ee. 2021. "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment." *American Economic Review*, 111(3): 831–870.

Lin, Hause, Jana Lasser, Stephan Lewandowsky, Rocky Cole, Andrew Gully, David G Rand, and Gordon Pennycook. 2023. "High level of correspondence across different news domain quality rating sets." *PNAS Nexus*, 2(9): pgad286.

Martin, Gregory J, and Ali Yurukoglu. 2017. "Bias in cable news: Persuasion and polarization." *American Economic Review*, 107(9): 2565–2599.

Mastrorocco, Nicola, Arianna Ornaghi, Matteo Pograxha, and Stephane Wolton. 2025. "Man Bites Dog: Editorial Choices and Biases in the Reporting of Weather Events."

Mullainathan, Sendhil, and Andrei Shleifer. 2005. "The Market for News." *American Economic Review*, 95(4): 1031–1053.

Mummolo, Jonathan, and Erik Peterson. 2019. "Demand effects in survey experiments: An empirical assessment." *American Political Science Review*, 113(2): 517–529.

Newman, Nic, Amy Ross Arguedas, Craig T. Robertson, Rasmus Klein Nielson, and Richard Fletcher. 2025. "Reuters Institute Digital News Report 2025."

Pennycook, Gordon, and David G Rand. 2021. "The Psychology of Fake News." *Trends in Cognitive Sciences*, 25(5): 388–402.

Peterson, Erik, Sharad Goel, and Shanto Iyengar. 2021. "Partisan Selective Exposure in Online News Consumption: Evidence from the 2016 Presidential Campaign." *Political Science Research and Methods*, 9(2): 242–258.

Pew Research Center. 2024. "Social Media and News Factsheet." *Pew Research Center*.

Rathje, Steve, Clara Pretus, James He, Trisha Harjani, Jon Roozenbeek, Kurt Gray, Sander van der Linden, and Jay Van Bavel. 2024. "Unfollowing hyperpartisan social media influencers durably reduces out-party animosity."

**Stray, Jonathan, Ian Baker, George Beknazar-Yuzbashev, Ceren Budak, Julia Kamin, Kylan Rutherford, Mateusz Stalinski, Tin Acosta, Chris Bail, Michael Bernstein, Mark Brandt, Amy Bruckman, Anshuman Chhabra, Soham De, Kayla Duskin, Sara Fish, Beth Goldberg, Andy Guess, Dylan Hadfield-Menell, Muhammed Haroon, Safwan Hossain, Michael Inzlicht, Gauri Jain, Yanchen Jiang, Alexander P. Landry, Yph Lelkes, Hongfan Lu, Peter Mason, Jennifer McCoy, Smitha Milli, Paul Resnick, Emily Saltz, Martin Saveski,

Lisa Schirch, Max Spohn, Siddarth Srinivasan, Alexis Tatore, Luke Thorburn, Joshua A. Tucker, Robb Willer, Magdalena Wojcieszak, Manuel Wüthrich, and Sylvan Zheng. 2026. "The Prosocial Ranking Challenge: Reducing Polarization on Social Media without Sacrificing Engagement."

Sunstein, Cass R. 2017. *#Republic: Divided Democracy in the Age of Social Media.* Princeton University Press.

Thaler, Michael. 2024. "The Fake News Effect: Experimentally Identifying Motivated Reasoning Using Trust in News." *American Economic Journal: Microeconomics*, 16(2).

Thaler, Richard H, and Cass R Sunstein. 2008. "Nudge: Improving decisions about health, wealth, and happiness."

Ursu, Raluca M., Andrey Simonov, and Eunkyung An. 2025. "Online Advertising as Passive Search." *Management Science*, 71(2).

U.S. Census Bureau. 2024. "American Community Survey."

Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The spread of true and false news online." *Science*, 359(6380): 1146–1151.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political Effects of the Internet and Social Media." *Annual Review of Economics*, 12(1).

Zizzo, Daniel John. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics*, 13: 75–98.

Tables and Figures

Table 1: Sample Sizes

Phase	Sample Size Description
Baseline	N = 3,353 consented to participate and completed the baseline survey.
Midline	N = 2,182 participated in the midline survey and were randomized, of which: N = 1,872 were assigned to the control group, the re-optimization treatments (R, RS, and RR), or the Reliability Info treatment (NR). N = 1,864 were in the **midline impact evaluation sample**, which means we successfully collected the set of pages they followed on Facebook at midline. N = 310 were assigned to the No-Experimenter Demand treatment (NED).
Endline	N = 1,860 participated in the endline survey, of which: N = 1,608 were in the control group, the re-optimization treatments (R, RS, and RR), or the Reliability Info treatment (NR). N = 1,580 were in the **endline impact evaluation sample**, which means we successfully collected the set of pages they followed on Facebook at the beginning of the endline survey. N=252 were in the No-Experimenter Demand treatment (NED)

Notes: This table shows the size of our sample at different stages of the experiment. The main results on the effect of our treatments on the set of news pages that participants follow on Facebook rely on the midline impact evaluation sample. The endline impact evaluation sample is primarily used to show persistence.

Table 2: Distance to Bliss Points

	Slant (1)	Reliability (2)	Euclidean (3)
Re-optimization (R)	-0.119***	-0.050***	-0.119***
	(0.029)	(0.019)	(0.022)
Re-optimization + Slant Info (RS)	-0.151***	-0.021	-0.117***
	(0.029)	(0.017)	(0.022)
Re-optimization + Reliability Info (RR)	-0.092***	-0.103***	-0.120***
	(0.027)	(0.017)	(0.022)
Reliability Info (NR)	0.000	0.002	-0.000
	(0.014)	(0.009)	(0.011)
p-value, R=RS	0.42	0.25	0.96
p-value, R=RR	0.47	0.03	0.97
p-value, RR=NR	0.00	0.00	0.00
Control Mean	0.00	0.00	-0.00
Control SD	1.00	1.00	1.00
Observations	1864	1864	1864
R-squared	0.78	0.91	0.86

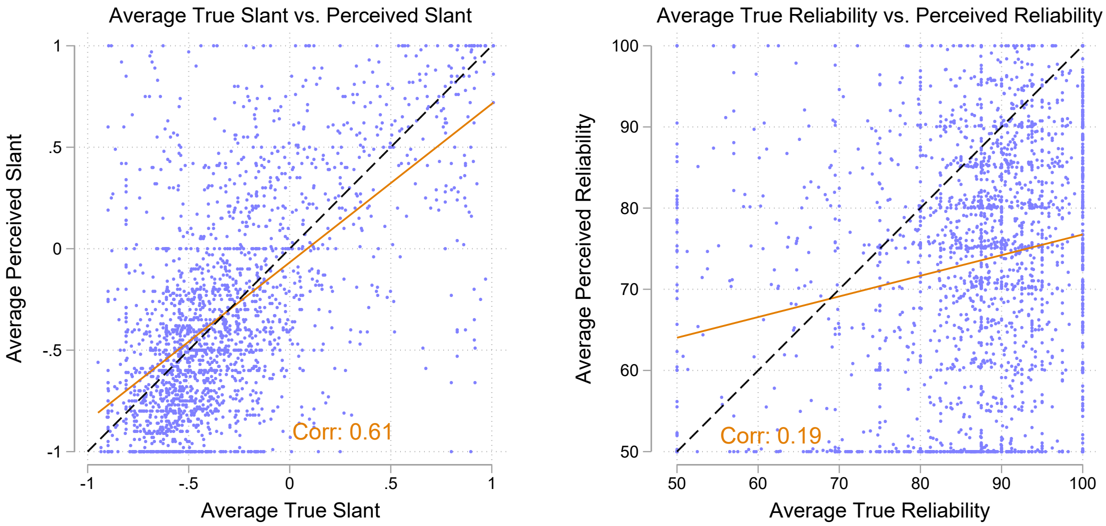
Notes: This table explores the effect of our intervention on the distance between the slant and reliability of participants' news portfolios on Facebook and participants' stated bliss points in the slant-reliability space. Specifically, it presents estimates of coefficients β_j from Equation 1 with distance to bliss point as the outcome variable. Column (1) estimates Equation 1 on the standardized absolute difference between a participant's stated bliss point on the slant dimension and the average slant of the portfolio of news pages the participant follows on Facebook. Column (2) estimates Equation 1 on the standardized absolute difference between a participant's stated bliss point on the reliability dimension and the average reliability of the portfolio of news pages the participant follows on Facebook. Column (3) estimates Equation 1 on the Euclidean distance between a participant's stated bliss point in the two-dimensional slant-reliability space and the average position in that space of the portfolio of news pages the participant follows on Facebook. The standardization was performed as follows: for each participant, we subtracted from the outcome the average outcome among participants in the control group, and divided the resulting number by the standard deviation of the outcome among participants in the control group. The regressions are run on the midline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table 3: Effects on Page Visits

	(1)	(2)	(3)	(4)	(5)
	First Stage	Extensive margin	IHS(Number of Visits)		
	Page Followed	(Facebook referred)	(Facebook referred)	Extensive margin	IHS(Number of Visits)
Pooled Reoptimization Treatments	0.195***				
	(0.010)				
Page Liked		0.040**	0.047**	0.105**	0.118
		(0.019)	(0.021)	(0.050)	(0.075)
First-stage F-stat	404.95				
Control Mean, Dep Var	0.00	0.01	0.01	0.09	0.16
Control SD, Dep Var	0.00	0.10	0.11	0.29	0.65
Observations	6156	6156	6156	6156	6156
R-squared	0.07	0.03	0.04	0.19	0.55

Notes: This table explores the effects of our intervention on online news consumption for the subset of participants who installed the Chrome browser extension that tracks their behavior on news websites. The analysis is run on a dataset where an observation is a user-page. Pooled treatments refer to the three main re-optimization treatments, namely R, RS, and RR. Belonging to any of those treatments is used as an instrument for following a page in the re-optimization interface. Column (1) displays the results of the first-stage regression on the instrument and the same set of controls as in Equation 2. Columns (2) through (5) display the results of the second-stage regression; i.e., estimates of parameter τ in Equation 2. Columns (2) and (3) consider visits to news websites originating from Facebook. Columns (4) and (5) consider all visits to news websites, independently of whether the visit originates from Facebook. Columns (2) and (4) present extensive margin results: the outcome variable is an indicator that takes value one if a user visited a news page at least once in the 31 days after the midline survey and zero otherwise. Columns (3) and (5) capture both the extensive and intensive margins: the outcome variable is the inverse hyperbolic sign of the total number of visits by a user to a page in the 31 days after the midline survey. Standard errors are clustered at the participant level.

Figure 1: Perceptions vs. Reality

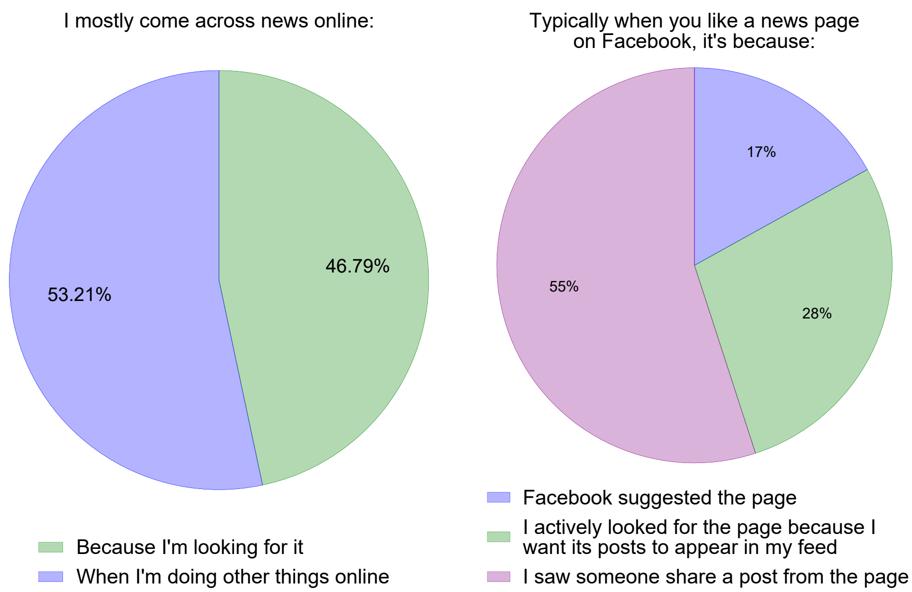


Notes: This figure explores the relationship between participants' perceptions and the actual average slant and reliability of their Facebook news portfolios. The left panel presents a scatter plot where the x-axis is the actual average slant and the y-axis is the participant's perception of the average slant. The right panel presents a scatter plot where the x-axis is the actual average reliability and the y-axis is the participant's perceptions of the average reliability. Each dot represents an individual. The solid orange lines are lines of best fit. The dashed black lines are 45-degree lines. The figure also reports Pearson correlations. To improve the clarity of the figure, we winsorize reliability values below 50 to 50. The Pearson correlation is still calculated using the non-winsorized data.

Figure 2: Bliss Points vs. Reality

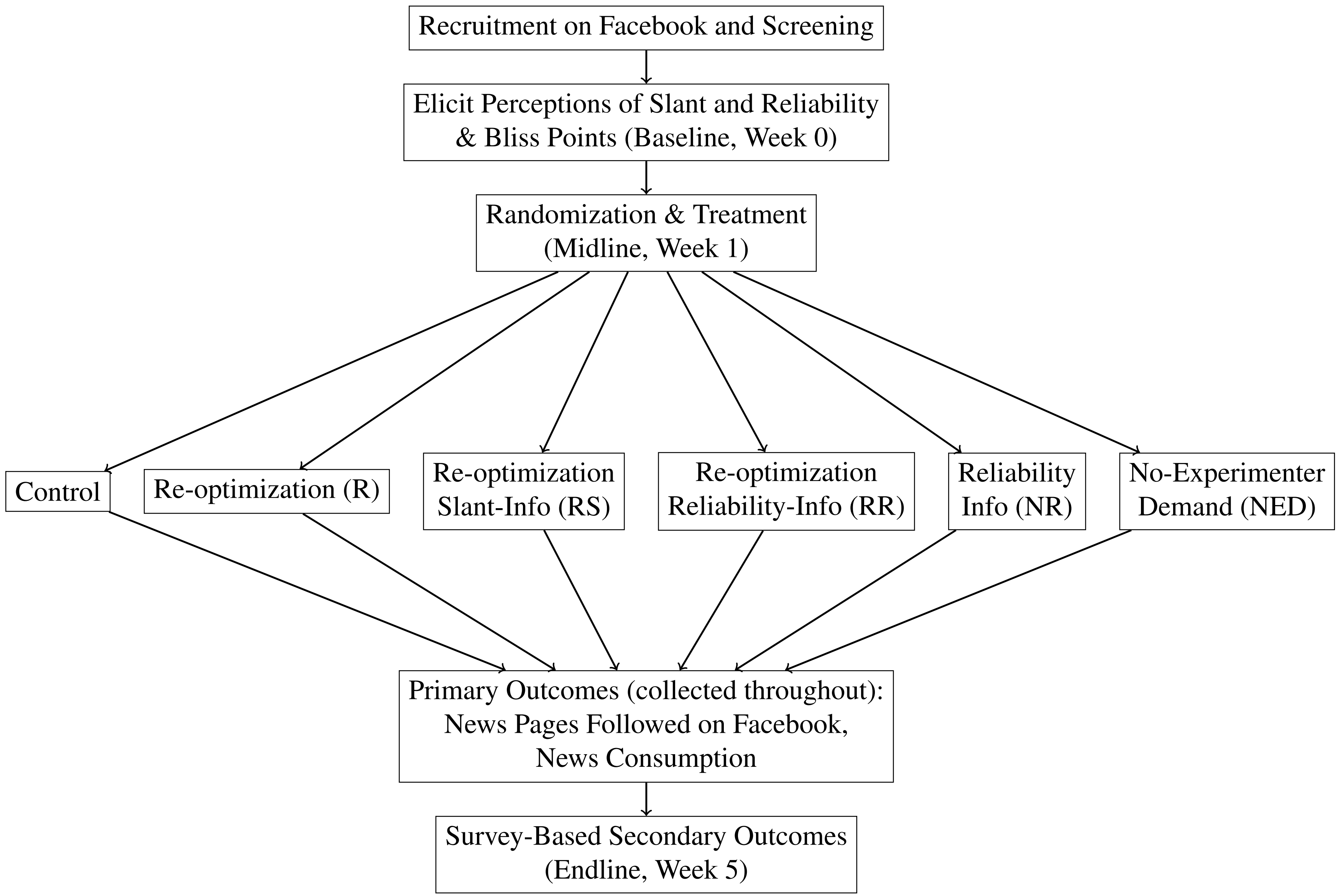
Notes: This figure explores the relationship between participants' stated bliss points and the actual average slant and reliability of their Facebook news portfolios. The left panel presents a scatter plot where the x-axis is the actual average slant and the y-axis is the participant's stated bliss point in the slant dimension. The right panel presents a scatter plot where the x-axis is the actual average reliability and the y-axis is the participant's stated bliss point in the reliability dimension. Each dot represents an individual. The solid orange lines are lines of best fit. The dashed black lines are 45-degree lines. The figure also reports Pearson correlations. To improve the clarity of the figure, we winsorize reliability values below 50 to 50. The Pearson correlation is still calculated using the non-winsorized data.

Figure 3: Evidence of Passive News Consumption Online



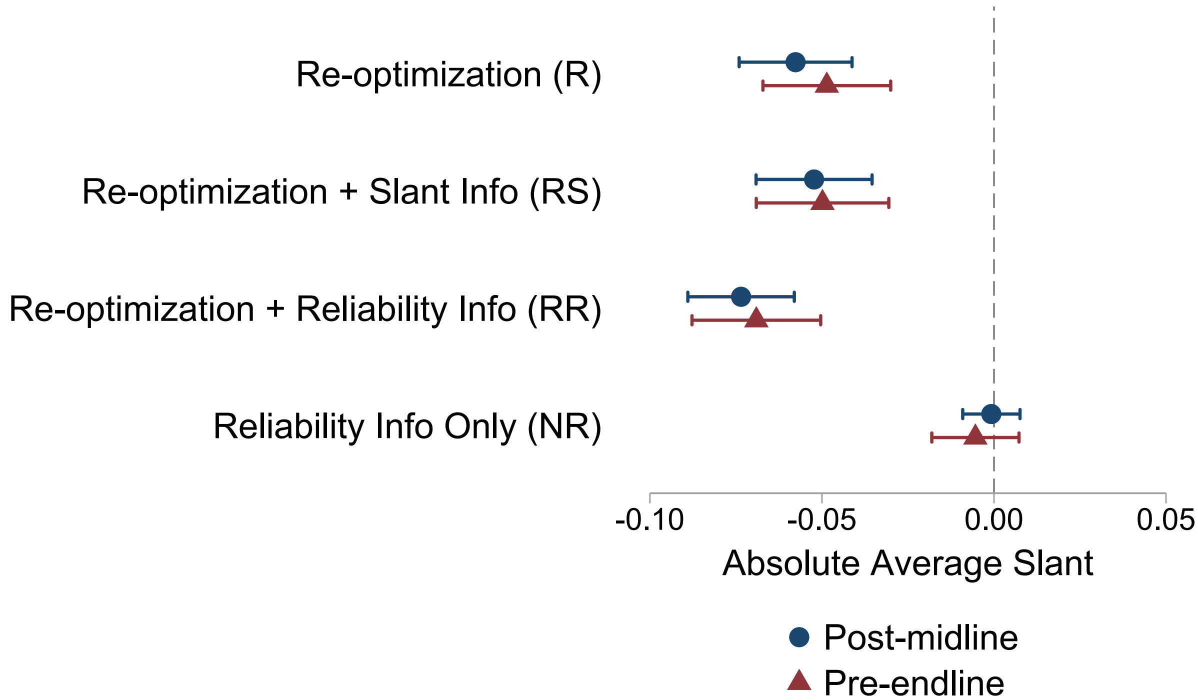
Notes: This figure presents the distribution of responses to two questions about participants' interactions with the online news ecosystem. The first question asked participants: "Which statement best describes how you get news online, whether on a computer, phone, or tablet, even if neither is exactly right?" The answer options were: "I mostly come across news online because I'm looking for it; I mostly come across news online when I'm doing other things online." The second question asked participants: "Typically when you like a news page on Facebook, does this happen because:" The answer options were: "I saw someone share a post from the page; I actively looked for the page because I want its posts to appear in my feed; Facebook suggested the page; I do not like Facebook pages; None of the above." We dropped from this analysis participants who said 'none of the above'. We also dropped 4% of the remaining participants who answered that they did not like any Facebook pages, because we knew those participants were mistaken. Specifically, had they really not liked any pages on Facebook, they would have automatically been screened out prior to this question.

Figure 4: Experimental Design



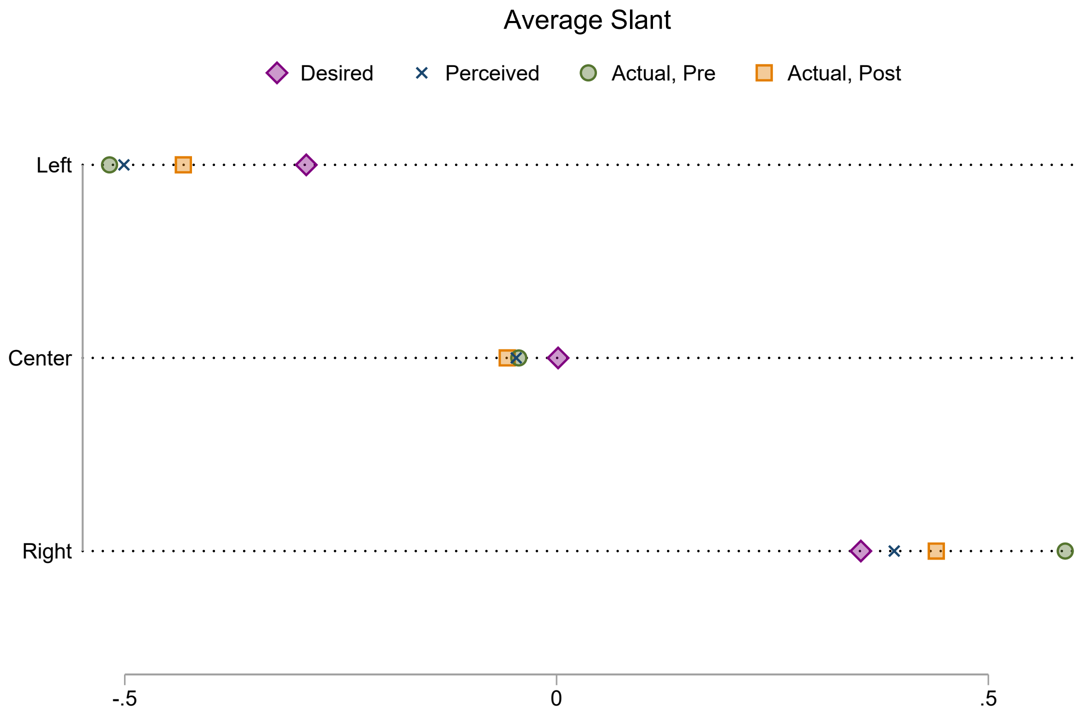
Notes: This figure illustrates the key components of our experimental design.

Figure 5: ATE on Absolute Average Slant



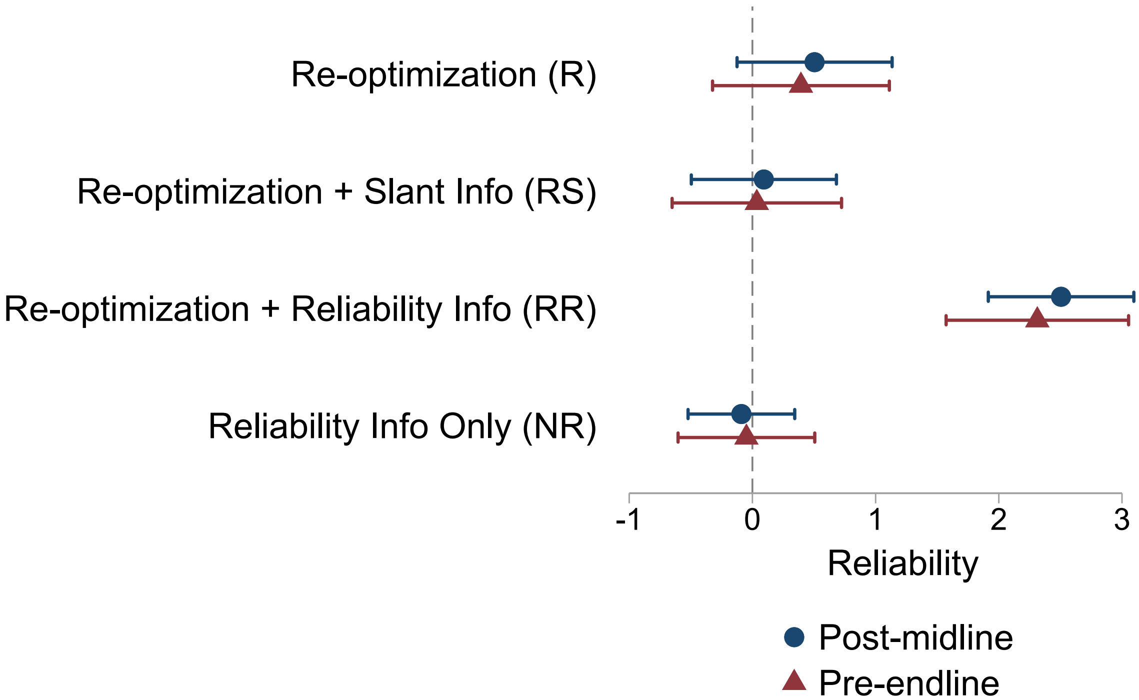
Notes: This figure explores the effect of our intervention on the absolute average slant of participants' news portfolios on Facebook. Specifically, it presents estimates of coefficients β_j from Equation 1 with the absolute average slant of a participant's news portfolio on Facebook as the outcome variable. The blue circles present estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. The red triangles estimate Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. The figure presents 95% confidence intervals using robust standard errors.

Figure 6: Average Desired, Perceived, Pre-Treatment, and Post-Period Slant by Baseline Portfolio



Notes: This figure shows the average portfolio desired slant at baseline (purple diamond), the average perceived portfolio slant at baseline (navy X), the average actual portfolio slant at baseline (green circle), and the average actual portfolio slant at midline, post-treatment (orange square). The averages are shown for participants in the re-optimization treatments (R, RS, RR) and computed separately for participants whose portfolios were left-leaning (average slant of less than -0.2), participants whose portfolios were moderate (average slant between -0.2 and 0.2), and participants whose portfolios were right-leaning (average slant of greater than 0.2).

Figure 7: ATE on Reliability



Notes: This figure explores the effect of our intervention on the reliability of participants' news portfolios on Facebook. Specifically, it presents estimates of coefficients β_j from Equation 1 with the reliability of a participant's news portfolio on Facebook as the outcome variable. The blue circles present estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. The red triangles estimate Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. The figure presents 95% confidence intervals using robust Standard errors.

Appendix For Online Publication

Table of Contents

A. Appendix Figures and Tables

Table A.1: Confusion Matrix: Slant

		Perceived		
		Left-leaning	Centrist	Right-leaning
Actual	Left-leaning	0.81	0.12	0.07
	Centrist	0.43	0.28	0.29
	Right-leaning	0.13	0.13	0.74

Notes: This table explores the relationship between participants' perceptions of the average slant of their Facebook news portfolios and the actual average slant of those portfolios. The table is row-stochastic, meaning the entries of each row sum to one. The entries of the first row show the conditional probabilities that a participant whose actual Facebook news portfolio is, on average, left-leaning perceives the average slant of her portfolio as left-leaning, centrist, and right-leaning, respectively. The other rows show similar probabilities, but conditional on the actual portfolio being, on average, centrist (row 2) or right-leaning (row 3). A portfolio is considered left-leaning if its average slant < -0.2, centrist if $-0.2 \leq$ average slant ≤ 0.2, and right-leaning if its average slant > 0.2.

Table A.2: Confusion Matrix: Reliability

		Perceived		
		Low Reliability	Medium Reliability	High Reliability
Actual	Low Reliability	0.35	0.34	0.31
	Medium Reliability	0.33	0.39	0.27
	High Reliability	0.19	0.40	0.41

Notes: This table explores the relationship between participants' perceptions of the average reliability of their Facebook news portfolios and the actual average reliability of those portfolios. The table is row-stochastic, meaning the entries of each row sum to one. The entries of the first row show the conditional probabilities that a participant whose actual Facebook news portfolio is, on average, low-reliability perceives the average reliability of her portfolio as low-reliability, medium-reliability, or high reliability, respectively. The other rows show similar probabilities, but conditional on the actual portfolio being, on average, medium-reliability (row 2) or high-reliability (row 3). A portfolio is considered low-reliability if its average reliability < 60, medium-reliability if $60 \leq$ average reliability ≤ 80, and high-reliability if its average reliability > 80.

Table A.3: Correlations between Slant and Reliability Perceptions

	(1) Democrats Perceived Reliability	(2) Republicans Perceived Reliability
Perceived Slant	-20.449***	17.110***
	(0.592)	(1.254)
Observations	7050	1952
R-squared	0.23	0.24
Clusters	1188	330

Notes: This table shows the results from a regression at the individual-outlet level of the perception of reliability on the perception of slant. We run this regression separately for Democrats and for Republicans and cluster standard errors at the participant level.

Table A.4: Baseline Portfolio Statistics

	(1) Democrats	(2) Republicans	(3) All
Avg. Number News Pages Followed	6.52	4.86	5.77
Avg. Portfolio Slant [-1, +1]	-0.46	0.19	-0.30
Avg. Portfolio Reliability [0, 100]	88.98	76.70	86.14
Observations	1028	281	1872

Notes: This table presents descriptive statistics about the baseline portfolios of news pages that participants in our experiment follow on Facebook. The table shows these statistics for our midline impact evaluation sample: participants who were assigned to the control group or to one of the main treatments, and for whom we successfully collected the set of pages they followed on Facebook during the midline survey (see Section 4.4 for details). Column (1) shows descriptive statistics for self-identified Democrats. Column (2) does the same for self-identified Republicans. Column (3) includes all participants. The number of observations in column (3) is not simply the sum of the observations in columns 1 and 2, because the former columns exclude independents.

Table A.5: Sample Demographics

	(1) Midline impact evaluation sample	(2) US Facebook users	(3) US population
Female	0.73	0.57	0.51
White	0.94	0.73	0.72
Age under 50	0.44	0.57	0.54
College	0.62	0.35	0.34
Democrat	0.55		0.20
Republican	0.15		0.15

Notes: This table compares the demographic characteristics of our midline impact evaluation sample to those of U.S. adult Facebook users and of the overall U.S. adult population. Column (1) presents the average demographics for participants in our sample (N=1,864): participants who were assigned to the control group or to one of the main treatments, and for whom we successfully collected the set of pages they followed on Facebook during the midline survey (see Section 4.4 for details). Column (2) presents our estimate of the average demographics of U.S. adults with a Facebook account. The numbers in column (2) are inferred from a Pew Research Center survey of social media use by demographic group (Pew Research Center, 2024). Column (3) presents the average demographics of American adults. The top four numbers are from the 2023 American Community Survey (U.S. Census Bureau, 2024), and the Republican and Democrat shares are from the 2024 American National Election Study (American National Election Study, 2024).

Table A.6: Balance at Midline

	Group Means						p-value treatment-control difference				
	Control	R	RS	RR	NR	NED	p(R)	p(RS)	p(RR)	p(NR)	p(NED)
Female	0.75	0.72	0.70	0.75	0.77	0.72	0.33	0.16	0.95	0.57	0.46
White	0.94	0.94	0.94	0.94	0.96	0.93	0.92	0.86	0.82	0.26	0.56
Age	51.89	53.43	53.60	52.73	51.49	53.25	0.11	0.07	0.39	0.69	0.18
College Grad	0.62	0.65	0.65	0.62	0.57	0.65	0.37	0.42	0.96	0.23	0.31
Democrat	0.57	0.54	0.56	0.56	0.50	0.52	0.42	0.76	0.67	0.05	0.14
Republican	0.11	0.17	0.16	0.15	0.16	0.16	0.02	0.06	0.13	0.06	0.06
News SM Often	0.82	0.83	0.85	0.83	0.84	0.82	0.64	0.20	0.62	0.40	0.96

Notes: This table presents balance checks on the following baseline characteristics: gender (indicator for identifying as female), race (indicator for identifying as white), age, education (indicator for being a college graduate), political ideology (indicators for identifying as Democrat or Republican), and baseline consumption of news on social media (indicator for consuming news on social media often or sometimes as rarely or never). The first six columns present means across our six study arms. The last five columns present p-values related to the null hypothesis of no difference in means between each treatment arm and the control arm. The sample includes all individuals who were randomized.

Table A.7: Differential Attrition

	(1) Pre-endline likes collected	(2) Completed Endline + post-endline likes collected
Pooled Reoptimization Treatments	0.019	0.018
	(0.021)	(0.022)
Re-optimization (R)	0.053**	0.050**
	(0.024)	(0.025)
Re-optimization + Slant Info (RS)	0.016	0.014
	(0.026)	(0.026)
Re-optimization + Reliability Info (RR)	-0.012	-0.010
	(0.027)	(0.027)
Reliability Info (NR)	0.019	0.015
	(0.027)	(0.028)
No Experimenter Demand (NED)	-0.042	-0.052*
	(0.030)	(0.030)
Control Mean	0.84	0.84
Observations	2182	2182
R-squared	0.01	0.01

Notes: This table presents the results of a battery of tests for differential attrition in our experiment. The independent variables are each treatment. We also run a separate regression where the independent variable is the pooled re-optimization treatments (R, RS, and RR). In column (1), the dependent variable is an indicator that takes the value one if and only if we were able to collect the set of news pages followed by a participant at the beginning of the endline survey. In column (2), the outcome variable is an indicator for completing the endline survey and the collection of post-endline likes. The sample includes all participants who reached the randomization stage. Standard errors in parentheses are robust.

Table A.8: ATE on Absolute Average Slant

	(1) Absolute Average Slant (post-midline)	(2) Absolute Average Slant (pre-endline)
Re-optimization (R)	-0.058***	-0.049***
	(0.008)	(0.009)
Re-optimization + Slant Info (RS)	-0.052***	-0.050***
	(0.009)	(0.010)
Re-optimization + Reliability Info (RR)	-0.073***	-0.069***
	(0.008)	(0.010)
Reliability Info (NR)	-0.001	-0.005
	(0.004)	(0.006)
p-value, R=RS	0.64	0.92
p-value, R=RR	0.15	0.09
p-value, RR=NR	0.00	0.00
Control Mean	0.45	0.46
Control SD	0.24	0.24
Observations	1864	1580
R-squared	0.70	0.67

Notes: This table explores the effect of our intervention on the absolute average slant of participants' news portfolios on Facebook. Specifically, it presents estimates of coefficients β_j from Equation 1 with the absolute average slant of a participant's news portfolio on Facebook as the outcome variable. Column (1) estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. Column (2) estimates Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.9: PATE on Absolute Average Slant

	(1) Absolute Average Slant (post-midline)	(2) Absolute Average Slant (pre-endline)
Re-optimization (R)	-0.081***	-0.074***
	(0.015)	(0.016)
Re-optimization + Slant Info (RS)	-0.080***	-0.086***
	(0.017)	(0.019)
Re-optimization + Reliability Info (RR)	-0.072***	-0.073***
	(0.010)	(0.011)
Reliability Info (NR)	-0.013*	-0.021**
	(0.008)	(0.010)
p-value, R=RS	0.95	0.59
p-value, R=RR	0.54	0.97
p-value, RR=NR	0.00	0.00
Control Mean	0.45	0.46
Control SD	0.24	0.24
Observations	1864	1580
R-squared	0.75	0.73

Notes: This table explores the population average treatment effects of our intervention on the absolute average slant of participants' news portfolios on Facebook. To create these estimates, we run the same regression we present in Table A.8 but add weights to make our sample similar to the US population. We create the weights using entropy balancing (Hainmueller, 2012) to match the following characteristics to the US population: gender (male or female), age (bins for 20-39, 40-59, and 60+), race (White or not), Hispanic (binary), party affiliation (Democrats, Independents, Republicans) and education (Bachelor's degree and above or not). The US population shares for party affiliation are based on the 2024 American National Election Survey; all other variables are based on 2024 American Community Survey 1-year estimates. We included the one participant who was 19 yearsold in the 20-year-old bin in order to not exclude any observations.

Table A.10: PATE on Reliability

	(1) Average Reliability (post-midline)	(2) Average Reliability (pre-endline)
Re-optimization (R)	0.853	0.609
	(0.540)	(0.593)
Re-optimization + Slant Info (RS)	0.284	0.261
	(0.545)	(0.594)
Re-optimization + Reliability Info (RR)	2.784***	2.352***
	(0.518)	(0.517)
Reliability Info (NR)	0.555	0.582
	(0.442)	(0.462)
p-value, R=RS	0.40	0.64
p-value, R=RR	0.00	0.01
p-value, RR=NR	0.00	0.00
Control Mean	86.52	86.48
Control SD	12.18	12.15
Observations	1864	1580
R-squared	0.83	0.82

Notes: This table explores the population average treatment effects of our intervention on the average reliability of participants' news portfolios on Facebook. To create these estimates, we run the same regression we present in Table A.15 but add weights to make our sample similar to the US population. We create the weights using entropy balancing (Hainmueller, 2012) to match the following characteristics to the US population: gender (male or female), age (bins for 20-39, 40-59, and 60+), race (White or not), Hispanic (binary), party affiliation (Democrats, Independents, Republicans) and education (Bachelor's degree and above or not). The US population shares for party affiliation are based on the 2024 American National Election Survey; all other variables are based on 2024 American Community Survey 1-year estimates.

Table A.11: Balance at Endline

	Group Means						**p-value treatment-control difference**				
	Control	R	RS	RR	NR	NED	p(R)	p(RS)	p(RR)	p(NR)	p(NED)
Female	0.75	0.71	0.69	0.75	0.76	0.70	0.27	0.12	0.95	0.66	0.20
White	0.93	0.94	0.94	0.93	0.95	0.92	0.54	0.69	0.89	0.30	0.57
Age	51.23	53.19	53.35	52.01	50.56	52.85	0.06	0.04	0.46	0.53	0.15
College Grad	0.62	0.66	0.65	0.62	0.58	0.69	0.24	0.52	0.90	0.35	0.10
Democrat	0.56	0.55	0.54	0.57	0.50	0.50	0.72	0.61	0.91	0.12	0.15
Republican	0.11	0.16	0.17	0.14	0.15	0.16	0.06	0.03	0.29	0.15	0.11
News SM Often	0.85	0.84	0.85	0.85	0.84	0.81	0.50	0.99	0.92	0.58	0.13

Notes: This table presents balance checks on the following baseline characteristics: gender (indicator for identifying as female), race (indicator for identifying as white), age, education (indicator for being a college graduate), political ideology (indicators for identifying as Democrat or Republican), and baseline consumption of news on social media (indicator for consuming news on social media often or sometimes as rarely or never). The first six columns present means across our six study arms. The last five columns present p-values related to the null hypothesis of no difference in means between each treatment arm and the control arm. The sample includes all individuals for whom we were able to collect the set of news pages they followed at the beginning of the endline survey.

Table A.12: ATE on Absolute Average Slant Among Participants Who Finished The Endline Survey

	(1) Absolute Average Slant (post-midline)	(2) Absolute Average Slant (pre-endline)
Re-optimization (R)	-0.057***	-0.049***
	(0.009)	(0.009)
Re-optimization + Slant Info (RS)	-0.055***	-0.050***
	(0.009)	(0.010)
Re-optimization + Reliability Info (RR)	-0.077***	-0.069***
	(0.009)	(0.010)
Reliability Info (NR)	-0.003	-0.005
	(0.005)	(0.006)
p-value, R=RS	0.88	0.92
p-value, R=RR	0.10	0.09
p-value, RR=NR	0.00	0.00
Control Mean	0.45	0.46
Control SD	0.24	0.24
Observations	1577	1577
R-squared	0.69	0.67

Notes: This table explores the effect of our intervention on the absolute average slant of participants' news portfolios on Facebook, among participants for whom we have data from both the midline and endline surveys. Specifically, it presents estimates of coefficients β_j from Equation 1 with the absolute average slant of a participant's news portfolio on Facebook as the outcome variable. Column (1) estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. Column (2) estimates Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.13: Effect on Slant, Channels

	(1) Absolute Average Slant (post-midline)	(2) Average Absolute Slant (post-midline)	(3) Fraction Counter-attitudinal Pages Followed (post-midline)
Pooled Reoptimization Treatments	-0.061***	-0.005	0.058***
	(0.005)	(0.004)	(0.004)
Control Mean	0.45	0.51	0.02
Control SD	0.24	0.21	0.06
Observations	1550	1550	1277
R-squared	0.64	0.72	0.30

Notes: This table presents the effect of our intervention on the absolute average slant of participants' news portfolios on Facebook (our primary outcome), the average absolute slant of those portfolios, and the fraction of counter-attitudinal pages in those portfolios. Specifically, it presents the estimate of the β coefficient for a version of Equation 1 in which all three main re-optimization treatment arms are pooled together. A page is deemed counter-attitudinal if its slant and the average slant of the participant's baseline news portfolio on Facebook have opposite signs. The regressions whose outcomes are reported in columns 1 and 2 use data collected right after the midline survey; thus, they employ the midline impact evaluation sample. The regression whose outcome is reported in column (3) uses the subset of the midline impact evaluation sample that does not have an initial portfolio with a centrist slant (i.e., their baseline average slant does not fall between -0.2 and 0.2). The regression includes pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.14: Effect on Slant, Channels by Party

	(1) Democrats Absolute Average Slant	(2) Republican Absolute Average Slant	(3) Democrats Average Absolute Slant	(4) Republican Average Absolute Slant	(5) Democrats Fraction Counter-attitudinal Pages Followed	(6) Republican Fraction Counter-attitudinal Pages Followed
Pooled Reoptimization Treatments	-0.055***	-0.051**	-0.005	0.007	0.050***	0.096***
	(0.007)	(0.023)	(0.005)	(0.015)	(0.005)	(0.022)
Control Mean	0.47	0.39	0.50	0.55	0.02	0.04
Control SD	0.21	0.31	0.20	0.26	0.06	0.09
Observations	867	229	867	229	774	152
R-squared	0.63	0.64	0.71	0.76	0.29	0.32

Notes: This table presents the effect of our intervention on the absolute average slant of participants' news portfolios on Facebook, the average absolute slant of those portfolios, and the fraction of counter-attitudinal pages in those portfolios. The table presents the effects separately for Democrats and Republicans. Specifically, it presents the estimate of the β coefficient for a version of Equation 1 in which all three main re-optimization treatment arms are pooled together. A page is deemed counter-attitudinal if its slant and the average slant of the participant's baseline news portfolio on Facebook have opposite signs. The regressions whose outcomes are reported in columns 1 and 2 use data collected right after the midline survey; thus, they employ the midline impact evaluation sample. The regression whose outcome is reported in column (3) uses the subset of the midline impact evaluation sample that does not have an initial portfolio with a centrist slant. The regression includes pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.15: ATE on Reliability

	(1) Avg. Reliability (post-midline)	(2) Avg. Reliability (pre-endline)
Re-optimization (R)	0.505	0.394
	(0.321)	(0.366)
Re-optimization + Slant Info (RS)	0.093	0.036
	(0.300)	(0.351)
Re-optimization + Reliability Info (RR)	2.505***	2.313***
	(0.301)	(0.378)
Reliability Info (NR)	-0.089	-0.049
	(0.221)	(0.283)
p-value, R=RS	0.31	0.41
p-value, R=RR	0.00	0.00
p-value, RR=NR	0.00	0.00
Control Mean	86.52	86.48
Control SD	12.18	12.15
Observations	1864	1580
R-squared	0.82	0.79

Notes: This table explores the effect of our intervention on the average reliability of participants' news portfolios on Facebook. Specifically, it presents estimates of coefficients β_j from Equation 1 with reliability of a participant's news portfolio on Facebook as the outcome variable. Column (1) estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. Column (2) estimates Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.16: ATE on Reliability Among Participants Who Finished The Endline Survey

	(1) Average Reliability (post-midline)	(2) Average Reliability (pre-endline)
Re-optimization (R)	0.507	0.399
	(0.354)	(0.367)
Re-optimization + Slant Info (RS)	0.111	0.038
	(0.337)	(0.351)
Re-optimization + Reliability Info (RR)	2.632***	2.315***
	(0.341)	(0.378)
Reliability Info (NR)	-0.008	-0.047
	(0.250)	(0.284)
p-value, R=RS	0.37	0.41
p-value, R=RR	0.00	0.00
p-value, RR=NR	0.00	0.00
Control Mean	86.52	86.48
Control SD	12.18	12.15
Observations	1577	1577
R-squared	0.81	0.79

Notes: This table explores the effect of our intervention on the average reliability of participants' news portfolios on Facebook, among participants for whom we have data from both the midline and endline surveys. Specifically, it presents estimates of coefficients β_j from Equation 1 with reliability of a participant's news portfolio on Facebook as the outcome variable. Column (1) estimates Equation 1 using data collected right after the midline survey; thus, it employs the midline impact evaluation sample. Column (2) estimates Equation 1 using data collected at the beginning of the endline survey; thus, it employs the endline impact evaluation sample. All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.17: ATE on Secondary Outcomes: News Satisfaction and Trust

	News Satisfaction (endline)	News Trust (endline)
Pooled Reoptimization Treatments	0.099*	0.083**
	(0.053)	(0.039)
Re-optimization (R)	0.077	0.090*
	(0.066)	(0.047)
Re-optimization + Slant Info (RS)	0.096	0.083*
	(0.065)	(0.047)
Re-optimization + Reliability Info (RR)	0.125*	0.082*
	(0.065)	(0.048)
p-value, R=RS	0.77	0.87
p-value, R=RR	0.45	0.87
Control Mean	2.98	2.53
Control SD	0.88	0.66
Observations	1608	1608
R-squared	0.16	0.21

Notes: This table explores the effect of our intervention on the secondary outcomes of news satisfaction and trust described in Section 4.3. Specifically, it presents estimates of coefficients β_j from Equation 1 for each secondary outcome variable, as well as the coefficient on a version of the regression that pools our main re-optimization treatments. The sample is the set of individuals who were randomized and who completed endline. Column (1) estimates Equation 1 using a measure of news satisfaction on Facebook as the outcome variable. The measure is constructed from the following question: "How satisfied are you with the news appearing in your Facebook feed?" The answer options were: "Very dissatisfied (coded as 0); Dissatisfied (coded as 1); Neutral (coded as 2); Satisfied (coded as 3); Very satisfied (coded as 4)." Column (2) estimates Equation 1 using a measure of trust in news on Facebook as the outcome variable. The measure is constructed from the following question: "How much, if at all, do you trust the news you get from your Facebook feed?" The answer options were: "Not at all (coded as 0); Not too much (coded as 1); Some (coded as 2); A lot (coded as 3)." All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.18: ATE on Secondary Outcomes: Perceived Slant and Reliability

	Perceived Slant (endline)	Perceived Reliability (endline)
Pooled Reoptimization Treatments	-0.014	0.067
	(0.060)	(0.047)
Re-optimization (R)	-0.033	0.004
	(0.073)	(0.059)
Re-optimization + Slant Info (RS)	0.059	0.079
	(0.072)	(0.056)
Re-optimization + Reliability Info (RR)	-0.064	0.128**
	(0.072)	(0.056)
p-value, R=RS	0.19	0.20
p-value, R=RR	0.65	0.03
Control Mean	-0.11	2.82
Control SD	0.88	0.72
Observations	1417	1608
R-squared	0.02	0.02

Notes: This table explores the effect of our intervention on the secondary outcomes of perceived slant and reliability described in Section 4.3. Specifically, it presents estimates of coefficients β_j from Equation 1 for each secondary outcome variable, as well as the coefficient on a version of the regression that pools our main re-optimization treatments. The sample is the set of individuals who were randomized and who completed endline. Column (1) estimates Equation 1 using a measure of the slant of the news in one's Facebook feed. The measure is constructed from the following question: ”In the last four weeks, relative to your typical Facebook feed, was the news you saw in your news feed...?” The answer options were: ”Much more liberal than usual; Somewhat more liberal than usual; Had a similar slant compared to what I am used to; Somewhat more conservative than usual; Much more conservative than usual.” If, at baseline, a participant reported primarily consuming conservative (liberal) news on Facebook, the answer options above were ranked in ascending (descending) order. This way, a negative treatment effect always indicates exposure to more counter-attitudinal news. Column (1) has fewer observations than the other columns, because we exclude individuals who, at baseline, reported primarily consuming moderate news. Column (2) estimates Equation 1 using a measure of reliability of the news in one's Facebook feed. The measure is constructed from the following question: ”In the last four weeks, relative to your typical Facebook feed, was the news you saw in your news feed...?” The answer options were: ”Much less reliable than usual (coded as a 0), Somewhat less reliable than usual (coded as a 1), Had a similar reliability compared to what I am used to (coded as a 2), Somewhat more reliable than usual (coded as a 3), Much more reliable than usual (coded as a 4).” All regressions include pre-specified controls. Our controls consist of: baseline value of the outcome variable (when available), baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization strata fixed effects. Standard errors in parentheses are robust.

Table A.19: Effects on (Self-reported) Changes to Facebook Feed

	(1) First Stage Page Followed	(2) Number of Times Seen in Feed
Pooled Reoptimization Treatments	0.165***	
	(0.005)	
Page Followed (instrumented)		0.542**
		(0.253)
Control Mean, Dep Var	0.00	0.76
Control SD, Dep Var	0.00	1.46
Observations	18636	18636
R-squared	0.05	0.23

Notes: This table explores the effects of our interventions on the content participants reported seeing on their Facebook feeds. The analysis is run on a dataset where an observation is a user-page. The users are the individuals who completed the endline survey and were assigned either to the control group or to one of our three main re-optimization treatments, namely R, RS, and RR. The pages are the pages users were shown or the potential pages they would have been shown had they been assigned to one of the three re-optimization treatments. Pooled treatments refer to the three main re-optimization treatments. Belonging to any of those treatments is used as an instrument for following a page in the re-optimization interface. Column (1) displays the results of the first-stage regression of following a news page on Facebook on the instrument and the same set of controls as in Equation 2. Column (2) displays the results of the second-stage regression. The outcome variable in column (2) is a participant's self-report of the number of times she encountered a post from the page in her Facebook feed. Standard errors are clustered at the participant level.

Table A.20: Reduced Form Effects on Page Visits

	(1) Extensive margin (Facebook referred)	(2) IHS(Number of Visits) (Facebook referred)	(3) Extensive margin	(4) IHS(Number of Visits)
Pooled Reoptimization Treatments	0.008**	0.009**	0.020**	0.023
	(0.004)	(0.004)	(0.010)	(0.014)
Control Mean, Dep Var	0.01	0.01	0.09	0.16
Control SD, Dep Var	0.10	0.11	0.29	0.65
Observations	6156	6156	6156	6156
R-squared	0.01	0.03	0.17	0.54

Notes: This table explores the reduced form effects of our intervention on online news consumption for the subset of participants who installed the Chrome browser extension that tracks their behavior on news websites. The analysis is run on a dataset where an observation is a user-page. The users are the individuals who installed the extension. The pages are the pages users were shown or the pages they would have been shown had they been assigned to one of the three re-optimization treatments. Pooled treatments refer to the three main re-optimization treatments, namely R, RS, and RR. The table presents estimates of the reduced-form of the IV specification detailed in Equation 2; in other words, we regress $V_{i,k}$ on an indicator for assignment to the pooled re-optimization treatments and include the same set of controls. Columns (1) and (2) consider visits to news websites originating from Facebook. Columns (3) and (4) consider all visits to news websites, independently of whether the visit originates from Facebook. Columns (1) and (3) present extensive margin results: the outcome variable is an indicator that takes value one if a user visited a news page at least once in the 31 days after the midline survey and zero otherwise. Columns (2) and (4) capture both the extensive and intensive margins: the outcome variable is the inverse hyperbolic sign of the total number of visits by a user to a page in the 31 days after the midline survey. Standard errors are clustered at the participant level.

Table A.21: ATE on Slant and Reliability of Visited Pages

	Extensive Margin High-reliability	IHS(visits) High-reliability	Extensive Margin Counter-Attitudinal	IHS(visits) Counter-Attitudinal
Pooled Reoptimization Treatments	0.058	0.142	0.034	0.065*
	(0.039)	(0.121)	(0.021)	(0.034)
Re-optimization (R)	0.009	0.132	0.030	0.041
	(0.047)	(0.142)	(0.028)	(0.040)
Re-optimization + Slant Info (RS)	0.077*	0.217	0.074**	0.122***
	(0.047)	(0.145)	(0.032)	(0.046)
Re-optimization + Reliability Info (RR)	0.086*	0.069	-0.008	0.026
	(0.048)	(0.150)	(0.026)	(0.053)
Reliability Info (NR)	0.059	0.006	0.004	0.035
	(0.074)	(0.205)	(0.041)	(0.055)
Control Mean	0.74	2.51	0.04	0.04
Control SD	0.44	2.10	0.19	0.23
Observations	538	538	451	451
R-squared	0.26	0.71	0.28	0.65

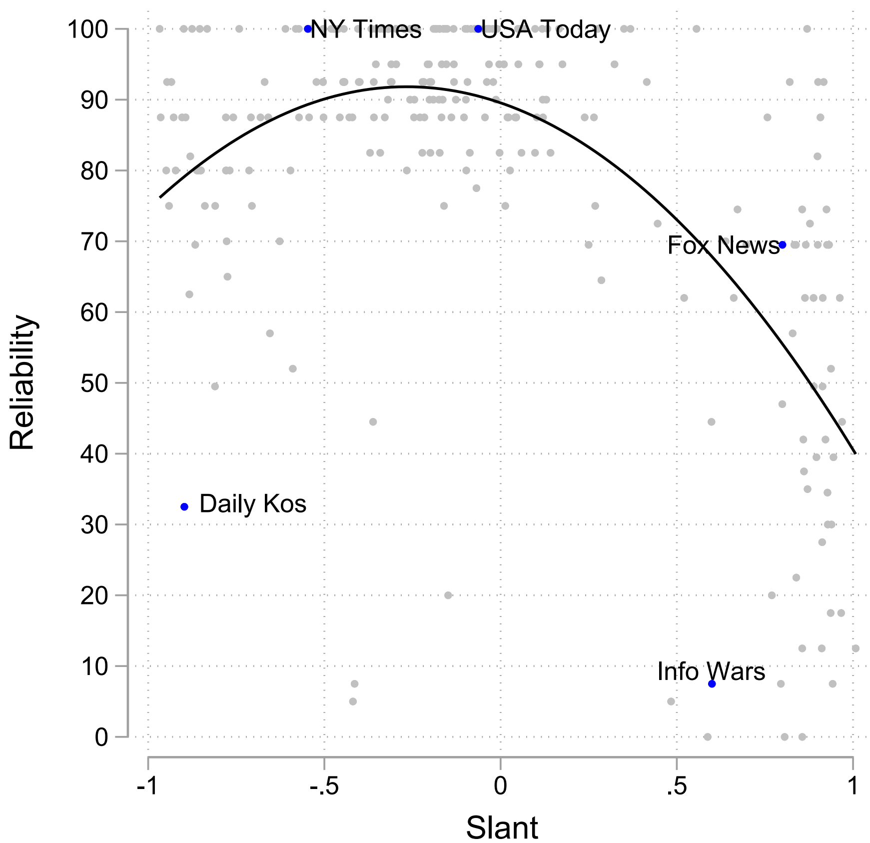
Notes: This table examines the effect of our intervention on the slant and reliability of news websites visited by participants who installed our Chrome extension. Because treatment assignment affects the likelihood of visiting any news website at all, directly regressing the average slant or reliability of visited pages on treatment indicators could yield biased estimates, as the data would be missing not at random. We therefore construct outcome variables defined for all Chrome extension users and estimate treatment effects using Equation 1. The table reports estimates of the coefficients β_j for each outcome, as well as the coefficient from a specification pooling the main re-optimization treatments. The outcome in column (1) is an indicator for visiting at least one high-reliability news website (defined as the website of an outlet with a reliability score above 80) in the 31 days after the midline survey. Column (2) uses the inverse hyperbolic sine (IHS) transformation of the number of visits to high-reliability news websites. Column (3) is an indicator for visiting at least one counter-attitudinal news website during the same period, where a website is considered counter-attitudinal if its slant has the opposite sign of the average slant of the participant's baseline Facebook news portfolio. Column (4) reports the IHS of the number of visits to counter-attitudinal news websites using the same definition. Columns (3) and (4) exclude participants with a baseline slant between -0.2 and +0.2, as counter-attitudinal news sites do not exist for such participants. All regressions include the pre-specified controls: baseline political ideology, baseline reliance on Facebook for news, baseline political interest, wave fixed effects, device fixed effects (desktop or mobile), and recruitment and randomization fixed effects. Standard errors in parentheses are robust.

Table A.22: ATE of the No-Experimenter Demand (NED) Treatment

	Absolute Average Slant (pre-endline) (1)	Average Reliability (pre-endline) (2)
Re-optimization (R)	-0.048***	0.417
	(0.010)	(0.366)
Re-optimization + Slant Info (RS)	-0.050***	0.052
	(0.010)	(0.351)
Re-optimization + Reliability Info (RR)	-0.069***	2.325***
	(0.010)	(0.375)
Reliability Info (NR)	-0.006	-0.033
	(0.007)	(0.283)
No Experimenter Demand (NED)	-0.078***	1.784***
	(0.012)	(0.388)
p-value, R=RS	0.90	0.40
p-value, R=RR	0.09	0.00
p-value, RR=NED	0.54	0.24
Control Mean	0.46	86.48
Control SD	0.24	12.15
Observations	1822	1822
R-squared	0.64	0.79

Notes: This table explores the effect of our intervention on the absolute average slant and average reliability of participants' news portfolios on Facebook. Specifically, it presents estimates of coefficients β_j from Equation 1, including not only the three re-optimization treatment arms, but also the No-Experimenter Demand (NED) treatment. Column 1 estimates Equation 1 with the absolute average slant of a participant's news portfolio on Facebook as the outcome variable. Column 2 estimates Equation 1 with the average reliability of a participant's news portfolio on Facebook as the outcome variable. The sample is the endline impact evaluation sample + the sample of participants assigned to the NED treatment for whom we could collect the pages they follow on Facebook during the endline survey. We need to rely on the endline impact evaluation sample rather than the midline impact evaluation sample, because, by design, permissions to observe the set of news pages followed by participants in the NED treatment were revoked at midline and re-obtained at endline. All regressions include pre-specified controls. Our controls consist of: baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, device fixed effects (desktop or mobile), recruitment and randomization fixed effects. Standard errors in parentheses are robust.

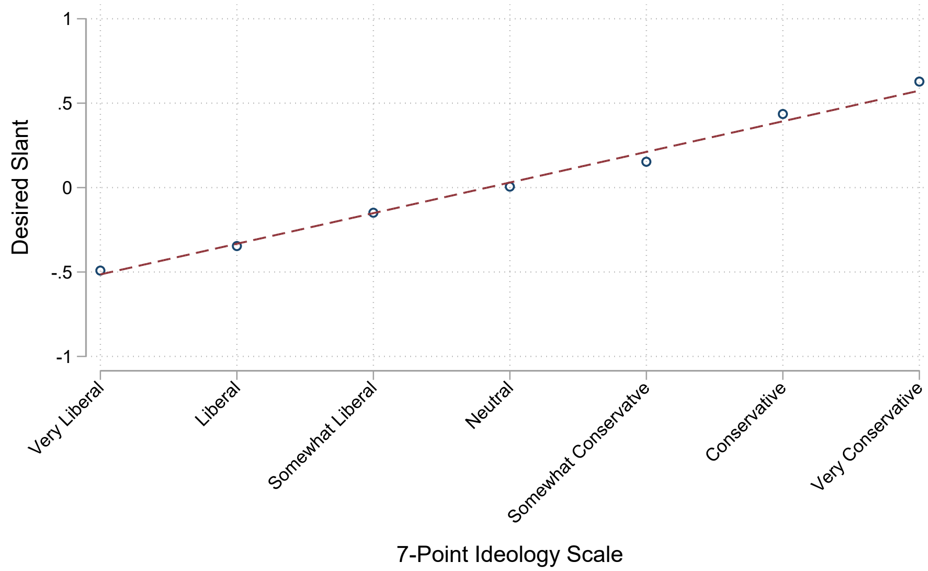
Figure A.1: Joint Distribution of Slant and Reliability for News Outlets



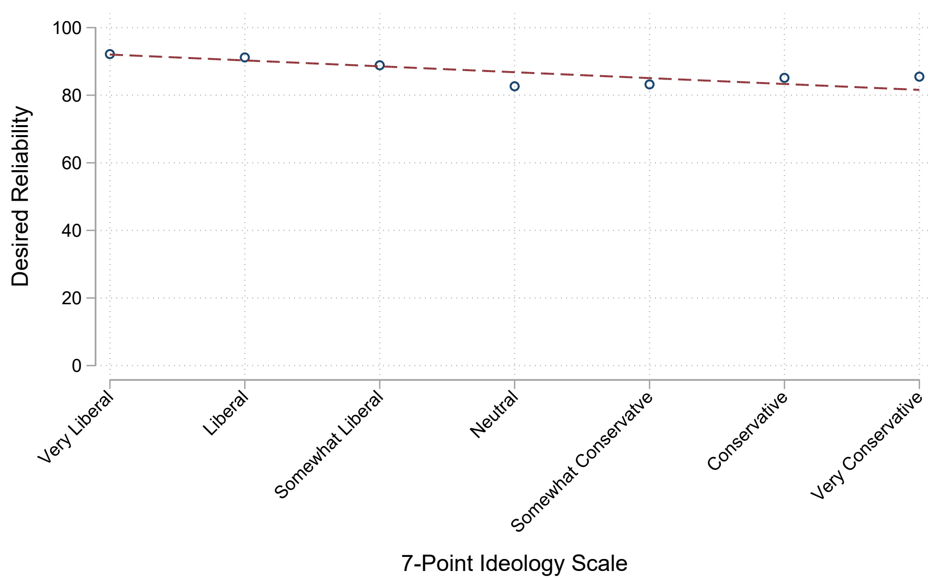
Notes: This figure shows the joint distribution of slant and reliability of the 262 news outlets for which we have both a slant rating from Bakshy, Messing and Adamic (2015) and a reliability measure from NewsGuard.

Figure A.2: Bliss point and Political Ideology

(a) Slant

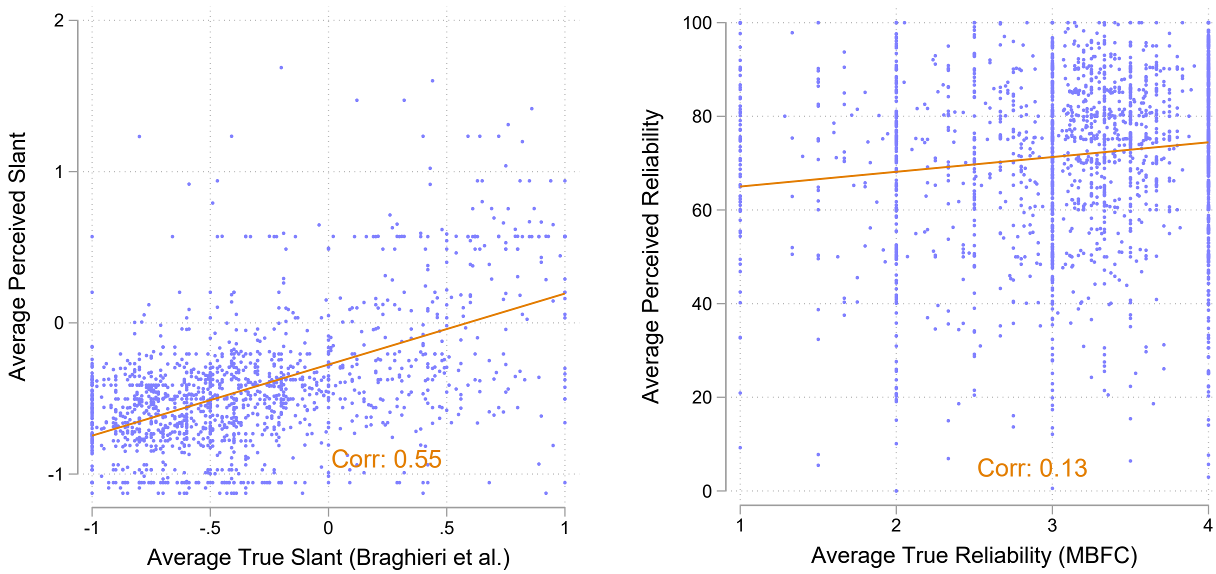


(b) Reliability



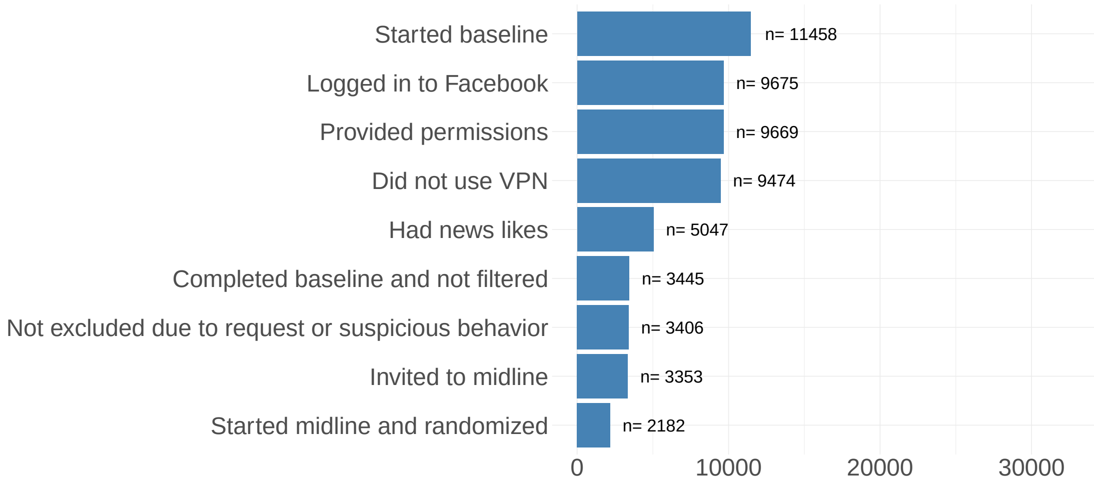
Notes: These figures explore the relationship between participants' stated bliss points on the slant and reliability dimensions and their self-reported political ideology. Specifically, the x-axis shows seven bins of self-reported political ideology, and the y-axis shows the average slant and reliability bliss points for participants whose self-reported ideology falls in the bin on the x-axis.

Figure A.3: Comparing Perceptions to Alternative Definitions of Slant and Reliability



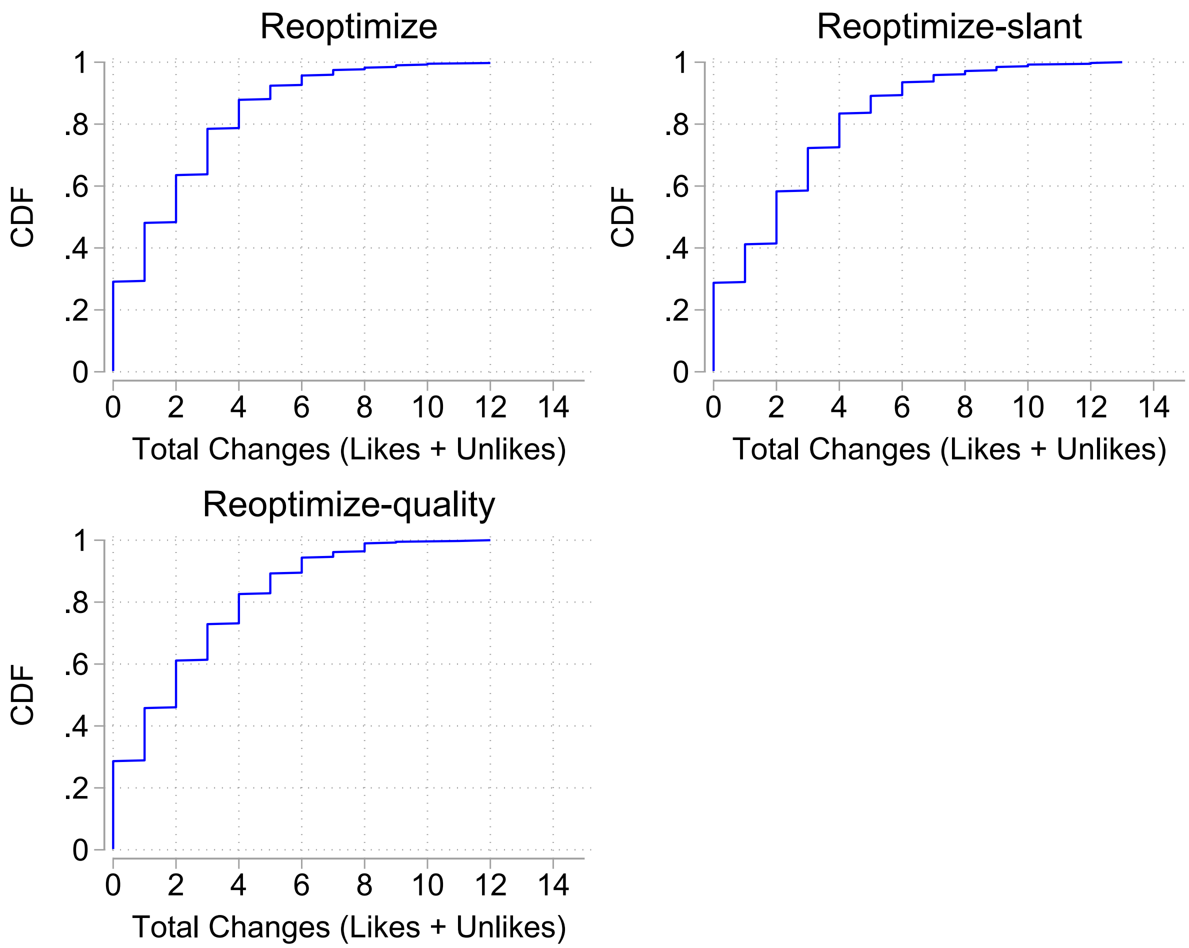
Notes: This figure explores the relationship between participants' perceptions of the average slant and reliability of their news portfolios on Facebook and the actual average slant and reliability of those portfolios, based on alternative measures of slant and reliability. The left panel focuses on slant and the right panel focuses on reliability. Each dot in each figure represents an individual. The y-axis indicates the participants' perceptions of the average reliability and slant of their news portfolio. Slant and reliability perceptions are based on our baseline survey. When these perceptions were elicited, reliability and slant were defined in the same way as in Figure 1, with slant measured based on Bakshy, Messing and Adamic (2015) and reliability based on NewsGuard. The x-axis presents the actual slant and reliability based on alternative measures. On the left panel, actual slant is measured according to Braghieri et al. (2025), where the slant of an outlet is the average slant of all articles published by the outlets. On the right panel, actual reliability is based on Media Bias Fact Check (MBFC), with outlets with low and very low reliability according to MBFC receiving a reliability score of 1, outlets with mixed reliability receiving a score of 2, mostly factual outlets receiving a score of 3, and high or very high reliability outlets receiving a score of 4. The figure also reports Pearson correlations.

Figure A.4: Participants Screened Out Before Randomization



Notes: This figure presents the number of participants in different stages of recruitment. ”Completed baseline and not filtered” refers to respondents who live in the USA, self-report getting at least some news on Facebook, passed an attention check, submitted an email address, and did not take the survey before. ”Not excluded due to request or suspicious behavior” refers to respondents who did not ask to have their Facebook data deleted and were suspected of being a fraudulent or duplicate response.

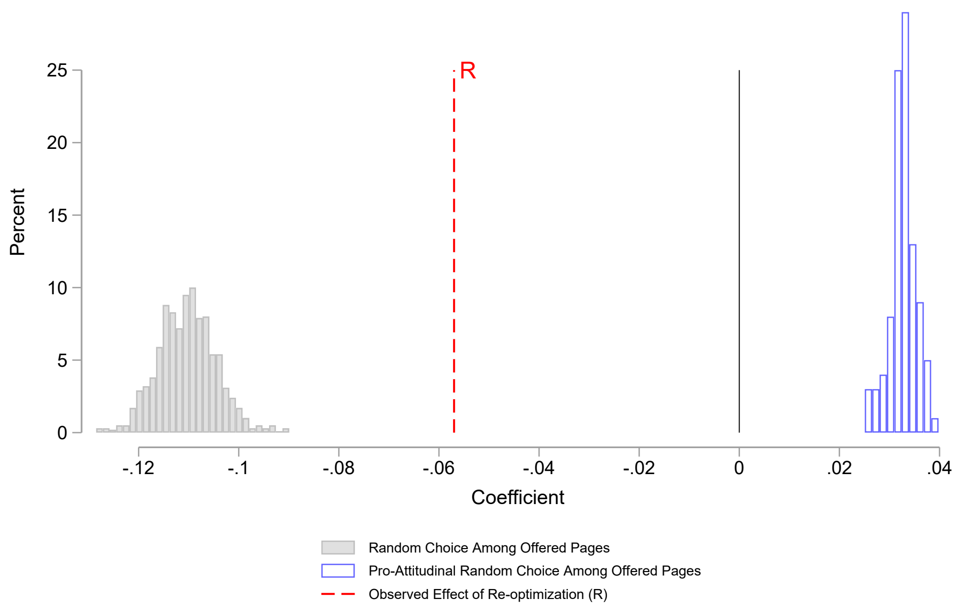
Figure A.5: CDF of portfolio changes by participants at midline



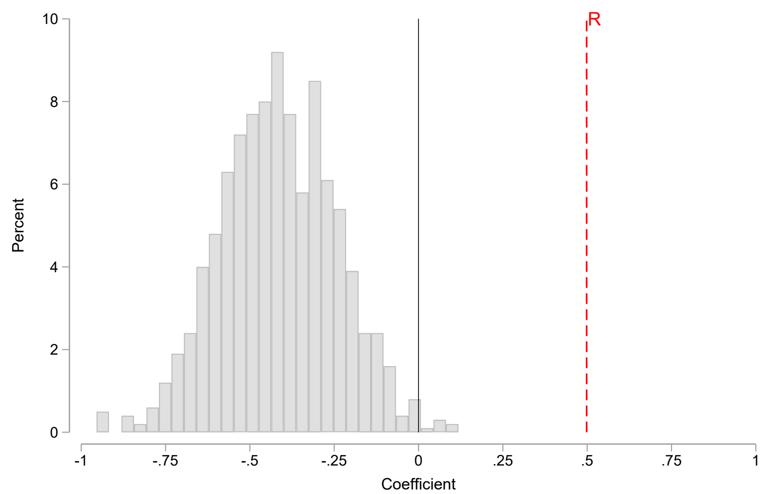
Notes: This figure shows the cumulative distribution function (CDF) of the number of total portfolio changes by participants in the re-optimization treatments during the midline survey. Following F new pages to one's portfolio and unfollowing U pages from one's original portfolio is counted as $F+U$ changes.

Figure A.6: Simulated Estimates of ATE on Absolute Average Slant and Reliability

(a) Absolute Average Slant

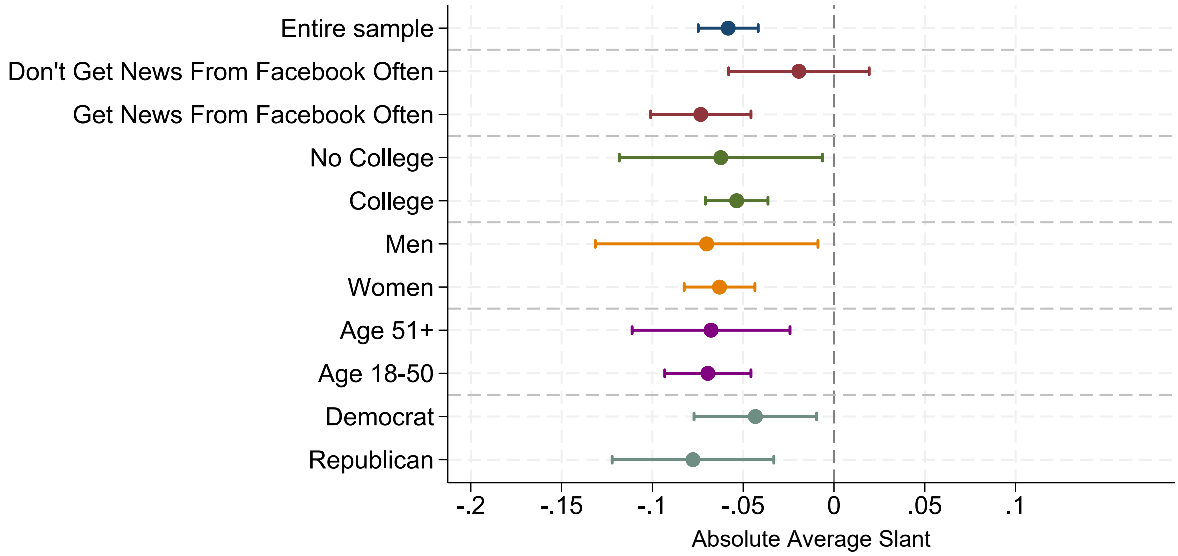


(b) Average Reliability



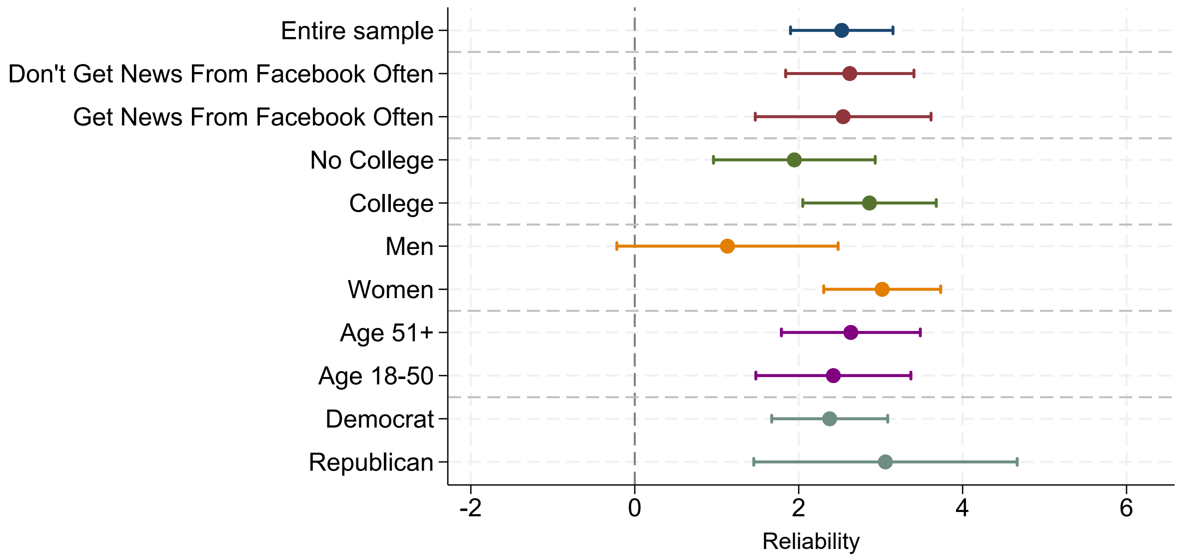
Notes: The gray bars show the distribution of simulated treatment effects on absolute average portfolio slant (Panel a) and average portfolio reliability (Panel b) when pages in our re-optimization interface are followed and unfollowed at random. Specifically, for each participant in the Re-optimization (R) treatment, we first determine how many pages they followed (F) and unfollowed (U) at midline. Next, we randomly select F pages from the set of potentially offered pages as "followed" and randomly select U pages from the participant's original set as "unfollowed." Using these simulated choices, we calculate the resulting absolute average slant and average reliability for each participant's portfolio and re-estimate Equation 1. Each gray bar corresponds to an estimate of Equation 1. The purple bars in Panel a show the distribution of simulated treatment effects on actual absolute average portfolio slant when pro-attitudinal pages are followed. Specifically, for each participant in the Re-optimization (R) treatment, we randomly select F pages from the set of pro-attitudinal potentially offered pages. This means that participants who at baseline had portfolios with average slant < -0.2 (> 0.2) randomly follow F pages from the set of potentially offered pages with slant < -0.2 (> 0.2). Participants who at baseline had portfolios with an average slant between -0.2 and 0.2 randomly follow F pages from the full set of potentially offered pages. Unfollows are not pro-attitudinal: as before, we randomly select U pages from the participant's original set as "unfollowed." Using these simulated choices, we calculate the resulting absolute average slant for each participant's portfolio and re-estimate Equation 1. Each purple bar corresponds to an estimate of Equation 1. We perform 1000 simulations of each type. In both panels, the red dotted vertical line indicates the actual estimated treatment effects on absolute average portfolio slant (Panel a) and average portfolio reliability (Panel b) for the Re-optimization treatment.

Figure A.7: Effects of the Re-optimization Treatment on Absolute Average Slant, by Group



This figure shows the effect of the Re-optimization (R) treatment on post-midline absolute average slant for different groups. The first row shows the effect among the entire sample and is identical to Table 5. The rest of the rows present regressions run for each group separately. The figure presents point estimates along with 95% confidence intervals.

Figure A.8: Effects of the Re-optimization + Reliability Info Treatment on Reliability, by Group



This figure shows the effect of the Re-optimization + Reliability Info (RR) treatment on post-midline average reliability for different groups. The first row shows the effect among the entire sample and is identical to Table 7. The rest of the rows present regressions run for each group separately. The figure presents point estimates along with 95% confidence intervals. Confidence intervals are calculated by estimating the main specification separately for each group using the post-midline average reliability variable, with 95% confidence intervals reported.

B. Measuring Slant and Reliability of News

In this section, we provide additional detail about our measures of slant and reliability.

Slant We measure news slant using the methodology introduced by Bakshy, Messing and Adamic (2015), which determines slant based on the average ideology of Facebook users who share content from a particular news domain. This measure has been widely validated and correlates well with alternative slant metrics (Gentzkow and Shapiro, 2010; Braghieri et al., 2025). It has also been extensively adopted in prior literature studying news slant (Guess, 2021; Peterson, Goel and Iyengar, 2021). The primary advantage of the Bakshy, Messing and Adamic (2015) measure is its comprehensiveness, covering 500 news outlets.

Reliability We measure news reliability using NewsGuard's Source Reliability Ratings from January 2024.[33] NewsGuard employs a team of analysts who assess outlets based on nine criteria of journalistic quality. Each criterion carries a specific number of points:

1. Does not repeatedly publish false content (22 pts)
2. Gathers and presents information responsibly (18 pts)
3. Regularly corrects or clarifies errors (12.5 pts)
4. Handles the difference between news and opinion responsibly (12.5 pts)
5. Avoids deceptive headlines (10 pts)
6. Website discloses ownership and financing (7.5 pts)
7. Clearly labels advertising (7.5 pts)
8. Reveals who is in charge, including possible conflicts of interest (5 pts)
9. Provides the names of content creators, along with either contact or biographical information (5 pts)

If an outlet meets a given criterion, it receives the full points allocated; otherwise, it receives zero points. Analysts prepare a detailed "Nutrition Report" documenting and justifying each rating. NewsGuard also reaches out to outlets for comment, and each assessment is reviewed by an experienced editor. Ultimately, each outlet receives a reliability rating ranging from 0 to 100. NewsGuard ratings are also highly comprehensive, covering thousands of outlets, and have been increasingly utilized in recent studies on news reliability (Pennycook and Rand, 2021; Allen, Arechar, Pennycook and Rand, 2021; Bhadani et al., 2022).

We merge the datasets containing slant and reliability measures, manually matching them to active Facebook pages. The resulting dataset used in this study consists of 262 unique domains.[34]

[33] www.newsguardtech.com/solutions/news-reliability-ratings

[34] We combine outlets with similar domains, such as huffingtonpost.ca and huffingtonpost.com, into a single domain.

C. Conceptual Framework: Deliberative Optimization vs. Passive Portfolio Formation

This appendix presents a simple conceptual framework that organizes the empirical analysis.

The first view is a *deliberative optimization* benchmark. In that benchmark, users hold stable preferences over slant and reliability, have a good sense of the menu of outlets available to them, and adjust their portfolios whenever useful information arrives. This benchmark captures the spirit of canonical models of news demand such as Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2006). It predicts that information about outlet quality should affect portfolio choices, while a prompt to reconsider one's portfolio—in the absence of new information—should have little effect.

The second view is a *passive portfolio formation* model. In that model, the news-portfolio-formation process is largely passive. Users do not continuously optimize over the full set of possible outlets. Instead, portfolio revision tends to occur only when a decision episode is triggered, and it is then shaped by the subset of outlets that is made especially easy to consider and act upon. This formulation is intentionally broad: it can accommodate several closely related forces, including limited attention, inertia, small hassle costs, active-choice effects, and the fact that portfolio revision on social media is often prompted by platform or friend cues rather than by deliberate search.

The contrast between these two views yields a set of comparative statics that map directly into our experimental design.

C.1 A Deliberative Optimization Benchmark

We begin with a simple benchmark that captures the main economic forces emphasized in canonical models of news demand. Consider periods indexed by $t = 1, 2, \ldots$. The economy consists of a finite set of users I that is constant across periods and a countable set of news pages J that, over the time span of the experiment, is also assumed to be constant across periods. Each user i has an ideological position $s_i \in \mathbb{R}$, and each page j has a slant $s_j \in \mathbb{R}$ and a true reliability $r_j \in \mathbb{R}$. In line with canonical models of news demand (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006), we assume that slant is relatively easy for users to perceive, whereas reliability is not perfectly observed. Over the time span of the experiment, the user's ideological position, as well as the slant and reliability of news outlets, are assumed not to vary.

Users derive utility from consuming outlets whose slant is close to their own ideological position and from consuming reliable news. Let $\hat{r}_{ijt}$ denote user i's subjective expectation of outlet j's reliability in period t. Information about reliability can change $\hat{r}_{ijt}$ over time. We assume the utility flow from following outlet j in period t is

$$u_{ijt} = \alpha \hat{r}_{ijt} - \gamma |s_i - s_j|, \tag{3}$$

where $\alpha, \gamma > 0$. The first term captures the expected value of reliability; the second captures a preference for ideologically congenial content.

Following pages is costly. To capture the idea that news content crowds out other valued content in the feed, we assume that following m pages entails a convex cost $C_i(m)$, with $C_i(0) = 0$,

$C_i'(m) > 0$, and $C_i''(m) > 0$. User i then chooses a portfolio $A_{it}^* \subseteq J$ to solve

$$A_{it}^* \in \arg\max_{A \subseteq J} \left\{ \sum_{j \in A} u_{ijt} - C_i(|A|) \right\}. \tag{4}$$

In the deliberative benchmark, the observed portfolio in period t is A_{it}^*.

This benchmark has three natural implications.

First, if users value the consumption of pro-attitudinal news and have accurate perceptions of slant, the chosen portfolio will tend to be pro-attitudinal on average.

Second, if users value reliability but perceive it imperfectly, their chosen portfolios at any time t may contain outlets that are less reliable than what they would choose under full information.

Third, and most importantly for our purposes, users are assumed to be optimizing over the relevant set of outlets given their current beliefs and preferences. Let $A_{i,t-1}^*$ denote the portfolio user i held in the previous period. If no new information arrives between $t-1$ and t, then $\hat{r}_{ijt} = \hat{r}_{ij,t-1}$ for all j, and the user's optimization problem is unchanged. With stable preferences, the user continues to choose the same portfolio: $A_{it}^* = A_{i,t-1}^*$. As a result, a prompt to reconsider one's portfolio that does not provide new information should have no effect. In contrast, new and useful information about outlet reliability arriving at time t changes $\hat{r}_{ijt}$, changes the optimization problem in (4), and should induce users to re-optimize their portfolios relative to the one they held at time $t-1$.

C.2 A Passive Portfolio Formation Model

We now consider an alternative view in which portfolio formation on social media is largely passive. The primitives of preferences and beliefs are the same as above, but the timing and scope of portfolio choice differ.

The first ingredient is that users do not revisit their news portfolios in every period t; rather, they revisit them only when a decision episode is triggered. Such episodes may be triggered by platform prompts, reminders, salient news events, or other shocks that make news-following choices top of mind. In contrast, procrastination, inertia, mild hassle costs, and the fact that users do not ordinarily think of followed pages as a portfolio to be optimized make it less likely for decision episodes to be triggered. Formally, let $A_{i,t-1}$ denote user i's inherited portfolio of followed news pages at the beginning of period t, and let $D_{it} \in \{0,1\}$ indicate whether a decision episode is triggered at time t. If $D_{it} = 0$, the user does not reconsider her portfolio and the status quo persists: $A_{it} = A_{i,t-1}$. If $D_{it} = 1$, the user is prompted to consider revising her portfolio.

The second ingredient is that, conditional on a decision episode, users do not optimize over the full set of possible outlets. Instead, they only choose from a subset of outlets that is easy to consider and act upon at that moment. Let $M_{it} \subseteq J$ denote user i's consideration set in period t. The user can add outlets in $M_{it} \setminus A_{i,t-1}$ and remove outlets in $M_{it} \cap A_{i,t-1}$, but does not actively consider outlets outside M_{it}. The consideration set may be shaped by recent exposure, platform suggestions, friend activity, familiarity, and the ease with which following or unfollowing can be implemented.

Conditional on entering a decision episode, user i chooses a revised portfolio by solving

$$\max_{A_i \subseteq M_i} \sum_{j \in A_i} u_{ij} - C_i(|A_i|). \tag{5}$$

If $D = 0$, the user simply keeps the existing portfolio.

This formulation generates several qualitative implications that differ from the deliberative benchmark.

First, users may hold portfolios that are systematically misaligned with their stated preferences. If portfolio revision is infrequent and consideration-set-dependent, users may continue to follow an inherited or incidentally accumulated set of pages. When platform suggestions are skewed toward pro-attitudinal and ideologically extreme outlets, these inherited portfolios will be more extreme than the portfolios users would choose in a re-optimization episode from a consideration set that is more representative of the overall distribution of slant and reliability.

Second, information has no effect when delivered outside a decision episode. A user may find reliability information useful in principle and update $\hat{r}_{ijt}$, yet fail to act on it unless a decision episode is triggered.

Third, an intervention that prompts active reconsideration and makes portfolio revision especially easy may generate substantial portfolio changes even without the provision of new information. The reason is not that preferences or information change, but that the intervention can trigger a decision episode and affect users' consideration sets.

C.3 Comparative Statics and Experimental Predictions

The contrast between the two frameworks yields three predictions that organize our empirical strategy.

Prediction 1: Re-optimization without new information. Consider an intervention that prompts users to reconsider the set of news pages they follow of Facebook and presents them with an easy-to-use interface for following or unfollowing outlets, but does not provide new information.

Under the deliberative optimization benchmark, this intervention should have little effect. Users are already at their optimum given their current beliefs and preferences, so merely prompting reconsideration should not substantially change behavior.

Under passive portfolio formation, the same intervention can have a large effect. By triggering a decision episode and lowering the behavioral frictions to revision, it enables users to act on preferences that were previously dormant or not implemented.

This prediction corresponds to our **R** treatment.

Prediction 2: Information without re-optimization. Consider an intervention that provides personalized information about outlet reliability, but does not prompt an active re-optimization moment and does not provide an interface that makes changes easy to implement.

Under the deliberative optimization benchmark, this intervention should affect behavior to the extent that users value reliability and were previously misinformed about it.

Under passive portfolio formation, the effect may be small or zero. Users may update beliefs without changing behavior if the information arrives outside a decision episode.

This prediction corresponds to our **NR** treatment.

Prediction 3: Information combined with re-optimization. Consider an intervention that both provides reliability information and prompts active re-optimization through an easy-to-use interface.

Both frameworks predict that this intervention can affect behavior. In the deliberative benchmark, the effect comes from improved information. In the passive portfolio formation model, the effect comes from the interaction between useful information and a prompted decision moment that makes action easy.

This prediction corresponds to our **RR** treatment.

D. Selection of News Pages

We provided each participant in our re-optimization treatments with up to 12 news pages that they could follow on Facebook. The selection of which pages to show each participant is described below.

Outlets were classified into six distinct categories based on their slant and reliability, and within each category, outlets were ranked according to their popularity on Facebook. We only included U.S.-focused outlets dedicated to hard news. For each category, we chose as defaults two prominent outlets among the top four most popular outlets. If a participant already followed one of these outlets, we offered the next most popular outlet in that category instead. We detail these categories and their rankings below. Table A.23 lists the top 3-4 outlets potentially offered in each category. In rare cases where a user already liked all or most of these outlets, they were offered the next most popular outlet in that category (not shown in the table) if such an outlet was available.

The six categories used for selecting outlets were as follows: higher-reliability conservative, lower-reliability conservative, higher-reliability moderate, lower-reliability moderate, higher-reliability liberal, and lower-reliability liberal.

As discussed, we measured slant according to the methodology of Bakshy, Messing and Adamic (2015). Outlets with a slant score greater than 0.25 were classified as conservative, those with a slant between -0.25 and 0.25 as moderate, and those with a slant below -0.25 as liberal.

We assessed reliability using NewsGuard's Source Reliability Ratings. Higher-reliability outlets were defined as those scoring 92.5 or above, meaning they failed at most one low-weight criterion as detailed in Appendix B. Lower-reliability outlets were defined as those scoring between 60 and 75. We did not include outlets with scores below 60 in the re-optimization interface to avoid promoting misinformation.

To determine popularity, we manually collected the number of Facebook follows (or likes) for each page as of January 2022. We also manually verified whether each outlet focused specifically on hard news (news covering recent events of general significance) and whether the outlet primarily targeted U.S. audiences.

Table A.23: Most common outlets offered

Name	Group	Rank	Reliability	Slant
The Wall Street Journal	Higher reliability conservative	1	100	0.28
National Review	Higher reliability conservative	2	92.5	0.9
The Daily Caller	Higher reliability conservative	3	100	0.87
Washington Examiner	Higher reliability conservative	4	92.5	0.82
Fox News	Lower reliability conservative	1	69.5	0.8
The Political Insider	Lower reliability conservative	2	69.5	0.9
Opposing Views	Lower reliability conservative	3	75	0.27
MRCTV	Lower reliability conservative	4	62	0.86
Business Insider	Higher reliability moderate	1	100	-0.06
USA TODAY	Higher reliability moderate	2	100	-0.06
Yahoo News	Higher reliability moderate	3	100	0.05
Today Show	Higher reliability moderate	4	95	-0.17
ABC News	Lower reliability moderate	1	75	-0.16
AOL	Lower reliability moderate	2	75	0.01
New York Post	Lower reliability moderate	3	69.5	0.25
The New York Times	Higher reliability liberal	1	100	-0.55
NPR	Higher reliability liberal	2	100	-0.61
NBC News	Higher reliability liberal	3	100	-0.27
TIME	Higher reliability liberal	4	100	-0.33
Mother Jones	Lower reliability liberal	1	69.5	-0.87
VICE	Lower reliability liberal	2	70	-0.63
Upworthy	Lower reliability liberal	3	75	-0.81
Atlanta Black Star	Lower reliability liberal	4	75	-0.71

Notes: This table presents the two outlets we offered participants, along with the top two alternatives. The top two outlets in each category are the ones we offered participants by default (these are always two of the four most popular outlets in each category). The rest of the outlets are ranked by popularity and were offered in case previous outlets were already followed. Only three outlets are shown for lower reliability moderate outlets because only three outlets fit our inclusion criteria. For more details, see Appendix D.

E. Implementation: Additional Details

Participants were invited to take the surveys in three sequential batches. Recruitment for the baseline survey took place in 2025 through targeted Facebook ads. Ads for the first batch ran from January 14th to January 24th, ads for the second batch from January 25th to February 13th, and ads for the final batch from February 21st to March 2nd.

On average, participants received invitations for the midline survey approximately one week after completing the baseline survey, and for the endline survey approximately one month after baseline. Invitations and multiple reminders for both midline and endline surveys were distributed via email.

Upon starting the baseline survey, participants were asked to grant us permission to collect information about the pages they follow on Facebook. We collected the data using three complementary methods:

1. Survey-based follows: Real-time collection of the full set of pages participants follow. The data was collected at the beginning and the end of each survey.

2. Webhooks-based follows: Notifications from Facebook's webhooks whenever a participant followed or unfollowed a page.

3. Snapshot-based follows: Weekly snapshots capturing all pages each participant followed and the corresponding dates they began following them.

Utilizing these three methods allowed us to maintain high-quality data and promptly detect and rectify any potential inconsistencies or errors.

We nonetheless encountered two issues during data collection:

1. We temporarily lost access to participants' follow data on Feb 13^{th} 2025 due to technical issues with our Facebook app. Consequently, we paused data collection during this period. Fortunately, most participants re-granted us access to their data upon completing the subsequent survey. This interruption means that for the first batch, data between the midline and endline surveys is partially missing, while for the second batch, data between baseline and midline surveys is partially missing (no webhooks-based follows or snapshots were collected during this period). Importantly, data on pages followed or unfollowed during the survey itself remained unaffected. We describe below how we addressed these gaps.

2. In certain cases, data obtained through survey-based and snapshot-based methods was incomplete. Whenever we suspected partial data, we cross-referenced with the webhooks data, which was not affected by this issue, to ensure accuracy.

We used these collected data to construct the following variables:

- Baseline follows: Determined using survey-based follow data. Participants for whom we could not collect survey-based data at baseline were not invited to participate in the midline survey. We cross-checked this data with later snapshots. In rare cases, we retrospectively identified partial baseline data, resulting in participants having been offered pages they already followed. These 53 participants were excluded from our final analysis dataset.

- Post-midline follows: We primarily relied on webhooks data to capture changes in follows between baseline and midline surveys, as well as pages followed or unfollowed during the midline survey. For the minority of participants affected by the temporary Facebook permissions issue (and therefore lacking data between baseline and midline), we assumed no changes occurred during this period. This assumption is reasonable, as data from unaffected users indicated minimal activity between baseline and midline surveys.

- Pre-endline and post-endline follows: To identify the list of follows before and after the endline survey, we first located the earliest complete snapshot taken after endline completion. We then combined this snapshot with webhooks data to account for any changes occurring between the beginning of the endline survey and the snapshot.[35] This approach ensured accuracy despite the temporary loss of Facebook access prior to the endline survey for some users. In the very few cases where this procedure did not yield complete data, we used the survey-based follow records as a fallback.

F. Perceptions of Slant and Reliability of Mainstream Outlets

In this appendix, we examine participants' perceptions of the slant and reliability of several mainstream news outlets. The findings echo those in Section 2.2, which focuses on the perceptions of the slant and reliability of the outlets participants actually follow. As before, individuals appear to have a better grasp of ideological slant than of reliability. We also find that Democrats and Republicans hold similar perceptions of outlets' slant, but differ markedly in their perceptions of reliability. In what follows, we describe the data and the results. We do not include the figures and tables themselves due to space limits, but they are available upon request.

In the baseline survey, in addition to eliciting perceptions of the average slant and reliability of participants' own news portfolios, we also asked about six major outlets: MSNBC, The New York Times, CNN, USA Today, The Wall Street Journal, and Fox News.

One potential concern is that participants may correctly understand which outlets are more conservative or more reliable relative to each other, but misperceive the overall scales; for instance, they might believe that all outlets are less reliable than they actually are. To address this, we also analyze the implied rankings of outlets. For each participant, we construct a rank order of the six outlets based on perceived slant or reliability, then compute the average rank for each outlet across participants.

We find that participants' perceptions of slant are broadly accurate. They correctly recognize, for instance, that Fox News is more conservative than The Wall Street Journal, which in turn is to the right of USA Today, while USA Today is more conservative than CNN, The New York Times, and MSNBC. Strikingly, Democrats and Republicans tend to agree on the relative slant of these outlets, even for those that are counter-attitudinal.

In contrast, we find that participants consistently underestimate the reliability of most outlets and misrank them relative to the truth. For example, many believe that MSNBC and CNN are at least as reliable as The Wall Street Journal, contrary to NewsGuard ratings. Moreover, Republicans and Democrats diverge sharply in their assessments of the reliability of The New York Times, CNN,

[35] To detect incomplete snapshots, we first determined the maximum number of pages followed in any snapshot. We considered any snapshot incomplete if it recorded at least 50 fewer pages than the maximum snapshot for that user.

Fox News, and MSNBC, with each group rating ideologically aligned outlets as more reliable (consistent with (Gentzkow and Shapiro, 2006))

We quantify the relative accuracy of slant and reliability perceptions by comparing participants' perceived rankings to the actual rankings using Kendall's tau-b, computed individually for each baseline respondent. Comparing the distributions of this measure for slant and reliability confirms that perceived rankings of slant are substantially more accurate than those for reliability, consistent with the analysis in Section 2.2.

Finally, we test whether our information treatments affected participants' perceptions. At midline, we re-elicited perceptions for five mainstream outlets—two already included at baseline (MSNBC, Fox News), and three new ones (The Atlantic, The Washington Post, The National Review). For each participant, we compute the absolute distance between perceived and actual slant and reliability, averaged across outlets, and estimate Equation 1. We find no detectable effect of the RS treatments on slant misperceptions. However, the RR treatment significantly reduces reliability misperceptions by 5.6 points. The NR treatment, which provides the same information without the re-optimization interface, also reduced reliability misperceptions. These findings reinforce the baseline pattern: participants already have reasonably accurate perceptions of slant, but substantial misperceptions about reliability. This helps explain why information about reliability—but not about slant—was a valuable addition to the re-optimization interface. Interestingly, the RR treatment increased slant misperceptions slightly, possibly reflecting a perceived correlation between slant and reliability.

G. Surveys

G.1 Baseline Survey

The baseline survey is available here.

G.2 Midline Survey

The midline survey is available here.

G.3 Endline Survey

The endline survey is available here.