

# Who Should Know What and Why?

## Curated Supply of Political Information

Yero Samuel Ndiaye\*

University of Cologne and Max Planck Institute for Behavioral Economics

March 27, 2026

*[EARLY DRAFT REFERRING TO EXPERIMENT 1 ONLY — EXPERIMENT 2 IN PROGRESS — COMMENTS WELCOME]*

*[Click here for most recent version](#)*

### Abstract

The information environment can be distorted not only by the spread of falsehoods but also by the selective supply of truths. I study how citizens curate which accurate information others receive based on political alignment. In a pre-registered experiment with 1,200 U.S. adults, senders decide whether to share verified immigration statistics with receivers who hold inaccurate beliefs. Senders provide accurate information at high rates (above 80%), but are approximately 8 percentage points more likely to share facts aligned with their political position than facts that contradict it, even when receivers do not take a downstream political action in the experiment. This selectivity approximately doubles when receivers make a political decision, with the amplification operating primarily through withholding of unaligned information. The interpretation of my framework is that baseline selectivity reflects a preference for belief congruence, while its amplification under political action is consistent with an instrumental component. This “curated truth”—the selective provision of accurate information that favors the sender’s political position—distorts the information environment through a channel not targeted by current policy interventions focused on false content.

*JEL classification:* D82, D83, D91

*Key words:* Information Provision, Selective Disclosure, Political Information, Immigration

## 1 Introduction

Citizens’ beliefs shape political preferences and behavior. In democracies, these preferences translate into collective decisions. Citizens thus face incentives to shape what others believe as they benefit when others’ political actions align with their interests. At the societal level, however, when beliefs are inaccurate, as they widely are,<sup>1</sup> collective decisions can be distorted. Governments and platforms have responded with a growing anti-misinformation architecture including regulatory frameworks such as the EU’s Digital Services Act and the EU Code of Conduct on Disinformation, and platform interventions such as third-party fact-checking and Community Notes. These interventions aim at inaccurate beliefs that arise from

---

\*I am grateful for valuable comments and suggestions from Arno Appfelstaedt, Lara Maria Berger, Khanh Hoang, Christoph Feldhaus, Markus Dertwinkel-Kalt, Max Rainer Pascal Grossmann, Melisa Kurtis, Felix Mindl, Johannes Münster, Axel Ockenfels, Eva Maria Oess, Malte Rattenborg, Christopher Roth, Nicolas Röver, Christoph Schottmüller, Sebastian Sieglösch, Till Stange, Johannes Wohlfahrt, and Tom Zimmermann, as well as by participants of seminars and conferences at the C-SEB Early Ideas Workshop and IMPRS BeSmart Thesis Workshop. I thank Jan-Luca Bauss, Frauke Nora Gerster, Linda Plehn, Lena Schmelig, Anselm Bührlin, and Moritz Lakenbrink for their excellent research assistance. Support from the Center for Social and Economic Behavior (C-SEB) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2126/1–390838866 is gratefully acknowledged.

<sup>1</sup>For example, in October 2024 90% of Republicans believed there was more crime in the U.S. than the year before, while FBI data indicate a 4.5% decline in violent crime ([Gallup \(2024\)](#) and [Federal Bureau of Investigation \(2025\)](#)). Conversely, in survey data a majority of self-identified “very liberal” American respondents estimated that police killed over 1,000 unarmed Black men in 2019, while the actual figure is on the order of tens ([Skeptical Research Center \(2021\)](#) and [The Washington Post \(2026\)](#)).

exposure to *false information*. This paper documents a second, distinct channel: citizens selectively withhold *true information* that could correct others' inaccurate beliefs, and this selectivity intensifies precisely when the recipient's beliefs carry political consequences.

In a pre-registered experiment with 1,200 U.S. adults recruited on Prolific with quota-balancing on age, gender, ethnicity, and party identification, I show that people biasedly select which *accurate* information others receive based on political alignment and that this selectivity doubles when the recipient's beliefs translate into a downstream political action. This *curated truth*, the selective provision of accurate information that favors the sender's political position, distorts the information environment through a channel not directly targeted by interventions focused on false content.

Prior work on politically distorted information has focused on two margins that leave the supply side largely unexplored. A first, extensive literature documents biased information *demand*: individuals seek out confirmatory information, engage in motivated reasoning, and selectively expose themselves to attitude-consistent content (Bénabou and Tirole 2002, Chopra et al. 2024, Golman et al. 2022, Peterson and Iyengar 2021). This work demonstrates why people *acquire* beliefs that may be inaccurate. A second literature establishes that information *can* affect and shift political attitudes and preferences (Alesina et al. 2023, Haaland and Roth 2020, Grigorieff et al. 2020, Facchini et al. 2022, Lergetporer et al. 2021).<sup>2</sup> Together, these findings imply that control over which information reaches whom has real political consequences. On the other hand, less is known about how individuals decide which information to *supply* to whom and why.

Related evidence documents politically selective information provision in several settings, including among fact-checkers, politicians, and citizens. Louis-Sidois (2025) shows that fact-checkers, who are institutionally committed to non-partisanship, exhibit selective coverage that reflects their affiliated outlet's political slant, and that this selectivity intensifies before elections. Freitag et al. (2021) show that politicians selectively share news items aligned with their party's political position on social media, and Shin and Thorson (2017) show that partisans on Twitter selectively share fact-checking messages that favor their own candidate and denigrate the opponent. If selective information provision persists even among actors with institutional neutrality commitments, it likely exacerbates among actors who face fewer procedural constraints. Financial incentives can further cause people to strategically supply motivated beliefs (Hagmann and Loewenstein 2017, Thaler 2021). But the observational evidence cannot identify *why* selective provision occurs: whether it reflects expressive preferences, instrumental calculation, veracity concerns, or institutional incentives remains open. This paper uses experimental variation in downstream political action to distinguish a baseline alignment effect from an action-contingent amplification — a pattern consistent with distinct projective and instrumental components. Thus, I provide evidence on micro-level behavioral mechanisms that contribute to selective provision of accurate political information and provide experimental evidence that downstream political action amplifies it.

I study the supply and allocation of accurate political information when receivers cannot observe that information was withheld. The theoretical framework distinguishes three motives for disclosure. First, senders may value the receiver holding accurate beliefs regardless of political direction (*veracity motive*). Second, senders may prefer the receiver to hold beliefs aligned with their own political position, even absent any downstream action, an expressive motive that extends the logic of projective paternalism (Ambuehl et al. 2021) from intervention in choices to the curation of beliefs (*projective motive*). Third, when the receiver acts on political preferences shaped by those beliefs, senders may care about the resulting political outcome (*instrumental motive*). The framework departs from classical disclosure models (Grossman 1981, Milgrom 1981) in one key respect: receivers do not know that a disclosure opportunity existed. Non-disclosure therefore carries no inferential content, blocking the unraveling logic that sustains full revelation in standard models. This assumption captures the defining feature of modern information environments, social media feeds, messaging apps, and editorial selection, where receivers cannot identify what information was available but not shared.

I test the framework's predictions in a pre-registered online experiment. 1,200 U.S. adults recruited on Prolific with quota-balancing on age, gender, ethnicity, and party identification each act as a sender. I choose the domain of U.S. immigration, where information effects on attitudes are well-documented, opinions are polarized, and misconceptions are common. Senders first learn the verified answers to four immigration statistics and reveal their policy preferences through an incentive-compatible donation split between a pro-immigration and an anti-immigration advocacy organization. Each sender is then matched with four receivers who were previously surveyed about the same statistics and either answered incorrectly or reported not knowing the answer. For each statistic, the sender observes the receiver's inaccurate belief state and decides whether to share the verified correct information. Crucially, receivers do not know that

---

<sup>2</sup>This includes immigration, the domain I study.

a disclosure decision has been made: shared information appears as a natural element of the survey, and withheld information leaves no trace.

The design generates two sources of identifying variation. Within each sender, two of the four statistics support and two undermine the sender’s revealed policy position, providing within-sender variation in political alignment. The key between-subject treatment varies whether receivers subsequently make a consequential political decision—the same incentive-compatible donation split between the two advocacy organizations. This provides exogenous variation in whether the receiver’s updated beliefs carry political consequences, allowing me to test whether selective disclosure intensifies in that setting. The interpretation of the resulting interaction as consistent with the instrumental motive rests on the maintained assumption that ACTION operates only through political consequentiality and not through changes in the salience of the sharing decision or the sender’s construal of the receiver’s interests.<sup>3</sup>

Three main results emerge. First, senders overwhelmingly share accurate information: provision rates exceed 80% even for information misaligned with senders’ policy preferences, indicating a strong baseline veracity norm. Second, provision is nevertheless selective: senders are approximately 8 percentage points more likely to share aligned than unaligned information ( $p < 0.01$ ), consistent with a positive projective motive. Third, and most importantly, this selectivity approximately doubles, to 16 percentage points, when receivers make a downstream political action ( $p < 0.05$ ), consistent with a positive instrumental motive. The increased selectivity operates primarily through *withholding* unaligned information (−4.6 pp) rather than through increased provision of aligned information (1.9 pp). The distortion concentrates in true information the receiver never sees—making it unobservable to both the receiver and any external monitoring system.

The data do not show that larger expected preference shifts amplify selectivity. However, as we do observe symmetrically and causally shifted beliefs, this inference is limited in strength. It is nonetheless consistent with the projective and instrumental motives responding primarily to the direction of the preference shift rather than its magnitude.

The heterogeneity suggests a theoretically informative interpretation. Baseline selectivity is concentrated among strong partisans, consistent with the projective motive operating through ideological conviction. The action-induced amplification is driven by political Independents who exhibit no baseline selectivity but become selective when political actions are introduced. Mechanism-consistent evidence is also consistent with this interpretation: veracity concerns predict provision levels but do not explain the action-induced amplification, and self-reported instrumental motivation predicts alignment bias *only* when receivers can take political action.

This paper makes four contributions. This paper documents a supply-side distortion in the provision of accurate political information. The existing literature has focused on the demand side of politically distorted information. My results establish that the supply side is independently distorted, and since receivers simultaneously demand congenial information and senders independently supply it, the two forces are likely complementary. Second, I use experimental variation in downstream political action to distinguish a baseline alignment effect from an action-contingent amplification. Under the model’s maintained assumptions, this maps to the projective and instrumental motives respectively. This distinction suggests that curated truth is potentially stronger in settings where it is most consequential—before elections, referenda, and other moments when beliefs translate into collective decisions. Third, citizen-level experimental evidence on selective provision of accurate information complements the observational evidence on politicians’ selective sharing (Freitag et al. 2021) and professional fact-checkers’ selective coverage (Louis-Sidois 2025), consistent with selective provision appearing at multiple levels, among citizens, politicians, and professional fact-checkers, though whether these patterns share a common mechanism is an open question. Fourth, I apply the logic of projective paternalism (Ambuehl et al. 2021) to a new domain: the selective disclosure of accurate political information. Where Ambuehl et al. show that decision-makers restrict others’ choice sets based on projection, I show that citizens curate the *information environment* on which choices rest. The projection operates upstream of the decision: senders shape what receivers *know*, not what receivers *can do*. This extends projection from direct intervention in choices to the curation of the informational inputs to those choices.

**Related Literature.** This paper contributes to five strands of literature. First, the paper relates to the literature on information provision and political attitudes. A growing body of work shows that factual information can shift beliefs, policy preferences, and politically relevant behavior, including in

---

<sup>3</sup>The revealed-preference design, which measures policy preferences through incentive-compatible donations rather than survey items, also addresses concerns about partisan expressive responding in political surveys raised by Bullock and Lenz (2019).

the domain of immigration that I study (Grigorieff et al. 2020, Haaland and Roth 2020, Facchini et al. 2022, Lergetporer et al. 2021, Alesina et al. 2023).<sup>4</sup> I take as given that factual beliefs can affect policy preferences and ask: how do people decide which belief-relevant information reaches others, and why? My contribution is to shift attention from the effects of information on receivers to the selective supply of information by senders.

Second, the paper speaks to the literature on motivated reasoning, selective exposure, and the demand for congenial information. Prior work shows that individuals disproportionately seek, consume, and value information that confirms their prior views, and often avoid disconfirming information (Golman et al. 2022, Peterson and Iyengar 2021, Chopra et al. 2022, 2024). Relatedly, people may prefer opinion-congruent narratives to neutral factual reporting (Bursztyjn et al. 2023). This literature studies the demand side of the information market. This paper complements it by studying the supply side: I document a supply-side analogue of motivated reasoning in which citizens curate which accurate information others are allowed to see. My paper closely relates to Thaler (2021), who shows that financial incentives cause people to strategically supply false motivated beliefs. I study verified accurate information without financial incentives, and additionally decompose expressive from instrumental motives through exogenous variation in downstream political action—a margin that has not been tested in prior work.

Third, the paper contributes to the literature on misinformation, fact-checking, and the governance of online information environments. Most existing work and policy interventions focus on false or misleading content (Barrera et al. 2020). This paper studies a distinct channel: the information in my setting is true and policy-relevant, and distortion arises from *selective allocation* rather than fabrication. In that sense, this paper is closer to the logic of selective gatekeeping than to the standard misinformation paradigm. It especially complements Louis-Sidois (2025), who provides the first comprehensive econometric study of political selectivity across fact-checkers.<sup>5</sup> This paper provides citizen-level experimental evidence on selective provision that is consistent with projective and instrumental motives in the framework.

Fourth, the paper relates to the literature on disclosure and strategic information transmission. Classical disclosure models predict unraveling toward full revelation because receivers infer bad news from silence (Grossman 1981, Milgrom 1981). Dye (1985) and Jung and Kwon (1988) show that unraveling need not obtain when the receiver assigns positive probability to the sender being uninformed, and experimental evidence confirms that unraveling is often incomplete in practice (Jin et al. 2021). My framework implements a stronger departure: receivers do not know that a disclosure opportunity existed at all. This captures modern information environments where receivers observe what is shared but not what was available and omitted, linking political information supply to disclosure theory under conditions where the standard unraveling logic breaks down entirely.

Fifth, the paper connects to the literature on paternalism, projection, and intervention in others' decisions. Ambuehl et al. (2021) show that individuals restrict others' choice sets based on projected aspirations. Related work shows that people substitute between providing information and directly intervening in others' choices (Grossmann 2024) and use both hard restrictions and soft information provision when given the opportunity to intervene (Bartling et al. 2023). My setting shuts down the direct-intervention channel by design: senders can only provide or withhold information, not override the receiver's decision. This makes the experiment a clean test of whether projective preferences operate upstream, through the curation of information on which choices rest, extending the domain of projective paternalism from direct intervention in choices to the informational inputs to those choices.

Relative to existing work on information demand, misinformation, and disclosure, this paper identifies a channel through which inaccurate political beliefs persist, selective withholding of accurate information, that is distinct from misinformation, not addressed by interventions focused on false content, and potentially stronger in settings where political stakes are highest.

The paper is structured as follows. Section 2 develops a framework of alignment-based selective disclosure. Section 3 describes the experiment. Section 4 presents experimental results. Section 5 discusses the results. Finally, Section 6 concludes.

---

<sup>4</sup>Haaland et al. (2023) provide a comprehensive survey of information provision experiments.

<sup>5</sup>Andreottola and Louis-Sidois (2025) develop a theoretical companion in which two politicians report facts and a fact-checker selects one to verify. They show that a neutral fact-checker must produce balanced conclusions even if politicians differ in their propensity to fabricate. The observed imbalances in PolitiFact's coverage are instead consistent with partisan bias in selection. Their model concerns the fact-checker's verification choice; mine concerns the citizen's disclosure choice over true information.

## 2 Theoretical Framework

I develop a theoretical framework of selective disclosure of true information. A sender who knows a fact decides whether to reveal it to a receiver whose beliefs are inaccurate. The sender values accuracy but also cares about which factual beliefs and policy preferences the receiver holds. Additionally, when the receiver acts on the political preferences those beliefs inform, the sender cares about the resulting political outcome. I nest these three motives in a single utility function and derive comparative statics that map onto my experimental treatments.

My framework departs from classical disclosure models (Grossman 1981, Milgrom 1981) in one key respect. Standard models predict unraveling as senders reveal all information because receivers infer bad news from silence. This logic requires receivers to know that the sender *could* have disclosed. In many information environments such as social media feeds, messaging apps, fact-checking, and editorial selection, the receiver does not know that a disclosure opportunity existed. Non-disclosure therefore carries no inferential content, and selective withholding can persist.<sup>6</sup> My model also applies the logic of projective paternalism (Ambuehl et al. 2021) to selective disclosure: the sender derives utility when the receiver holds factual beliefs and policy preferences aligned with the sender’s worldview, independent of any downstream action.

### 2.1 Environment

Consider two agents: a female sender  $s$  and a male receiver  $r$ . First, nature determines the true state of world  $\omega \in \{0, 1\}$ , the sender’s political preference, and whether the receiver will subsequently take a political action  $A = 1$  or will not take political action  $A = 0$ . Without loss of generality, I normalize the true state of the world  $\omega = 1$ . The sender then observes  $\omega$ , the receiver’s prior belief about the state of the world  $b = \Pr(\omega = 1) \in (0, 1)$ , and  $A$ . Next, the sender chooses to disclose the true state of the world ( $d = 1$ ) or withhold it ( $d = 0$ ). If  $d = 1$ , the receiver observes the disclosed information and updates his belief to  $b' \in (b, 1]$ . If  $d = 0$ , the receiver’s belief remains at  $b$ . Crucially, the receiver does not know that a disclosure opportunity existed. Finally, if  $A = 1$ , the receiver takes a political action based on his (possibly updated) beliefs.

Let  $\hat{p}_r(b) \in [0, 1]$  denote the sender’s assessment of the probability that the receiver also prefers the sender’s preferred political position when his belief is  $b$ .<sup>7</sup> The sender’s beliefs about the preference shift from disclosure is  $\delta = \hat{p}_r(b') - \hat{p}_r(b)$ . I say the state of the world is *aligned* with the sender’s preferences if  $\delta > 0$  and *unaligned* if  $\delta < 0$ .

### 2.2 Sender Utility

The sender’s utility from disclosure combines three motives. First, a *veracity motive* ( $\theta_V \geq 0$ ): the sender values the receiver holding accurate beliefs, independent of content. This reflects truth-telling norms, aversion to withholding, or prosocial preferences for informed receivers. Second, a *projective motive* ( $\theta_P \geq 0$ ): following the literature on projective paternalism (Ambuehl et al. 2021), the sender prefers the receiver to hold beliefs and preferences aligned with her own political position, even when the receiver takes no action. Third, an *instrumental motive* ( $\theta_A \geq 0$ ): when the receiver acts on his preference ( $A = 1$ ), the sender cares about the direction of his preference, because the receiver’s action affects political outcomes the sender values like donations and voting. For aligned information, disclosure moves the receiver’s preference toward the sender’s; for unaligned information, it moves preferences away.

Combining these motives, the sender’s utility from disclosure is:<sup>8</sup>

$$U(d) = \left( \theta_V + (\theta_P + \theta_A A) \cdot w(\delta) \right) \cdot d, \quad (1)$$

where  $w : \mathbb{R} \rightarrow \mathbb{R}$  is weakly increasing with  $w(0) = 0$ ,  $w(\delta) > 0$  for  $\delta > 0$ , and  $w(\delta) < 0$  for  $\delta < 0$ . The function  $w$  maps the sender’s beliefs about the preference shift into the utility weight the sender places on that shift. The weak conditions on  $w$  require only that the sender’s utility response is sign-preserving.

<sup>6</sup>This is related to possibly uninformed senders in Dye (1985) and Jung and Kwon (1988). In their models, the receiver assigns positive probability to the sender being uninformed, which is sufficient to block full unraveling. My design implements a stronger version: receivers do not know that information can be disclosed.

<sup>7</sup>Factual beliefs can causally affect policy preferences as is documented experimentally (see Alesina et al. (2023), Haaland and Roth (2020), Grigorieff et al. (2020)). I take this relationship as given and study who controls the supply of belief-relevant information.

<sup>8</sup>The additive structure follows Bénabou and Tirole (2006) in nesting prosocial, projective, and instrumental motives. Non-negativity of  $\theta_V$ ,  $\theta_P$ ,  $\theta_A$  rules out adversarial senders who prefer receivers to be misinformed. The model nests pure veracity ( $\theta_P = \theta_A = 0$ ), pure projection ( $\theta_V = \theta_A = 0$ ), and pure strategy ( $\theta_V = \theta_P = 0$ ) as special cases.

## 2.3 Disclosure Decision

The sender discloses if and only if  $U(d = 1) > \varepsilon$ , where  $\varepsilon$  is an idiosyncratic utility shock drawn from a distribution  $F$  with full support on  $\mathbb{R}$ , mean zero, and strictly positive density  $f = F'$ .<sup>9</sup> Let  $\delta^+ > 0$  and  $\delta^- < 0$  denote the beliefs about the preference shifts for aligned and unaligned information, respectively. Write  $w^+ = w(\delta^+) > 0$  and  $w^- = w(\delta^-) < 0$ , and let  $c(A) = \theta_P + \theta_A A$ . The disclosure probabilities for aligned and unaligned information are:

$$\begin{aligned} p^+(A) &= \Pr(d = 1 \mid \delta^+, A) = F(\theta_V + c(A) \cdot w^+), \\ p^-(A) &= \Pr(d = 1 \mid \delta^-, A) = F(\theta_V + c(A) \cdot w^-). \end{aligned}$$

The veracity motive raises both probabilities. The projective and instrumental motives create a wedge between them, increasing  $p^+(A)$  and decreasing  $p^-(A)$ . This wedge widens when the receiver takes a political action.

## 2.4 Predictions

Next, I derive the predictions of my framework.

**Proposition 1** (Alignment-based selective disclosure). *If the sender values belief and preference alignment ( $\theta_P > 0$ ) and the sender believes the information is relevant for the receiver's preferences ( $\delta \neq 0$ ), the sender is strictly more likely to disclose aligned than unaligned information:*

$$p^+(A) > p^-(A) \quad \text{for all } A \in \{0, 1\}.$$

*Proof.* Since  $w^+ > 0 > w^-$  and  $c(A) > 0$  by  $\theta_P > 0$ , the argument of  $F$  is strictly larger for aligned information for all  $A \in \{0, 1\}$ . Since  $F$  is strictly increasing,  $p^+(A) > p^-(A)$ .  $\square$

Proposition 1 establishes selective disclosure as a consequence of projective preferences alone. Even absent any downstream political action and with only true information available, the sender is more likely to share facts aligned with her political worldview.

**Proposition 2** (Political action amplifies selectivity). *If the sender values the receiver's political action ( $\theta_A > 0$ ) and the sender believes the information is politically relevant ( $\delta \neq 0$ ), then the alignment-based selectivity in disclosure is strictly larger when the receiver acts:*

$$p^+(1) - p^-(1) > p^+(0) - p^-(0).$$

*Proof.* Define the alignment gap  $G(A) = F(\theta_V + c(A) \cdot w^+) - F(\theta_V + c(A) \cdot w^-)$ . Differentiating with respect to  $c$ :

$$\frac{\partial G}{\partial c} = w^+ \cdot f(\theta_V + c \cdot w^+) - w^- \cdot f(\theta_V + c \cdot w^-) > 0,$$

since  $w^+ > 0$ ,  $w^- < 0$ , and  $f > 0$  everywhere. Since  $\theta_A > 0$  implies  $c(1) > c(0)$ , we have  $G(1) > G(0)$ .  $\square$

Proposition 2 establishes selectivity increases in political action. When the receiver takes a political action, the sender gains additional utility from having shifted the receiver's preference in the sender's direction. This increases the utility of disclosing aligned information and the disutility of disclosing unaligned information, widening the gap.

**Proposition 3** (Beliefs about persuasiveness amplifies selectivity). *Fix  $A \in \{0, 1\}$  with  $\theta_P + \theta_A A > 0$ . If additionally  $|w(\delta)|$  is strictly increasing in  $|\delta|$ , then the alignment gap  $G(\delta^+, \delta^-) = p^+(A) - p^-(A)$  is strictly increasing in  $\delta^+$  and strictly increasing in  $|\delta^-|$ :*

$$\frac{\partial G}{\partial \delta^+} > 0, \quad \frac{\partial G}{\partial |\delta^-|} > 0.$$

<sup>9</sup>The shock  $\varepsilon$  captures components of the disclosure decision not modeled explicitly like mood, attention, idiosyncratic features of the sender-receiver pair, etc.

*Proof.* The alignment gap is  $G(\delta^+, \delta^-) = F(\theta_V + c \cdot w(\delta^+)) - F(\theta_V + c \cdot w(\delta^-))$  where  $c = c(A) > 0$ . Differentiating with respect to  $\delta^+$ :

$$\frac{\partial G}{\partial \delta^+} = c \cdot w'(\delta^+) \cdot f(\theta_V + c \cdot w(\delta^+)) > 0,$$

since  $|w(\delta)|$  strictly increasing in  $|\delta|$  implies  $w'(\delta^+) > 0$ , and  $c > 0$ ,  $f > 0$  everywhere. For the unaligned margin, since  $|\delta^-| = -\delta^-$ :

$$\frac{\partial G}{\partial |\delta^-|} = -\frac{\partial G}{\partial \delta^-} = c \cdot w'(\delta^-) \cdot f(\theta_V + c \cdot w(\delta^-)) > 0,$$

since  $|w(\delta)|$  strictly increasing in  $|\delta|$  for  $\delta < 0$  implies  $w'(\delta^-) > 0$ . □

Proposition 3 links selectivity to the magnitude of the beliefs about the preference shift through two independent channels: senders are more likely to share aligned information they believe to be more persuasive, and more likely to withhold unaligned information they believe to be more persuasive. Propositions 1 and 2 hold for any  $w$  satisfying the weak conditions in Section 2.2. Proposition 3 requires additionally that the sender’s utility responds not only to the *direction* but also to the *magnitude* of the beliefs about the preference shift. A rejection of this Proposition would indicate that senders categorize information as aligned or unaligned without scaling their response to the degree of the shift, i.e.,  $w$  closer to a sign function than a linear transformation.

## 2.5 Welfare Implications

I adopt an epistemic welfare benchmark: the receiver values making the decision he would make under accurate beliefs, where preferences follow the true mapping  $p_r(b)$ .<sup>10</sup> I remain agnostic about which policy is socially optimal.

The information provision literature establishes that accurate information can shift beliefs toward the truth and policy preferences.<sup>11</sup> If  $\omega = 1$  and disclosure shifts  $b' \geq b$ , under this benchmark any act of disclosure weakly improves the receiver’s welfare because it moves beliefs closer to the true state. However, the framework predicts the *probability* of receiving this improvement depends on the information’s political direction. By Proposition 1, the receiver is strictly more likely to learn a truth that favors the sender’s position than one that contradicts it. This distortion arises not because the receiver encounters false information, but because accurate information reaches him with probability that depends on its political direction. This is what distinguishes curated information from misinformation.

A final asymmetry deserves emphasis. If the veracity motive is strong enough to push aligned provision  $p^+(A)$  near its upper bound, the residual distortion operates primarily through reduced provision of unaligned information, since  $p^+(A)$  has limited room to increase. Then, the potential cost of curated truth comes from true information the receiver never sees.

## 3 Experiment

In this section, I present the experiment. First, I discuss the experimental design (section 3.1). Then, I discuss the identifying variations and map my theoretical framework into hypotheses (section 3.2). Next, I discuss my outcome measures and the empirical strategy (section 3.3). I conclude by discussing the sample and the experimental protocol (section 3.4).

### 3.1 Experimental Design

I test the predictions of Section 2 in a pre-registered online experiment with  $N = 1,200$  representative U.S. adults recruited on Prolific. I choose policy domain U.S. immigration as (i) information effects on attitudes are well-documented,<sup>12</sup> (ii) opinions are polarized, and (iii) misconceptions about factual information are common.

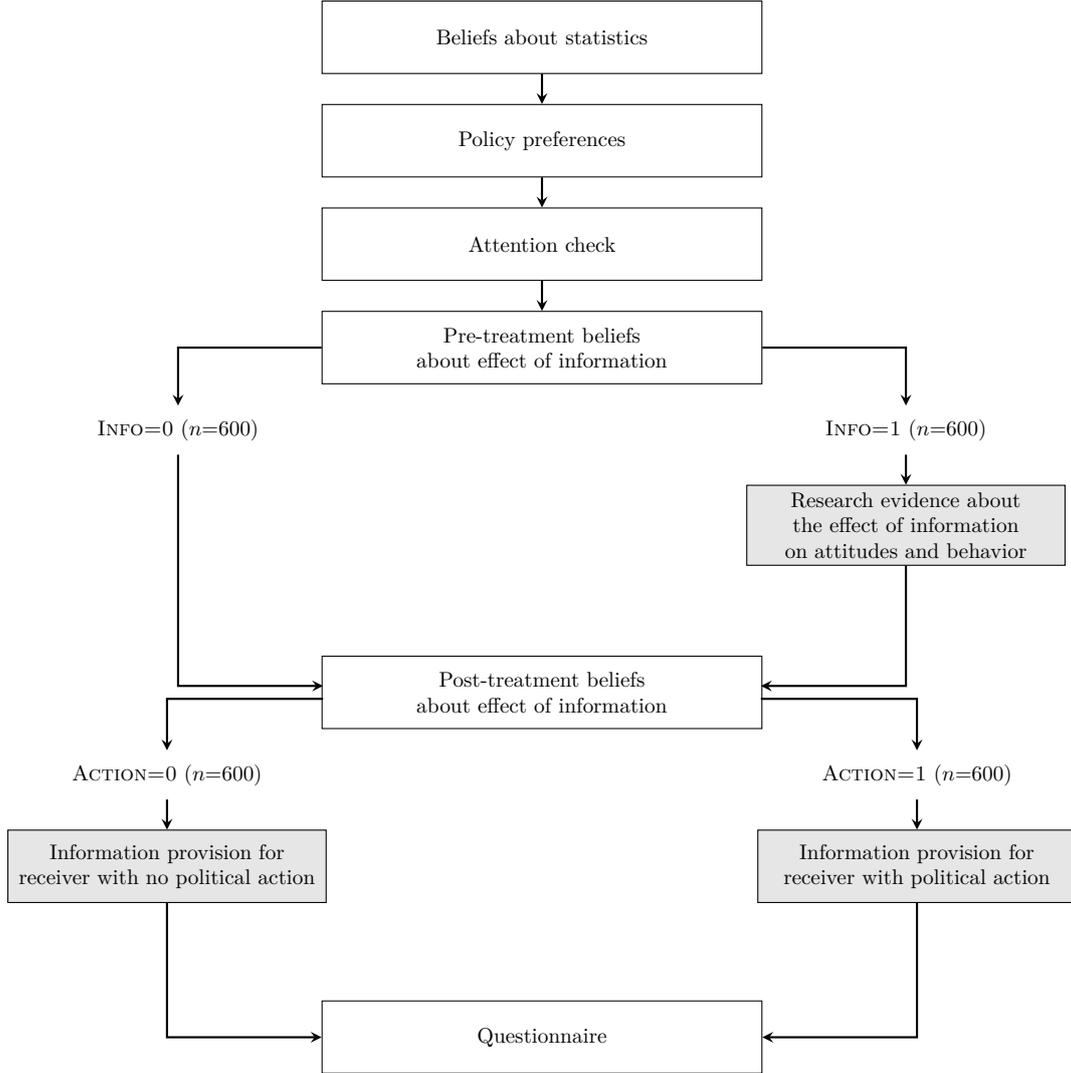


Figure 1: Experimental timeline and conditions

### 3.1.1 Experimental Timeline

Figure 1 summarizes the experimental flow, which proceeds in three parts. In Part 1, I elicit senders’ prior knowledge about four immigration statistics and measure their policy preferences through a donation decision. In Part 2, I elicit beliefs about the effect of information on political attitudes. Half of the senders receive information about the actual effect of information in the experimental study by [Haaland and Roth \(2020\)](#). In Part 3, senders make one information provision decision per statistic, each directed at a specific receiver. Half of the senders are randomly assigned to receivers who make a political donation decision. The other half are assigned to receivers who take no further action. The experiment concludes with a questionnaire on motivations, credibility assessments, and socio-demographics.

The experiment operationalizes each element of the conceptual framework as follows. Senders’ policy preferences are measured through an incentive-compatible donation split between a pro-immigration and an anti-immigration advocacy organization. I operationalize the binary state structure across four immigration statistics, each framed as an invertible comparative statement. Two are pro-immigration (immigrant crime rates and labor force participation) and two are anti-immigration (unauthorized border crossings and net fiscal costs at the local level). The receiver’s prior is captured by the receiver’s

<sup>10</sup>This criterion is related to [Blackwell \(1953\)](#): an information structure that brings beliefs closer to the truth weakly dominates one that does not, in the sense that it yields weakly better expected payoffs in every decision problem.

<sup>11</sup>See [Haaland and Roth \(2020\)](#), [Grigorieff et al. \(2020\)](#), [Lergetporer et al. \(2021\)](#), [Alesina et al. \(2023\)](#), [Facchini et al. \(2022\)](#), [Haaland et al. \(2023\)](#)

<sup>12</sup>See [Alesina et al. \(2023\)](#), [Haaland and Roth \(2020\)](#), [Grigorieff et al. \(2020\)](#).

incentivized belief  $b$ . Receivers either hold a false belief or “don’t know” which the sender observes before deciding. The sender’s binary disclosure decision  $d$  maps directly onto the provision choice for each statistic. Crucially, senders know that receivers are unaware that a disclosure decision has been made. When information is provided, it appears as part of the survey flow, and when it is withheld, receivers have no way of knowing that additional information was available. This design feature implements the key assumption that non-disclosure carries no inferential content, blocking the unraveling logic of standard disclosure models. Finally, the sender’s belief about the disclosure-induced preference shift  $\delta$  is measured through an elicitation in which senders predict how receiving each statistic would shift the attitudes of hypothetical respondents.

My identification strategy varies two dimensions within-subject: alignment between sender and statistic as well as receiver belief state. Additionally, I vary two dimensions between-subject: political action and information about the effect of information. The primary within-subject dimension uses variation in alignment: each sender faces two aligned and two unaligned statistics, providing a test of the alignment-based selective disclosure implied by Proposition 1. The primary between-subject dimension varies whether the receiver subsequently takes a consequential political action ( $A$ ): in the ACTION condition, receivers make the same incentivized donation decision as senders; in the no-ACTION condition, receivers complete the experiment without further decisions. This provides exogenous variation to test Proposition 2’s prediction of political-action-induced amplification of alignment-based selectivity.

The secondary between-subject and within-subject dimensions probe whether the magnitude of the beliefs about the preference shift further amplifies alignment-based selectivity, as implied by the additional restriction on  $w$  in Proposition 3.<sup>13</sup> The remainder of this section describes the experimental implementation in detail.

### 3.1.2 Statistics about Immigration

The choice of statistics involves unavoidable discretion. I was systematic along four criteria: veracity, perceived credibility, policy relevance, and balance across policy directions. Both positions must be reflected: the set includes statistics hypothesized to favor increased immigration and statistics hypothesized to favor decreased immigration. I selected statistics from highly authoritative sources (Bureau of Justice Statistics, Bureau of Labor Statistics, Customs and Border Protection, National Academies of Sciences) and validated them in a pre-registered survey ( $N = 100$ ) conducted ahead of the main experiment, which tested a larger pool of candidate statistics against these criteria.<sup>14</sup>

I framed all statistics as binary comparisons between immigrants and U.S.-born Americans or between recent and historical averages using invertible statements (e.g., “higher” versus “lower”). This design has two advantages. First, comparative framing avoids unanchored numerical estimates and difficult cognitive processing; the statements are easy to understand. Second, invertibility ensures that beliefs are either correct or incorrect rather than varying in degree. Finally, it yields a clean mapping to the binary state in my framework. Based on my validation survey and the existing literature, I classify the crime and labor statistics as pro-immigration and the border and fiscal statistics as anti-immigration. I verify this classification empirically using senders’ expected attitude shifts.<sup>15</sup> In the experiment, I provide the four statistics as these statements:

1. *Crime*: The crime rate of immigrants is *lower* than that of U.S.-born Americans.
2. *Labor*: The labor force participation of immigrants is *higher* than that of U.S.-born Americans.
3. *Border*: The average number of unauthorized border crossings is *higher* in the last five years compared to historical averages.
4. *Fiscal*: The net fiscal costs of first-generation immigrants at the state and local level are *higher* than those of U.S.-born Americans.

### 3.1.3 Receiver Belief States

Each of the four statistics is assigned to exactly one receiver who was previously asked about that statistic and either holds a false belief (completed the comparative statement incorrectly) or reports that they do

<sup>13</sup>As I document in Section 4.2, the between-subject manipulation shifted beliefs asymmetrically, limiting its power as an identification strategy. I supplement with within-sender variation in beliefs about the magnitude of the preference shift as correlational evidence.

<sup>14</sup>Appendix B.1 provides a detailed description of the validation survey and the full set of candidate statistics.

<sup>15</sup>See Section 4.2.

not know. The sender observes which case applies before making the provision decision. Assignment of receiver-statistic pairs is randomized across senders.

These two receiver types correspond to different regions of the prior in the framework. In both cases, disclosure can shift beliefs toward the true state and thereby policy preference. I hypothesize the magnitude of the preference shift is larger for the false-belief type. If senders believe such a difference exists, the additional restriction on  $w$  in Proposition 3 implies larger alignment-based selectivity for false-belief receivers. I examine this as a pre-registered heterogeneity analysis.<sup>16</sup>

The central design feature is that senders know that receivers are unaware a disclosure decision has been made. When a sender provides information, it appears to the receiver as a natural element of the survey flow. When a sender withholds information, the receiver cannot distinguish “no relevant information exists” from “relevant information was withheld.” This blocks inference from non-disclosure through which selective withholding unravels in standard disclosure models (Grossman 1981, Milgrom 1981). It also provides a closer approximation to information environments like social media feeds, messaging apps, and editorial choices, where receivers typically cannot identify what information was available but not shared.

### 3.1.4 Beliefs about Statistics

I first elicit senders’ prior knowledge about the four statistics. Senders complete each comparative statement by choosing “higher,” “lower,” or “don’t know,” in random order. Responses are incentivized: correct completions receive a bonus, incorrect completions do not, and “don’t know” responses receive the bonus with 25% probability.<sup>17</sup> After completing all four statements, senders are informed of the correct answer for each statistic. However, senders who already knew a statistic was true may hold it with greater conviction. I interpret prior belief accuracy as modulating the effective veracity motive and control for it throughout.

### 3.1.5 Policy Preferences

I measure senders’ policy preferences using an incentive-compatible revealed preference task. Each sender decides how to allocate \$500 between two advocacy organizations: the *American Immigration Council* (AIC), which advocates for increased immigration, and *NumbersUSA*, which advocates for decreased immigration. Both are well-established organizations representing opposing positions in the immigration debate. I briefly describe the stated goals of each organization and present their missions in their own words, with the display order randomized. Participants are informed that one donation decision from the experiment will be randomly selected and implemented, and they receive a verification link to confirm the donation was placed. This approach obtains a continuous measure of policy preference, and therefore of alignment, that captures both the direction and the intensity of each sender’s position.

### 3.1.6 Beliefs about the Effect of Providing Information

I elicit senders’ beliefs about the effect of information on political attitudes and behavior. Senders are informed about a study by Haaland and Roth (2020) that analyzes the effect of information about labor market outcomes on attitudes and behavior towards immigration policy. I then ask participants, with incentive-compatible payments, to estimate the percentage difference in signatures on a White House petition in favor of increased immigration between the control and treatment groups, and to guess whether the attitudes of participants who initially (falsely) believed that immigration had harmed U.S. workers changed more or less than those of participants who initially believed that immigration had not harmed U.S. workers, or whether both groups changed by the same amount.<sup>18</sup>

Afterwards, I measure the effect of the information treatment on senders’ expected impact of providing specific statistics to receivers. For each of the four immigration statistics, I ask each sender to predict how 10 hypothetical participants would respond after receiving that statistic. Specifically, senders must allocate these 10 hypothetical participants among three categories: those who become more supportive of

---

<sup>16</sup>See Section 4.4.2.

<sup>17</sup>The 25% payoff for “don’t know” is strictly dominated under expected utility maximization with binary alternatives, where uninformed guessing yields a 50% expected bonus. The option may nonetheless be selected by highly ambiguity-averse participants who prefer an objective 25% lottery over an ambiguous 50% guess. This implies that reported “don’t know” rates likely understate genuine ignorance among expected-utility maximizers.

<sup>18</sup>Specifically, I present the Mariel boatlift study, in which half of participants learned that a wave of Cuban immigration had virtually no effect on wages or employment for U.S.-born workers. I provide a brief summary of the study design in the experiment. Details of the implementation can be found in Appendix B.2.

immigration, those who become less supportive of immigration, and those whose opinions do not change after receiving the information. I do this for each statistic and for receivers who either did not know about the statistic or held false beliefs about it.

## 3.2 Treatments and Hypotheses

### 3.2.1 Treatments

The experiment varies four dimensions, two within subjects and two between subjects. Table 1 summarizes the design.

Table 1: Experimental design and identifying variation

Dimension	Variation	Levels	Identification
Alignment with statistic	Within-subject	<i>Aligned</i> and <i>Unaligned</i>	Proposition 1
Political action (ACTION)	Between-subject	ACTION = 1 and ACTION = 0	Proposition 2
Effect of information (INFO)	Between-subject	INFO = 1 and INFO = 0	Proposition 3
Receiver belief state	Within-subject	<i>False belief</i> and <i>Don't know</i>	Proposition 3

Each sender makes four provision decisions: one per statistic directed at four distinct receivers. By construction, two statistics are aligned and two are unaligned with the sender’s revealed policy preference, and each statistic is paired with a receiver who either holds a false belief or reports not knowing.<sup>19</sup> These within-sender variations drive identification as they exploit variation across a sender’s four decisions while holding all sender-level heterogeneity constant. The between-sender dimensions provide exogenous experimental variation in the returns to selective disclosure.

Senders are randomly assigned to the no-ACTION condition or the ACTION condition. Senders know in which condition they are before making their provision decisions. In the no-ACTION condition, receivers complete the experiment without taking any further action after potentially receiving information. In the ACTION condition, receivers make the same incentive-compatible donation decision as senders, splitting \$500 between the pro-immigration and anti-immigration advocacy organizations.

This condition provides exogenous variation in the key moderator of Proposition 2. Randomization holds the veracity motive  $\theta_V$ , the projective motive  $\theta_P$ , and the distribution of the beliefs about effect of providing information constant across conditions. Any difference in alignment-based selectivity between ACTION and no-ACTION is therefore consistent with  $\theta_A > 0$ , under the maintained assumption that the ACTION treatment operates only through the downstream political consequence.<sup>20</sup>

Senders are randomly assigned to the no-INFO condition or the INFO condition. In the no-INFO condition, senders receive no feedback after making their predictions about the Haaland and Roth (2020) study. In the INFO condition, senders learn the actual results, a 69% increase in petition signatures and significantly larger attitude shifts among those with initially incorrect beliefs, presented as a comparison with their own predictions to maximize salience.<sup>21</sup>

I employ a  $2 \times 2$  between-subjects factorial design with  $n = 300$  participants per treatment cell. TREATMENT BASE combines the no-ACTION and no-INFO conditions. TREATMENT A combines the ACTION and no-INFO conditions. TREATMENT I combines the no-ACTION and INFO conditions. TREATMENT A×I combines the action and INFO conditions.

### 3.2.2 Hypotheses

I translate the propositions from Section 2 into four pre-registered hypotheses. The first is a maintained assumption. Hypotheses 1 and 2 correspond directly to the model’s comparative statics (Propositions 1 and 2). Hypothesis 3 probes the additional restriction about the functional form of  $w$  in Proposition 3. Section 3.3 maps each hypothesis to specific coefficients in the pre-registered empirical specifications.

**Hypothesis 0** (Baseline provision). *Provision rates for both aligned and unaligned information exceed the indifference benchmark of 50% in all treatments.*

<sup>19</sup>Strictly, this is only true for senders who reveal a directional policy preference.

<sup>20</sup>I note that ACTION may also shift the salience of the decision and the sender’s construal of the task.

<sup>21</sup>This condition was designed to exogenously shift senders’ beliefs about the persuasiveness, providing evidence on the additional restriction in Proposition 3. As I document in Section 4.2, the manipulation shifted beliefs only for pro-immigration statistics, consistent with limited generalization from the pro-immigration content of the Haaland and Roth (2020) study, which limits its power as an identification strategy.

Hypothesis 0 is a maintained assumption rather than a derived comparative static. It holds if the veracity motive  $\theta_V$  is sufficiently large relative to  $\theta_P$  and  $\theta_A$ .

**Hypothesis 1** (Alignment-based selectivity). *Senders provide aligned information at strictly higher rates than unaligned information in treatment BASE.*

Hypothesis 1 follows from Proposition 1. In BASE, the only source of differential provision is the projective motive  $\theta_P$ , since the receiver takes no political action.

**Hypothesis 2** (Political action amplifies selectivity). *The alignment gap in provision increases when receivers take political actions.*

Hypothesis 2 follows from Proposition 2. Randomization of ACTION holds  $\theta_V$ ,  $\theta_P$ , and the distribution of  $\delta$  constant across conditions. Any widening of the alignment gap under the ACTION condition is consistent with  $\theta_A > 0$ , under the maintained assumption that the treatment operates only through its effect on the political consequentiality of the receiver’s beliefs.

**Hypothesis 3** (Magnitude of preference shift amplifies selectivity). *The alignment gap in provision is increasing in the beliefs about the magnitude of the disclosure-induced preference shift.*

Hypothesis 3 probes the additional restriction in Proposition 3: that the sender’s utility responds not only to the *direction* but also to the *magnitude* of the beliefs about the preference shift ( $|w(\delta)|$  strictly increasing in  $|\delta|$ ). I test it through three approaches. First, I test effect of the between-subject INFO treatment designed to exogenously shift beliefs about the effect of information as pre-registered. Second, variation in receiver belief state (false belief vs. don’t know) provides a pre-registered within-sender test if senders perceive larger persuasive impact for misinformed receivers. Third, within-sender variation in senders’ expected attitude change provides a correlational test. As noted in Section 2.4, rejection is informative. It implies that  $w$  is closer to a sign function than a linear transformation, meaning senders categorize information as aligned or unaligned without calibrating to the expected degree of preference change.

### 3.3 Outcome Measures and Empirical Strategy

I test Hypotheses 0 through 3 at two pre-registered levels of analysis. The sender-level analysis captures aggregate selective behavior and tests whether senders share information in a biased manner without controlling for individual heterogeneity, statistics heterogeneity and prior knowledge. It provides an intuitive summary that I present first in the results. The alignment bias also serves as the dependent variable in the heterogeneity analysis (Section 4.4), where the moderators are sender-level characteristics. The decision-level analysis exploits within-sender variation across the four statistics to determine the drivers of each provision choice. The decision-level specification is my primary pre-registered test.

#### 3.3.1 Sender-Level Analysis: Alignment Bias

Let  $y_{i,j}$  be a binary variable equal to 1 if sender  $i$  provides statistic  $j$  to the receiver and let  $AIC_i \in [0, 500]$  denote sender  $i$ ’s donation to the pro-immigration advocacy group (AIC). I say sender  $i$ ’s policy position is pro-immigration if  $AIC_i > 250$  and anti-immigration if  $AIC_i < 250$ . Senders with  $AIC_i = 250$  reveal no directional preference. They are excluded from the sender-level analysis. Based on my validation survey and the literature, I classify  $\mathcal{J}^+ = \{\text{crime, labor}\}$  as the set of pro-immigration statistics and  $\mathcal{J}^- = \{\text{border, fiscal}\}$  as the set of anti-immigration statistics. The aligned statistic set for sender  $i$  is  $A_i = \mathcal{J}^+$  if pro-immigration and  $A_i = \mathcal{J}^-$  otherwise, with the unaligned set  $U_i$  defined analogously.

I define the sender’s alignment bias as:

$$\text{Alignment Bias}_i = \frac{1}{|A_i|} \sum_{j \in A_i} y_{i,j} - \frac{1}{|U_i|} \sum_{j \in U_i} y_{i,j}. \quad (2)$$

A positive alignment bias indicates selective provision favoring the sender’s own policy position. A zero alignment bias indicates no selectivity, which includes both sharing all information and sharing no information.<sup>22</sup>

<sup>22</sup>At the sender level, I test Hypothesis 0 using a one-sample  $t$ -test of the unaligned provision rate against 0.5; Hypothesis 1 using a one-sample  $t$ -test of  $\text{Alignment Bias}_i > 0$  in BASE; and Hypotheses 2 and 3 using directed two-sample  $t$ -tests comparing  $\text{Alignment Bias}_i$  across the relevant treatment conditions. I Holm-correct tests for Hypotheses 2 and 3.

### 3.3.2 Decision-Level Analysis: Information Provision

Each sender’s four information provision decisions  $y_{i,j}$  allow me to separate a generally high willingness to share true information from selective withholding that depends on political alignment, downstream political action, and the sender’s prior knowledge.

To exploit variation in preference intensity, I define a continuous alignment measure:

$$\text{Align}_{i,j} = \begin{cases} \frac{AIC_i - 250}{250} & \text{if } j \in \mathcal{J}^+ \\ \frac{AIC_i - 250}{250} \cdot (-1) & \text{if } j \in \mathcal{J}^- \end{cases}$$

This measure is bounded between  $-1$  and  $1$ . Positive values indicate that statistic  $j$  is aligned with sender  $i$ ’s policy preference (relative to the midpoint  $AIC_i = 250$ ), negative values indicate misalignment, and the magnitude  $|\text{Align}_{i,j}|$  captures preference intensity. By construction,  $\text{Align}_{i,j} = 0$  for indifferent senders ( $AIC_i = 250$ ), so the measure incorporates all senders.<sup>23</sup>

Let  $\text{Belief}_{i,j}$  denote sender  $i$ ’s prior belief accuracy about statistic  $j$  (coded as 1 if correct, 0.5 if “don’t know,” and 0 if incorrect). In the framework, this modulates the effective veracity motive: senders who already knew a statistic was true may hold it with greater conviction and therefore face a stronger accuracy norm when deciding whether to disclose.<sup>24</sup> I estimate the pre-registered linear probability model:

$$\begin{aligned} y_{i,j} = & \beta_0 + \beta_1 \text{Align}_{i,j} + \beta_2 \text{ACTION}_i + \beta_3 \text{INFO}_i + \beta_4 (\text{Align}_{i,j} \times \text{ACTION}_i) \\ & + \beta_5 (\text{Align}_{i,j} \times \text{INFO}_i) + \beta_6 (\text{ACTION}_i \times \text{INFO}_i) \\ & + \beta_7 (\text{Align}_{i,j} \times \text{ACTION}_i \times \text{INFO}_i) + \gamma_1 \text{Belief}_{i,j} + \gamma_2 \mathbf{S}_j + \gamma_3 \mathbf{X}_i + \varepsilon_{i,j}, \end{aligned} \quad (3)$$

with robust standard errors clustered at the sender level.  $\mathbf{S}_j$  are statistic fixed effects and  $\mathbf{X}_i$  is a vector of demographic and political controls. The coefficients map to the model primitives as follows; this mapping is a model-based interpretation.  $\beta_1$  captures baseline alignment-based selectivity, consistent with the projective motive  $\theta_P$  (Hypothesis 1).  $\beta_4$  captures whether alignment matters more when the receiver’s beliefs translate into a political action, consistent with the instrumental motive  $\theta_A$  (Hypothesis 2).  $\beta_5$  captures whether exogenously shifting beliefs about persuasiveness amplifies alignment-based selectivity (Hypothesis 3).  $\gamma_1$  captures veracity concerns: senders who knew a statistic was true before being informed may provide it at higher rates, reflecting a stronger effective  $\theta_V$ .

## 3.4 Sample, Incentives, and Protocol

I conducted the sender experiment between July 16 and August 12, 2025 on Prolific. The distribution of gender, ethnicity, and political affiliation is representative of the U.S. population according to the 2020 census.<sup>25</sup> This study was approved by the Ethics Council of the Max Planck Society within the framework of the Generalized Approval of Experiments Following the Protocol that is Standard in Experimental Economics (approval no. 2018\_3 – Renewal 2024\_28). It was pre-registered on AsPredicted (No. 238343). In total, 1,257 subjects started the experiment. As pre-registered, I stopped data collection once I had 300 subjects per treatment who passed the attention check,<sup>26</sup> did not switch tabs during the statistics questions, and did not time out. Subjects took a median of 17.26 minutes to complete the survey and earned \$3.32 on average. The receiver data was collected between September 1 and September 10, 2025 on Prolific. In total, 96 receivers took a median of 3 minutes to complete the survey and earned \$0.80 on average.

## 4 Results

In this section, I present the results of my experiments. First, I discuss the subjects’ beliefs about the statistics, their beliefs about the effect of information on behavior, and their policy preferences (section 4.1). Then, I analyze how receiving information about the effect of information (INFO) has shifted beliefs

<sup>23</sup>The pre-registration defined alignment on  $[0, 1]$ . I reparameterize to  $[-1, 1]$  by centering at the indifference point to ease interpretation. This is a one-to-one transformation that leaves all hypothesis tests invariant. Results under the pre-registered coding are therefore identical.

<sup>24</sup>The coding of beliefs was pre-registered.

<sup>25</sup>The sample characteristic and balance across treatment conditions is verified in Table A.1 in Appendix A.

<sup>26</sup>In the attention check senders were asked to indicate that they incorrectly believed that badminton is the most popular sport among immigrants in the United States.

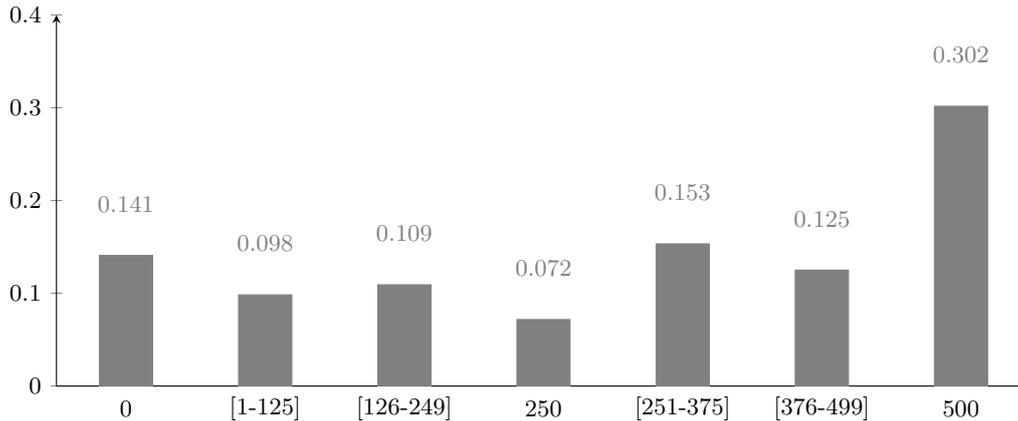


Figure 2: Distribution of senders’ donation allocations to the pro-immigration advocacy group

about effect of providing information (section 4.2). Next, I test how the treatments and alignment affect information provision (section 4.3) and investigate potential heterogeneity of the treatment effects (section 4.4). I conclude by providing correlation evidence on the behavioral drivers of my findings (section 4.5).

#### 4.1 Senders’ Beliefs and Policy Preferences

Before turning to treatment effects, I test the two prerequisites for my hypotheses: (i) senders have policy preferences and (ii) they believe information can shift receivers’ policy preferences. Additionally, I show that senders hold heterogeneous factual beliefs with partisan sorting.

**Senders’ Beliefs about Statistics** I find substantial heterogeneity in senders’ factual knowledge.<sup>27</sup> A majority of senders correctly identify facts about immigrants’ crime rates (70%), labor force participation (79%), and recent border crossings (61%). However, only 30% correctly assess the net fiscal impact of first-generation immigrants at the state and local level, and 14% report “don’t know” for this statistic compared to fewer than 8% for the other three.<sup>28</sup> I find systematic partisan differences: Democrats and immigration supporters demonstrate significantly higher accuracy on pro-immigration statistics (crime and labor), while Republicans and immigration opponents better identify anti-immigration statistics (border and fiscal; all differences  $p < 0.001$ ). This partisan sorting is consistent with motivated beliefs and validates my classification of statistics as pro- or anti-immigration. I control for senders’ prior beliefs throughout to identify senders’ veracity concerns about the chosen statistics.

**Policy Preferences** Figure 2 presents the distribution of senders’ donation decisions. 59% donate the majority of their endowment to the pro-immigration organization, and only 7% split equally. The distribution is markedly bimodal: 30.4% allocate their entire endowment to the pro-immigration group, and 14.1% give everything to the anti-immigration group, so that nearly half of participants hold extreme preferences. Donations correlate strongly with stated policy preferences ( $r = 0.56$ ,  $p < 0.001$ ) and with political affiliation ( $p < 0.001$ ) elicited at the end of the experiment. This polarized distribution provides the necessary variation in alignment to test for selective provision.

**Senders’ Beliefs about the Effect of Information on Political Preferences** Figure A.1 presents senders’ predicted treatment effects from the Haaland and Roth (2020) study. Roughly 90% expect information to increase petition signatures, fewer than 1% expect no effect, and 73% correctly identify that individuals with initially false beliefs update more strongly. Senders thus already believe that information strongly affects policy preferences. They do, however, underestimate the magnitude: the

<sup>27</sup>See Table A.2 for details.

<sup>28</sup>This gap may reflect the conceptual complexity of net fiscal cost, the counterintuitive finding that immigrants generate costs locally but benefits federally, and limited media coverage of technical fiscal analyses compared to more salient topics like crime. Despite lower prior knowledge, the fiscal statistic is rated as credible as the other three statistics ex post (see Table A.3) and is provided at rates comparable to the other anti-immigration statistic, indicating that the knowledge gap does not translate into differential behavior. Additionally, the results are robust to leaving out each statistic individually.

Table 2: Expected effects of information on political preferences

Statistic	Expect Effect of Information on Political Preferences			
	Crime (1)	Labor (2)	Border (3)	Fiscal (4)
<i>Panel (A): INFO</i>				
Constant	0.248*** (0.016)	0.168*** (0.016)	-0.338*** (0.018)	-0.322*** (0.020)
INFO	0.058*** (0.019)	0.038** (0.019)	-0.033 (0.021)	0.031 (0.023)
Estimated Effect Size	0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)
<i>Panel (B): Receiver Belief State</i>				
Constant	0.265*** (0.016)	0.185*** (0.017)	-0.352*** (0.018)	-0.313*** (0.020)
Receiver Belief False	0.020 (0.019)	0.004 (0.019)	-0.003 (0.021)	0.011 (0.023)
Estimated Effect Size	0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)
Clustered SE	No	No	No	No
Controls	No	No	No	No
<i>Panel (A): R<sup>2</sup></i>	0.009	0.018	0.017	0.022
<i>Panel (B) R<sup>2</sup></i>	0.002	0.015	0.015	0.021
<i>Panel (A): Adj. R<sup>2</sup></i>	0.007	0.016	0.015	0.020
<i>Panel (B) Adj. R<sup>2</sup></i>	0.000	0.013	0.013	0.019
Num. obs.	1200	1200	1200	1200

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ . The dependent variable is the sender's expected effect of each statistic on receivers' immigration preferences, measured on the interval  $[-1, 1]$ . Panel (A) estimates the effect of the INFO condition. Panel (B) estimates the effect of receiver false belief relative to "don't know."

true effect is a 69% increase in signatures, while senders predict 38% on average (median 40%). This leaves some room for the INFO condition to shift beliefs upward.

## 4.2 Expected Preference Effects of Information

Alignment-based selectivity requires that senders expect each statistic to shift receivers' preferences in a predictable direction. I now establish that this prerequisite holds and examine whether the INFO condition and receiver belief state moderate these expectations.

For each sender  $i$  and statistic  $j$ , I define the expected attitude change:

$$\Delta\text{Attitude}_{i,j} = \frac{1}{10} \left( n_{\text{more}}^{(i,j)} - n_{\text{less}}^{(i,j)} \right)$$

where  $n_{\text{more}}^{(i,j)}$  denotes the number of hypothetical receivers (out of 10) that sender  $i$  expects to become more supportive of immigration after receiving statistic  $j$ , and  $n_{\text{less}}^{(i,j)}$  denotes those expected to become less supportive. This measure ranges from  $-1$  to  $1$ , with positive values indicating an expected pro-immigration shift and negative values indicating an anti-immigration shift. For each statistic  $j$ , I estimate:

$$\Delta\text{Attitude}_{i,j} = b_0 + b_1 M_{i,j} + b_2 \text{EstimatedEffectSize}_{i,j} + \varepsilon_i \quad (4)$$

where  $M_{i,j}$  is the dimension of interest. I consider the between-subject INFO condition and the within-subject receiver belief state (equal to one if the receiver holds a false belief, zero if the receiver reports not knowing). I control for the sender's baseline belief about information persuasiveness, measured as

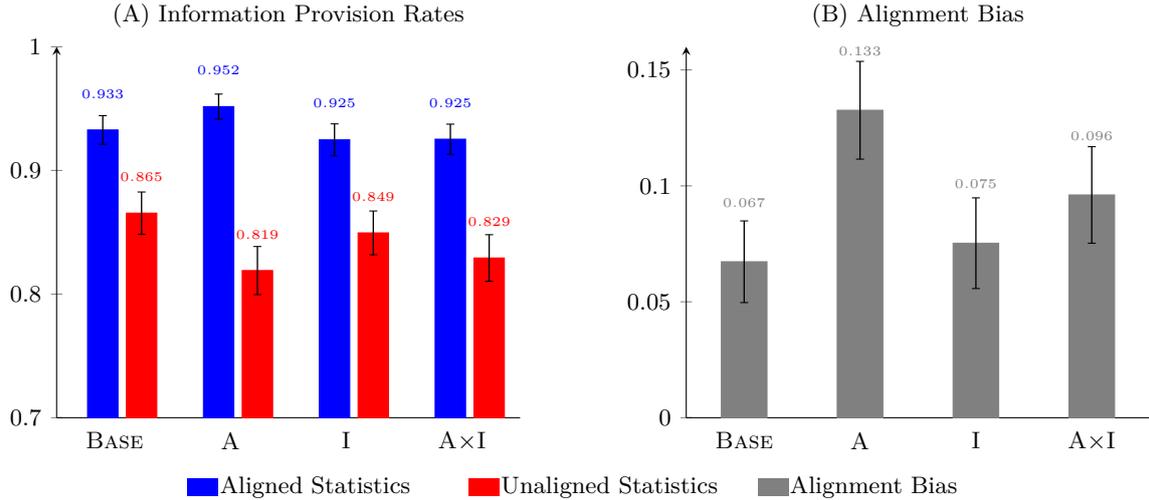


Figure 3: Provision rates for aligned and unaligned statistics by treatment

their incentivized estimate of the treatment effect size in [Haaland and Roth \(2020\)](#). If senders expect the statistic to shift preferences in the hypothesized direction,  $b_0 > 0$  for pro-immigration statistics and  $b_0 < 0$  for anti-immigration statistics. If the INFO or a receiver’s false belief state intensifies these expectations,  $b_1$  has the same sign as  $b_0$ . Table 2 presents the results. Panel (A) reports the effect of the INFO condition. Panel (B) reports the effect of receiver belief state.<sup>29</sup>

Panel (A) shows that senders expect information to shift preferences in the predicted directions. Providing the crime statistic is expected to shift preferences 25 pp toward pro-immigration; labor, 17 pp. Providing the border statistic is expected to shift preferences 34 pp toward anti-immigration; fiscal, 32 pp (all  $p < 0.001$ ). This validates my classification of statistics as pro- and anti-immigration and confirms that senders believe disclosure has the assumed directional effect for receivers’ policy preferences.

However, the INFO condition shifts beliefs about persuasiveness asymmetrically. The INFO condition significantly increases the expected persuasive effect of the pro-immigration statistics: crime ( $b_1 = 0.058$ ,  $p < 0.01$ ) and labor ( $b_1 = 0.038$ ,  $p < 0.05$ ). It does not affect the anti-immigration statistics: border ( $p = 0.114$ ) and fiscal ( $p = 0.185$ ). This asymmetry is consistent with limited generalization: the [Haaland and Roth \(2020\)](#) study demonstrates how pro-immigration information increased support, and senders appear to update about the persuasiveness of similar information without extrapolating to the opposite direction.<sup>30</sup> Overall, the effects are also modest in absolute terms, consistent with the limited room for upward revision discussed above.

Panel (B) shows that senders do not expect larger attitude shifts for receivers with false beliefs compared to uninformed receivers. The coefficients on receiver belief state are small and insignificant for all four statistics. Senders thus treat both receiver types as similarly persuadable, which I revisit in the heterogeneity analysis (Section 4.4.2).

Senders’ prior predictions about the magnitude of the [Haaland and Roth \(2020\)](#) effect only weakly predict statistic-specific attitude change expectations, suggesting that beliefs about the effect of individual statistics are not simply anchored to general beliefs about information effectiveness. Neither the INFO condition nor receiver belief state generates substantial variation in beliefs about the effect of information across the four statistics. Senders appear to hold firm beliefs about the directional impact of each statistic that are only weakly moderated by both experimental manipulation and contextual cues about the receiver. This limits the power of both between-subject and within-subject approaches to test Hypothesis 3 through exogenous variation. I therefore supplement with correlational evidence using within-sender variation in  $\Delta\text{Attitude}_{i,j}$  in Section 4.5.

### 4.3 Treatment Effects on Selective Provision of Information

I now test whether senders selectively provide accurate information and whether political action amplifies this selectivity. I first summarize the aggregate patterns using the sender-level alignment bias (2), then

<sup>29</sup>Table A.4 presents the results with controls.

<sup>30</sup>Notably, coefficient for the fiscal statistic is positive though insignificant.

Table 3: Treatment Effects on Information Provision

	Probability of Statistic Provision		
	(1)	(2)	(3)
Intercept	0.869*** (0.015)	0.851*** (0.017)	0.878*** (0.045)
Align	0.041*** (0.012)	0.036*** (0.012)	0.037*** (0.012)
INFO	-0.013 (0.017)	-0.012 (0.017)	-0.012 (0.017)
ACTION	-0.002 (0.016)	-0.002 (0.016)	-0.003 (0.017)
Belief Accuracy	0.043*** (0.011)	0.034*** (0.012)	0.033*** (0.012)
Align × INFO	0.006 (0.017)	0.004 (0.017)	0.004 (0.017)
Align × ACTION	0.041** (0.018)	0.041** (0.018)	0.041** (0.018)
INFO × ACTION	-0.007 (0.024)	-0.007 (0.024)	-0.008 (0.024)
Align × INFO × ACTION	-0.028 (0.026)	-0.027 (0.026)	-0.027 (0.026)
Clustered SE	Subject	Subject	Subject
Statistics FE	No	Yes	Yes
Controls	No	No	Yes
No. of Senders	1200	1200	1200
R <sup>2</sup>	0.030	0.033	0.040
Adj. R <sup>2</sup>	0.028	0.031	0.034
Num. obs.	4800	4800	4800

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

turn to the primary pre-registered decision-level specification. Figure 3 summarizes senders' provision decisions, restricting the sender-level analysis to the 1,114 senders (93%) who revealed a directional policy preference.<sup>31</sup>

Panel (A) shows that senders overwhelmingly share accurate information across all four treatment conditions. Unaligned provision rates range from 81.9% (A) to 86.5% (BASE), all significantly above 50% ( $p < 0.001$ , one-sample  $t$ -tests).

**Result 0** (Information provision is high). *Senders share accurate information at high rates, including information unaligned with their policy preferences ( $p < 0.001$  against the 50% benchmark in all treatments).*

Panel (B) reveals that provision is selective. The alignment bias is significantly positive in all four conditions ( $p < 0.01$ ), ranging from 6.7 to 13.3 pp. In both pairwise comparisons holding the INFO condition constant, the alignment bias is larger under ACTION.<sup>32</sup> These patterns motivate the more powerful decision-level analysis, which exploits within-sender variation and controls for beliefs about persuasiveness.

### 4.3.1 Decision-Level Analysis

I now turn to the decision-level specification (3). Table 3 reports the results.<sup>33</sup> Across all specifications, provision increases significantly in political alignment. A shift from maximally unaligned to maximally aligned increases provision by 7.2 to 8.2 pp ( $p < 0.003$ ). The magnitude is consistent with the sender-level estimates and is consistent with a positive projective motive  $\theta_P > 0$ . Senders prefer receivers to hold beliefs congruent with their own political position, even when the receiver takes no political action.

**Result 1** (Provision is higher for aligned than unaligned information). *Provision is selectively biased: senders are 7 to 8 pp more likely to share aligned information ( $p < 0.003$ ).*

The interaction  $\text{Align}_{i,j} \times \text{ACTION}_i$  is positive and statistically significant across all specifications ( $p < 0.05$ ). Political action widens the alignment gap by an additional 8.2 pp, approximately doubling the total effect from 7.2–8.2 pp to 15.4–16.4 pp. This is consistent with a positive instrumental motive  $\theta_A > 0$ : senders condition selective disclosure on whether the receiver’s updated beliefs translate into a downstream political action.<sup>34</sup>

A decomposition of the action effect reveals that political action widens the alignment gap primarily by reducing provision of *unaligned* information rather than by increasing provision of aligned information. In the sender-level data, moving from BASE to A reduces unaligned provision by 4.6 pp (from 86.5% to 81.9%) while increasing aligned provision by only 1.9 pp (from 93.3% to 95.2%). This asymmetry reflects, in part, a mechanical ceiling: aligned provision already exceeds 93%, so the margin for further increase is limited.<sup>35</sup>

**Result 2** (Selectivity is higher when receivers take a downstream political action). *Alignment-based selectivity approximately doubles when receivers make a downstream political action ( $p < 0.05$ ). The action effect operates primarily through reduced provision of unaligned information rather than increased provision of aligned information.*

Senders who directly know a statistic to be true provide it at higher rates (3.3 to 4.3 pp,  $p < 0.01$ ). This is consistent with prior knowledge modulating the effective veracity motive. Importantly, this effect is smaller in magnitude than the alignment effect and does not attenuate the  $\text{Align} \times \text{ACTION}$  interaction.<sup>36</sup>

Neither the main effect of INFO nor its interactions with alignment are statistically distinguishable from zero, including in specifications that allow asymmetric effects for the pro-immigration statistics where INFO successfully shifted beliefs about the effect of information on preferences.

**Result 3** (Beliefs about the effect of information do not affect selectivity). *Information about the effect of information does not affect the likelihood of overall information provision nor the selectivity of information provision.*

Combined with the null effect of receiver belief state on expected attitude shifts (Table 2), this pattern suggests that the primary margin is whether receiver beliefs have political consequences, not the magnitude of any individual statistic’s persuasive impact.

## 4.4 Heterogeneous Treatment Effects

The previous section established that alignment-based selectivity exists at baseline and approximately doubles under political action. I now investigate *who* drives each component: the baseline selectivity documented in Result 1 and the action-induced amplification documented in Result 2. I examine heterogeneity by political polarization (Section 4.4.1) and by receiver belief state (Section 4.4.2) as pre-registered.

Table 4: Sender-level Heterogeneous Treatment Effects

Heterogeneity	Alignment Bias					
	<i>Partisanship</i>		<i>Political Party</i>		<i>Issue Polarization</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.005 (0.036)	-0.105 (0.096)	0.132*** (0.032)	0.042 (0.095)	0.008 (0.045)	-0.102 (0.100)
A	0.118** (0.052)	0.120** (0.052)	0.030 (0.046)	0.028 (0.046)	-0.006 (0.067)	0.003 (0.068)
Partisanship	0.140** (0.058)	0.142** (0.059)				
A × Partisanship	-0.102 (0.082)	-0.101 (0.083)				
Independent			-0.132*** (0.048)	-0.130*** (0.048)		
Republican			-0.075 (0.050)	-0.069 (0.050)		
A × Independent			0.146** (0.068)	0.150** (0.069)		
A × Republican			-0.039 (0.069)	-0.025 (0.071)		
Alignment Intensity					0.088 (0.061)	0.104* (0.061)
A × Alignment Intensity					0.088 (0.087)	0.079 (0.088)
Controls	No	Yes	No	Yes	No	Yes
R <sup>2</sup>	0.020	0.058	0.036	0.069	0.027	0.067
Adj. R <sup>2</sup>	0.015	0.026	0.028	0.035	0.022	0.035
Num. obs.	554	554	554	554	554	554

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

#### 4.4.1 Partisans Drive Baseline Selectivity While Independents Drive Action-Induced Amplification

I present this heterogeneity analysis at the sender level, using the alignment bias (2) as the dependent variable. Given the null effect of INFO on all margins (Result 3), I restrict attention to BASE and treatment A.<sup>37</sup> I estimate:

$$\text{Alignment Bias}_i = \alpha_0 + \alpha_1 A_i + \alpha_2 \mathbf{Z}_i + \alpha_3 A_i \mathbf{Z}_i + \gamma \mathbf{X}_i + \varepsilon_i, \quad (5)$$

where  $\mathbf{Z}_i$  is the moderator and  $\mathbf{X}_i$  a vector of demographic and political controls.

Table 4 reports heterogeneous treatment effects on the alignment bias by three measures of polarization: a continuous partisanship index (Columns 1 and 2), political party affiliation (Column 3 and 4), and immigration-specific preference intensity (Column 5 and 6). The baseline alignment bias is driven by politically polarized senders. Non-partisans, Independents, and senders without a directional immigration preference exhibit no alignment bias in BASE. Moving from non-partisanship to full partisanship increases the alignment bias by 14 pp ( $p = 0.017$ ). Democrats exhibit significantly larger alignment bias than Independents (13.2 pp,  $p = 0.006$ ) and Republicans exhibit a positive but insignificant difference (5.7 pp,  $p = 0.268$ , Wald test). Senders with full alignment to the pro- or anti-immigration organization exhibit 9.6 pp alignment bias ( $p < 0.001$ , Wald test).<sup>38</sup> Selective provision in the absence of political action thus appears to be an ideological phenomenon, concentrated among those with strong prior positions.

The action-induced amplification, by contrast, is driven by those with moderate political preferences. The effect of political action on alignment bias is significantly stronger for Independents (14.6 pp,  $p = 0.006$ ) and insignificantly weaker for strong partisans ( $-10.2$  pp,  $p = 0.214$ ). Immigration-specific preference intensity does not significantly moderate the action effect (8.8 pp,  $p = 0.309$ ).<sup>39</sup>

#### 4.4.2 Receiver Belief State Does Not Drive Provision Decisions

Senders face receivers who either hold a false belief or report not knowing. If senders were motivated a desire to correct misinformation specifically akin to epistemic paternalism, I would expect higher provision for misinformed receivers. If senders believe preference shifts are larger for misinformed receivers, Proposition 3 predicts larger alignment-based selectivity for this group.

Table A.9 shows the patterns do not obtain. Senders are not more likely to provide information to receivers with false beliefs ( $-2.5$  pp,  $p = 0.17$ ). Alignment increases provision for both uninformed (5.2 pp,  $p = 0.067$ ) and misinformed receivers (9.0 pp,  $p = 0.011$ ), with no significant difference between the two ( $p = 0.39$ ).<sup>40</sup> The action-induced amplification of selectivity does not differ significantly across receiver types.<sup>41</sup>

<sup>31</sup>Senders with  $AIC_i = 250$  are excluded from the sender-level analysis.

<sup>32</sup>In the no-INFO conditions, which provide the cleanest test of Proposition 2: the alignment bias in BASE is 6.7 pp and nearly doubles to 13.3 pp under A ( $p = 0.011$ ). Under INFO: the bias rises from 7.5 pp (I) to 9.6 pp (A×I;  $p = 0.241$ ). Sender-level regressions confirming these patterns with demographic and political controls are reported in Appendix Table A.5.

<sup>33</sup>Next to the pre-registered specification, in Appendix Table A.7, I also report specifications allowing asymmetric INFO effects by statistic direction to accommodate the asymmetric effect of INFO. This does not affect my results. The additional terms  $\text{INFO}_i \times \text{PROSTAT}_j$  and  $\text{Align}_{i,j} \times \text{INFO}_i \times \text{PROSTAT}_j$  are not statistically significant. Wald tests fail to find evidence that INFO changes the alignment gradient for pro-immigration statistics ( $p = 0.403$ ) or anti-immigration statistics ( $p = 0.606$ ), and fail to reject that INFO changes overall provision for pro-immigration statistics ( $p = 0.390$ ).

<sup>34</sup>Table A.6 shows that the results are robust to logistic regression estimation.

<sup>35</sup>The withholding margin accounts for roughly two-thirds of the total action effect.

<sup>36</sup>I return to veracity as an alternative explanation in Section 4.5.

<sup>37</sup>Appendix Table A.8 pools across the INFO condition; the results are unchanged. Appendix Table A.10 reports the decision-level specification with triple interactions of alignment, ACTION, and each moderator; the conclusions are identical in direction and significance.

<sup>38</sup>The Wald-tested linear combinations for alignment bias by subgroup in BASE are: Independents: 0.0 pp ( $p = 0.986$ ); Democrats: 13.2 pp ( $p = 0.006$ ); Republicans: 5.7 pp ( $p = 0.268$ ); senders with full immigration alignment: 9.6 pp ( $p < 0.001$ ). For the total alignment bias under ACTION by subgroup: Independents: 17.6 pp ( $p = 0.006$ ); Democrats: 16.2 pp ( $p < 0.001$ ); Republicans: 4.8 pp ( $p = 0.647$ ).

<sup>39</sup>This pattern is theoretically interpretable. In the model, the alignment gap  $G(A) = F(\theta_V + c(A)w^+) - F(\theta_V + c(A)w^-)$  is concave in  $c$  for large  $c$  when  $F$  is logistic or normal (standard for discrete-choice models). If partisans have large  $\theta_P$ , so  $c(0)$  is already high and the marginal contribution of  $\theta_A$  to the alignment gap is smaller. If Independents have small  $\theta_P$ , the same  $\theta_A$  generates a larger *change* in the gap. The observation can be explained by Proposition 2 with a concave  $F$ .

<sup>40</sup>These are Wald-tested linear combinations from the full specification in Table A.9, which includes all interactions of alignment, ACTION, INFO, and receiver belief state.

<sup>41</sup>The four-way interaction  $\text{Align} \times \text{ACTION} \times \text{INFO} \times \text{Receiver Belief}$  is not significant ( $p > 0.2$ ). The core findings that alignment-based selectivity and its amplification under political action are robust to conditioning on receiver belief state.

This null is informative. Both receiver types represent opportunities to shift attitudes from an inaccurate to an accurate state. That senders treat them similarly suggests they respond to the *direction* of information, whether a statistic supports or undermines their preferred position, not to the epistemic status of the receiver. This is inconsistent with epistemic paternalism but consistent with my main finding that the primary margin is whether receiver beliefs have political consequences (Result 2), not the magnitude of any individual statistic’s impact on preferences (Result 3).

## 4.5 Mechanism-Consistent Correlational Evidence

The observed selective provision could reflect veracity concerns. Senders share information they find credible and withhold what they doubt. As they only believe information aligns with their preferences veracity concerns generate selective provision. Alternatively, strategic selectivity can arise if senders condition disclosure on whether information advances their political interests. I present correlational evidence further separating these channels. Veracity measures predict provision in levels but do not explain selectivity in the absence of action nor the action-induced amplification of alignment selectivity. Strategic motives, by contrast, fit with the pattern documented in Section 4.3: alignment-based selectivity that intensifies under political action. I begin with veracity (Section 4.5.1), then turn to beliefs about the effect of disclosure on preference (Section 4.5.2), and conclude with self-reported motivations that provide a further test of the strategic channels (Section 4.5.3).

### 4.5.1 Veracity Measures Predict Provision Levels but Not Alignment-Based Selectivity

Senders’ alignment with a statistic predicts both perceived credibility (3.5 pp,  $p < 0.001$ ) and prior belief accuracy (3.6 pp,  $p < 0.001$ ). This pattern is consistent with motivated reasoning in belief formation and raises the concern that alignment-based selectivity could reflect differential veracity rather than strategic motives.

My main specification already addresses the most cleanly identified component of this concern. I always control for incentivized prior belief accuracy that were elicited before senders knew they would make a disclosure decision. This control captures the possibility that alignment bias reflects partisan sorting in factual knowledge: senders may simply happen to know aligned statistics are true. The Align  $\times$  ACTION interaction is robust to this control throughout. Since prior belief accuracy is elicited pre-treatment with incentive-compatible payoffs, it is plausibly exogenous to the provision decision and provides a clean test of the differential-prior-knowledge channel.

I now extend the analysis to include a second veracity measure: unincentivized posterior credibility assessments, collected at the end of the experiment. These capture a distinct channel: motivated credibility discounting, where senders perceive unaligned statistics as less believable. This differential credibility, rather than alignment per se, could drive selective provision. Notably, credibility is measured after the provision decision and may reflect post-hoc rationalization. I do not interpret credibility patterns causally. Together, the two measures cover veracity concerns: prior knowledge captures what senders knew to be true *ex ante* with incentive-compatible elicitation, and credibility captures what they found believable *ex post*. The first is cleanly identified; the second is informative but potentially endogenous.

Columns (1) and (2) of Table 5 present results from specifications that add these veracity measures and their interactions with ACTION to the main model. I find that both Align and the Align  $\times$  ACTION interaction remains significant and stable in magnitude (6 to 6.2 pp,  $p < 0.05$  and 9 to 9.4 pp,  $p < 0.05$ ) across all specifications. Alignment-based selectivity and instrumental amplification of selectivity under political action is not attenuated by accounting for veracity concerns.

In levels, the veracity measures behave as expected. Credibility strongly predicts provision: statistics perceived as credible are 25.6 to 26.9 pp more likely to be shared ( $p < 0.01$ ). Prior belief accuracy also predicts provision (2.5 to 3.2 pp,  $p = 0.06$  and  $p = 0.0346$ ). I also test whether political action induces epistemic caution, i.e., whether senders become more reliant on credibility assessments or prior knowledge when receivers make political decisions. I find no evidence of such moderation. The interactions between political action and credibility ( $-2.2$  pp,  $p = 0.872$ ) and between political action and prior belief accuracy ( $-1.4$  pp,  $p = 0.535$ ) are both null. When receivers make political decisions, senders do not raise their epistemic standards. Still, they become more selective along the alignment dimension. This pattern is not attenuated by controlling for the veracity measures available in this design.

### 4.5.2 Beliefs about Persuasive Impact Weakly Predict Selective Provision

Strategic selectivity predicts that disclosure should depend on senders’ beliefs about how information will shift receivers’ attitudes. I test this using the signed measure of expected impact on preferences  $\text{AlignPersuasion}_{i,j}$ ,

$$\text{AlignPersuasion}_{i,j} = \begin{cases} \Delta\text{Attitude}_{i,j} & , \text{if } AIC_i > 250 \\ (-1) \Delta\text{Attitude}_{i,j} & , \text{if } AIC_i < 250 \\ 0 & , \text{else,} \end{cases}$$

which is positive when sender  $i$  expects statistic  $j$  to move receiver attitudes toward the sender’s preferred position and negative when the expected shift opposes the sender’s position.

Columns (3) and (4) of Table 5 replace the continuous alignment measure with the sender’s belief about the disclosure-induced preference shift. I interact  $\text{AlignPersuasion}$  with  $\text{ACTION}$  but not with  $\text{INFO}$ . The  $\text{INFO}$  treatment was designed to shift beliefs about persuasiveness; since  $\text{AlignPersuasion}$  measures those beliefs, interacting the two would condition on a post-treatment outcome.<sup>42</sup>

Column (3) uses the continuous measure. The main effect of  $\text{AlignPersuasion}$  is small and insignificant (2.1 pp,  $p = 0.19$ ). Absent political action, the magnitude of senders’ beliefs about a statistic’s persuasive impact does not predict whether they share it. The interaction  $\text{AlignPersuasion} \times \text{ACTION}$  is positive and marginally significant (4.2 pp,  $p = 0.079$ ). The point estimate suggests that beliefs about persuasiveness matter more when receivers act, but the effect is imprecisely estimated.

Column (4) replaces the continuous measure with a categorical specification. I define  $\text{AlignPersuasion}(+) = \mathbf{1}\{\text{AlignPersuasion}_{i,j} > 0\}$  and  $\text{AlignPersuasion}(-) = \mathbf{1}\{\text{AlignPersuasion}_{i,j} < 0\}$ , with senders who expect no directional shift as the reference category.<sup>43</sup> The coefficient on  $\text{AlignPersuasion}(+)$  is positive and marginally significant (4.0 pp,  $p < 0.10$ ): senders who believe a statistic will shift the receiver toward their preferred position provide it at somewhat higher rates, even after controlling for credibility and prior belief accuracy. The coefficient on  $\text{AlignPersuasion}(-)$  is essentially zero (0.6 pp,  $p = 0.81$ ): expecting a statistic to shift preferences against the sender’s position does not reduce provision relative to neutral. Neither dummy interacts significantly with  $\text{ACTION}$  ( $p > 0.5$ ), indicating that the marginally significant continuous interaction in Column (3) does not survive a less parametric decomposition.

The contrast between Columns (1) to (2) and Columns (3) to (4) is informative. The  $\text{Align} \times \text{ACTION}$  interaction is robust and significant across all veracity controls ( $p < 0.05$ , Columns 1–2). The  $\text{AlignPersuasion} \times \text{ACTION}$  interaction is at best marginal ( $p = 0.079$ , Column 3) and does not survive the categorical specification (Column 4). This is consistent with the instrumental motive operating through the sender’s *policy alignment*, captured by the incentive-compatible donation measure, rather than through fine-grained beliefs about a statistic’s persuasive impact. Senders condition disclosure on whether the receiver’s beliefs have political consequences, not on how much any particular statistic is expected to shift preferences. In terms of the theoretical framework, this pattern is consistent with the projective and instrumental motives responding to the direction of the preference shift rather than to its magnitude.

### 4.5.3 Self-Reported Motivations Align with Hypothesized Motivations

I classify participants’ open-ended responses about their decision-making into theory-derived motive categories using LLM-based coding with Open AI’s GPT procedure validated against Anthropic’s Claude.<sup>44</sup> Among respondents who referenced the sharing decision, veracity is the most frequently cited primary motive (61%), followed by projection (12%) and instrumental motives (5%).<sup>45</sup>

Table 6 shows that these self-reports predict observed behavior in a pattern that maps directly onto the model. Senders whose primary motive is classified as veracity exhibit 9.2 to 9.5 pp less alignment bias ( $p < 0.01$ ), consistent with  $\theta_V$  attenuating selectivity. Senders citing projection exhibit 21.6 to 21.9 pp more alignment bias ( $p < 0.01$ ), consistent with  $\theta_P$  driving baseline selectivity.

Self-reported instrumental motivation predicts alignment bias *only* when receivers can take political action: the main effect is insignificant ( $-7.6$  pp,  $p = 0.65$ ), while the interaction with  $\text{ACTION}$  is large and

<sup>42</sup>The  $\text{INFO}$  main effect and  $\text{INFO} \times \text{ACTION}$  enter as controls. Including the full three-way  $\text{AlignPersuasion} \times \text{INFO} \times \text{ACTION}$  produces qualitatively similar but noisier estimates.

<sup>43</sup>Figure A.2 in the Appendix shows the raw provision rates by bin of  $\text{AlignPersuasion}$ .

<sup>44</sup>For randomly selected examples of coding, see Table B.18 in the Appendix.

<sup>45</sup>About half of participants do not mention a motive related to the disclosure decision but rather to other parts of the experiment.

Table 5: Mechanisms

Mechanism Evidence	Probability of Statistic Provision			
	<i>Veracity</i>		<i>Persuasiveness</i>	
	(1)	(2)	(3)	(4)
Intercept	0.600*** (0.062)	0.592*** (0.091)	0.581*** (0.062)	0.566*** (0.065)
Align	0.030** (0.013)	0.031** (0.012)		
INFO	-0.014 (0.017)	-0.014 (0.017)	-0.014 (0.017)	-0.014 (0.017)
ACTION	0.010 (0.023)	0.023 (0.122)	0.001 (0.016)	0.012 (0.033)
Belief Accuracy	0.032* (0.017)	0.025** (0.012)	0.039*** (0.012)	0.038*** (0.012)
Credibility	0.256*** (0.060)	0.269*** (0.090)	0.264*** (0.060)	0.260*** (0.060)
Align × INFO	0.007 (0.017)	0.007 (0.017)		
Align × ACTION	0.047** (0.019)	0.045** (0.018)		
INFO × ACTION	-0.008 (0.023)	-0.008 (0.023)	-0.008 (0.023)	-0.008 (0.024)
ACTION × Belief Accuracy	-0.014 (0.022)			
Align × INFO × ACTION	-0.032 (0.025)	-0.032 (0.026)		
ACTION × Credibility		-0.022 (0.122)		
AlignPersuasion			0.021 (0.016)	
AlignPersuasion × ACTION			0.042* (0.024)	
AlignPersuasion(-)				0.006 (0.025)
AlignPersuasion(+)				0.040* (0.024)
AlignPersuasion(-) × ACTION				-0.022 (0.036)
AlignPersuasion(+) × ACTION				-0.004 (0.034)
Clustered SE	Subject	Subject	Subject	Subject
Statistics FE	Yes	Yes	Yes	Yes
Controls	No	No	No	No
No. of Senders	1200	1200	1200	1200
R <sup>2</sup>	0.051	0.051	0.041	0.041
Adj. R <sup>2</sup>	0.048	0.048	0.039	0.039
Num. obs.	4800	4800	4800	4800

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

Table 6: Self-Reported Motivations Effects on Alignment Bias

	Alignment Bias		
	(1)	(2)	(3)
Intercept	0.107*** (0.014)	0.089*** (0.017)	0.090*** (0.017)
<i>Veracity</i>	-0.094*** (0.022)	-0.095*** (0.022)	-0.092*** (0.022)
<i>Projection</i>	0.216*** (0.040)	0.216*** (0.040)	0.219*** (0.040)
<i>Instrumental</i>	0.249*** (0.068)	0.236*** (0.068)	-0.076 (0.168)
<i>Miscellaneous</i>	-0.041 (0.035)	-0.042 (0.035)	-0.041 (0.035)
ACTION		0.037* (0.020)	0.032 (0.020)
<i>Instrumental</i> × ACTION			0.372** (0.183)
Receiver Action	No	Yes	Yes
R <sup>2</sup>	0.059	0.062	0.065
Adj. R <sup>2</sup>	0.056	0.058	0.060
Num. obs.	1113	1113	1113

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

significant (37.2 pp,  $p < 0.05$ ; Column 3).<sup>46</sup> Notably, senders report instrumental motives overwhelmingly in the ACTION condition.<sup>47</sup> This asymmetry is itself consistent with the treatment activating instrumental considerations. The qualitative pattern suggests instrumental motivation predicting alignment bias only when the receiver has a downstream political action. Senders who say they wanted to influence the receiver’s downstream decision appear to bias their disclosure only when the receiver actually has a decision to make. This is the behavioral pattern predicted by  $\theta_A$ : the instrumental motive generates selective disclosure precisely when there is an action to influence.

Descriptive norms further corroborate the mechanism. Senders who more strongly agree that false beliefs should be corrected, that facts matter for opinion formation, or that social media should fact-check provide more information overall and provide it less selectively (Table A.11). Pro-accuracy norms predict *higher levels* of provision but do not eliminate the Align × ACTION interaction ( $\approx 4$  pp,  $p < 0.05$  across specifications). This rules out the alternative that strategic selectivity is simply the absence of accuracy norms: even among senders with strong pro-accuracy attitudes, political action amplifies alignment-based selectivity.

## 5 Discussion

This paper provides causal evidence on the drivers of the selective supply of accurate political information. In a pre-registered experiment, I show that citizens selectively share verified truths based on political alignment: senders are approximately 8 percentage points more likely to provide information aligned with their position than information that contradicts it. This selectivity approximately doubles to 16 percentage points when the receiver’s beliefs translate into a downstream political action. The amplification operates primarily through the withholding of unaligned information. The distortion concentrates in what receivers never see.

Four contributions emerge from these findings. First, I document a supply-side channel through which inaccurate political beliefs persist. The existing literature on politically distorted information has focused on the demand side—biased information seeking, motivated reasoning, and selective exposure (Chopra

<sup>46</sup>The distribution of stated motivations differs significantly between the two ACTION conditions ( $\chi^2(5) = 15.7$ ,  $p = 0.008$ ; see Figure A.3).

<sup>47</sup>Of 25 senders classified as instrumentally motivated, 21 are under ACTION and only 4 under no-ACTION.

et al. 2024, Golman et al. 2022, Peterson and Iyengar 2021). My results establish that the supply side is independently distorted: senders curate which accurate information reaches receivers based on political alignment. Since receivers simultaneously demand congenial information and senders independently supply it, the two forces are likely complementary. A receiver who seeks confirmatory content will probably encounter a sender who preferentially provides it. The resulting information environment is doubly filtered—by what receivers choose to see and by what senders choose to show. My one-shot design cannot identify whether these biases compound across a chain of senders and receivers, but the directional alignment of supply- and demand-side distortions suggests that models incorporating only one margin understate the aggregate bias.

Second, I provide experimental evidence distinguishing a baseline alignment effect from an action-contingent amplification. Under the model’s maintained assumptions, this maps to the projective and instrumental motives respectively. This distinction suggests that curated truth is potentially stronger in settings where it is most consequential—before elections, referenda, and other moments when senders know beliefs translate into collective decisions. Third, citizen-level experimental evidence on selective provision of accurate information complements the observational evidence on politicians’ selective sharing (Freitag et al. 2021) and professional fact-checkers’ selective coverage (Louis-Sidois 2025). Fourth, I apply the logic of projective paternalism (Ambuehl et al. 2021) to a new domain: the selective disclosure of accurate political information. Where Ambuehl et al. show that decision-makers restrict others’ choice sets based on projection, I show that citizens curate the *information environment* on which choices rest. The projection operates upstream of the decision: senders shape what receivers know, not what receivers can do. This extends projection from direct intervention in choices to the curation of the informational inputs to those choices.

**Withholding Asymmetry.** The decomposition of the action effect reveals an asymmetry that is central to the paper’s contribution. Political action widens the alignment gap primarily through reduced provision of unaligned information (−4.6 pp) rather than through increased provision of aligned information (1.9 pp). This partly reflects a mechanical ceiling, aligned provision already exceeds 93%, but the substantive implication extends beyond mechanics. In classical disclosure models (Grossman 1981, Milgrom 1981), non-disclosure is itself informative: receivers who know that a sender could have disclosed infer bad news from silence, generating an unraveling dynamic toward full revelation. In my setting, receivers do not know that a disclosure opportunity existed. Non-disclosure therefore carries no inferential content, and the withholding channel can operate without correction. Under the epistemic benchmark, the potential epistemic cost of curated truth is concentrated in information the receiver never sees and never knows was available—information that, had it been shared, would have moved beliefs weakly toward the truth. This is what distinguishes curated truth from misinformation: misinformation creates false content that can, in principle, be detected, labeled, and corrected. Curated truth preserves inaccurate beliefs by ensuring that the correction never arrives.

**Effect Sizes.** The alignment gap of 8 pp at baseline and 16 pp under political action represents a substantive distortion. These effects obtain despite verified sourcing from federal statistical agencies, no financial incentives to withhold, observation by the experimenter, and explicit labeling of all statistics as accurate. Each of these features biases toward sharing and against detecting selective provision. The alignment effects I estimate are therefore likely conservative along several of these dimensions, though the clean binary framing and explicit knowledge of receiver error are design features that may not translate directly to naturalistic settings. Two features of the setting amplify practical relevance. First, I capture a single binary provision decision per statistic. In naturalistic settings like social media feeds, messaging apps, or editorial selection, individuals face repeated opportunities for selective sharing, and per-decision biases of this magnitude may aggregate across many decisions and many senders. Whether and how such biases compound in a chain of selective intermediaries is an important open question that my one-shot design cannot resolve. Second, the bias is concentrated among senders with the strongest policy preferences—the individuals most likely to engage in political information sharing in the real world.

**Baseline Selectivity Appears More Sensitive to Direction than Magnitude** The combined evidence on Proposition 3 yields a finding that is suggestive though limited in strength by the partial first stage. The INFO treatment successfully shifted senders’ beliefs about the persuasiveness of pro-immigration statistics. Proposition 3 predicts that this increase in perceived persuasiveness should widen alignment-based selectivity for those statistics. However, it does not. The decision-level interaction is essentially zero, and the sender-level effect of TREATMENT I on the alignment bias is exactly zero with

controls. Within-sender variation in beliefs about persuasiveness does not predict provision at baseline. Receiver belief state does not moderate provision decisions. Notably, this is consistent with senders not expecting larger attitude shifts for misinformed relative to uninformed receivers. And the categorical decomposition reveals flat provision rates within each sign category: provision steps up discretely at zero, higher when information is expected to move the receiver toward the sender’s position, lower otherwise, but within each category the magnitude of the expected shift does not matter.

This pattern is consistent with the projective motive combining expressive and directional elements: senders share facts aligned with their political position partly as an act of political expression and partly because they value moving the receiver in their preferred direction, without calibrating to the expected magnitude of the shift. In a fully consequentialist version of projection, utility would scale with the magnitude of the receiver’s expected preference shift, predicting that perceived persuasiveness affects provision continuously even absent political action. The data do not support this. At the same time, senders who expect a statistic to move the receiver toward their preferred position do provide it at somewhat higher rates, indicating that they are not indifferent to the direction of the receiver’s updating. The instrumental motive, by contrast, shows stronger signs of magnitude sensitivity: when political downstream actions are at stake, the degree of expected impact begins to matter. The two motives are thus distinguished not only by what activates them, namely the presence of political action, but by what feature of information they respond to: direction for the projective motive, direction and magnitude for the instrumental motive. These interpretations should be read as suggestive given the partial first stage of the INFO treatment.

**Experimenter Demand and Social Desirability.** Social desirability bias (SDB) is a pervasive concern in survey and experimental research (see [Bursztyn et al. 2025](#)). In my setting, all three channels through which SDB operates, i.e., material costs, social-image concerns, and self-image concerns, likely push toward sharing rather than withholding. Sharing accurate, verified information is the socially desirable response; self-image pressure favors evenhanded disclosure. The SDB apparatus thus works against detecting selective withholding, suggesting that my estimates are likely conservative along this dimension. The more important question is whether SDB can explain *differential* selectivity across treatments. The ACTION condition does not change the social norm: sharing accurate information remains the right thing to do regardless of whether the receiver subsequently acts. If anything, ACTION makes the receiver’s welfare more salient, increasing social pressure to share and biasing against finding larger selectivity. The fact that selectivity approximately doubles under ACTION is difficult to reconcile with differential SDB. Three additional features reinforce this. First, the heterogeneity patterns of Independents driving the action effect and strong partisans driving the baseline are inconsistent with a uniform demand interpretation. Second, 87% of senders found the survey unbiased and only 7% correctly identified the research question.<sup>48</sup> Third, self-reported instrumental motives predict alignment bias only under political action—a pattern that maps onto the model rather than onto social desirability.

**Veracity as an Alternative Explanation.** Politically aligned statistics are perceived as more credible and are more likely to have been previously known. This raises the possibility that alignment-based selectivity reflects differential veracity. Senders simply share what they believe rather than having projective or instrumental motives. I address this through two complementary measures. Incentivized prior belief accuracy, elicited before senders knew they would face a disclosure decision, is included in all specifications. This pre-treatment measure is plausibly exogenous to the provision decision. Posterior credibility assessments, while potentially endogenous, capture a distinct channel: motivated credibility discounting. Credibility strongly predicts provision in levels but interactions remain significant and stable in magnitude across all specifications that include these controls. Critically, political action does not induce epistemic caution. When the receiver’s beliefs carry political consequences, senders do not raise their epistemic standards; instead they become more selective along the alignment dimension.

**Limitations, External Validity and Future Directions.** The experiment studies selective sharing of pre-verified facts in a controlled one-shot environment. I do not speak to equilibrium effects on the information environment, can only identify behavioral consistency with motives from model assumptions, limit the analysis to one policy domain in one country. The results establish a sender-side behavioral pattern under controlled conditions.

---

<sup>48</sup>Two RA’s independently coded all open-ended responses about the guessed research question with Cohen’s  $\kappa = 0.77$  and 82% agreement, indicating substantial inter-rater reliability.

On the other hand, several features of my design are conservative relative to naturalistic information environments, suggesting that my estimates likely understate selectivity along several dimensions, though the clean binary framing and explicit knowledge of receiver beliefs are design features that may amplify rather than attenuate the measured gap in other respects. Senders share verified, officially sourced information; in practice, information quality varies and senders exercise selection over which sources to engage with. Senders face no financial, reputation, or social incentives; in online environments, engagement incentives and ideological community norms can amplify selective sharing. Senders make a single binary share-or-withhold decision; real-world curation involves richer margins like framing, emphasis, contextualization, and sequencing that provide additional instruments for selective presentation. Extending the framework to incorporate framing is a natural direction for future work.

An important open question is whether institutional norms such as journalistic codes of conduct, editorial oversight, and professional training compress the projective and instrumental motive documented here. My results establish the behavioral pattern at the citizen level; [Louis-Sidois \(2025\)](#) documents analogous selectivity among professional fact-checkers at the institutional level. If professional information gatekeepers exhibit the same pattern, the supply-side distortion is structural and its interaction with demand-side bias determines the net information environment. If professional norms attenuate it, targeted institutional design becomes a viable policy lever. Extending the analysis to professional gatekeepers is a natural next step. Whether supply-side selectivity and demand-side selective exposure reinforce or partially offset each other has direct implications for regulatory design: interventions that address only one margin may be insufficient if both operate simultaneously.

**Policy Implications.** My analysis points to directions that differ fundamentally from the current regulatory approach. The existing architecture for political information focusing on fact-checking, content labeling, and platform moderation is designed to detect and remove *false* content. My results document a channel not directly targeted by this architecture: the selective provision of *true* content. The convergence of my citizen-level experimental evidence with observational evidence on citizens and politicians’ selective sharing ([Freitag et al. 2021](#), [Shin and Thorson 2017](#)) and professional fact-checkers’ selective coverage ([Louis-Sidois 2025](#)) is consistent with selective provision appearing at multiple points in the information process, among citizens, politicians, and professional fact-checkers, though the extent and mechanisms may differ across contexts.

The key implication is that information quality depends not only on accuracy but on completeness. Completeness is inherently harder to define and enforce than accuracy—accuracy is verifiable against ground truth, while completeness is relative to an open-ended universe of potentially relevant facts. Current interventions optimize for veracity while the dimension I identify, selectivity of coverage, remains unaddressed. My finding that curated truth intensifies under political action ([Proposition 2](#)) suggest that the distortion is potentially largest in settings where information quality matters most: before elections, referenda, and other moments when beliefs translate into collective decisions. Which policies could effectively mitigate the supply-side distortion documented here are open questions that require evidence from institutional and platform settings.

## 6 Concluding Remarks

Inaccurate political beliefs persist not only because people encounter falsehoods, but because truths that would correct them are selectively withheld. This paper develops a theoretical framework that nests three motives for selective disclosure: veracity, projection, and instrumental strategy. It tests the predictions in a pre-registered experiment with 1,200 U.S. adults recruited on Prolific with quota-balancing on age, gender, ethnicity, and party identification. Senders share verified immigration statistics at high rates, but are approximately 8 percentage points more likely to share facts aligned with their political position. This selectivity doubles when the receiver’s beliefs carry political consequences, with the amplification operating primarily through the withholding of unaligned information.

The theoretical framework organizes these results around a clean separation. The projective motive generates baseline selectivity that is both expressive and directional: senders share facts that affirm their political position categorically, without calibrating to expected persuasive magnitude. The instrumental motive doubles selectivity precisely when the receiver’s beliefs feed into a downstream political action. Together, the two motives produce a pattern of selective provision of accurate information that favors the sender’s position, *curated truth*, that is distinct from misinformation, not addressed by interventions focused on false content, and potentially stronger in settings where political stakes are highest.

The information provision literature establishes that accurate statistics of the kind used here shift beliefs toward the truth and policy preferences in the predicted directions (Haaland and Roth 2020, Grigorieff et al. 2020, Lergetporer et al. 2021, Alesina et al. 2023, Facchini et al. 2022). Under the epistemic benchmark, the potential epistemic cost of curated truth is borne by information the receiver never sees. In many information environments receivers observe what is shared but cannot identify what was available and withheld. This renders the standard unraveling logic inoperative and means that selective omission of true information is not addressed by fact-checking, content labeling, and platform moderation in their current form. Whether institutional norms and professional standards can attenuate this distortion, and how supply-side selectivity interacts with the well-documented demand for congenial information, are questions that the framework developed here is designed to address.

## References

- Abramitzky, R., Ager, P., Boustan, L., Cohen, E., and Hansen, C. W. (2024). The effect of immigration on the economy: Lessons from the 1920s border closure. *American Economic Review: Insights*.
- Alesina, A., Miano, A., and Stantcheva, S. (2023). Immigration and redistribution. *The Review of Economic Studies*, 90(1):1–39.
- Ambuehl, S., Bernheim, B. D., and Ockenfels, A. (2021). What motivates paternalism? an experimental study. *American economic review*, 111(3):787–830.
- Andreottola, G. and Louis-Sidois, C. (2025). Why us and not them? a theory of political fact-checking. SSRN Working Paper No. 5122378.
- Barrera, O., Guriev, S., Henry, E., and Zhuravskaya, E. (2020). Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of public economics*, 182:104123.
- Bartling, B., Cappelen, A. W., Hermes, H., Tungodden, B., and Skivenes, M. (2023). Free to fail? paternalistic preferences in the united states. *NHH Dept. of Economics Discussion Paper*, (09).
- Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The quarterly journal of economics*, 117(3):871–915.
- Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5):1652–1678.
- Blackwell, D. (1953). Equivalent comparisons of experiments. *The annals of mathematical statistics*, pages 265–272.
- Bullock, J. G. and Lenz, G. (2019). Partisan bias in surveys. *Annual Review of Political Science*, 22(1):325–342.
- Bursztyn, L., Haaland, I. K., Röver, N., and Roth, C. (2025). The social desirability atlas. Technical report, National Bureau of Economic Research.
- Bursztyn, L., Rao, A., Roth, C., and Yanagizawa-Drott, D. (2023). Opinions as facts. *The Review of Economic Studies*, 90(4):1832–1864.
- Chopra, F., Haaland, I., and Roth, C. (2022). Do people demand fact-checked news? evidence from us democrats. *Journal of Public Economics*, 205:104549.
- Chopra, F., Haaland, I., and Roth, C. (2024). The demand for news: Accuracy concerns versus belief confirmation motives. *The Economic Journal*, 134(661):1806–1834.
- Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of accounting research*, pages 123–145.
- Facchini, G., Margalit, Y., and Nakata, H. (2022). Countering public opposition to immigration: The impact of information campaigns. *European Economic Review*, 141:103959.
- Federal Bureau of Investigation (2025). UCR summary of reported crimes in the nation, 2024. Technical report, Federal Bureau of Investigation, Criminal Justice Information Services (CJIS). PDF report accessed via FBI Crime Data Explorer (CDE).
- Freitag, J., Kerkhof, A., and Münster, J. (2021). Selective sharing of news items and the political position of news outlets. *Information Economics and Policy*, 56:100926.
- Gallup (2024). Smaller majorities say crime in u.s. is serious, increasing. Gallup News.
- Golman, R., Loewenstein, G., Molnar, A., and Saccardo, S. (2022). The demand for, and avoidance of, information. *Management Science*, 68(9):6454–6476.
- Grigorieff, A., Roth, C., and Ubfal, D. (2020). Does information change attitudes toward immigrants? *Demography*, 57:1117–1143.
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of law and Economics*, 24(3):461–483.
- Grossmann, M. R. P. (2024). Knowledge and freedom: Evidence on the relationship between information and paternalism.

- Haaland, I. and Roth, C. (2020). Labor market concerns and support for immigration. *Journal of Public Economics*, 191:104256.
- Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of economic literature*, 61(1):3–40.
- Hagmann, D. and Loewenstein, G. (2017). Persuasion with motivated beliefs. In *Opinion Dynamics & Collective Decisions Workshop*.
- Jin, G. Z., Luca, M., and Martin, D. (2021). Is no news (perceived as) bad news? an experimental investigation of information disclosure. *American Economic Journal: Microeconomics*, 13(2):141–173.
- Jung, W.-O. and Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting research*, pages 146–153.
- Lergetporer, P., Piopiunik, M., and Simon, L. (2021). Does the education level of refugees affect natives’ attitudes? *European Economic Review*, 134:103710.
- Light, M. T., He, J., and Robey, J. P. (2020). Comparing crime rates between undocumented immigrants, legal immigrants, and native-born us citizens in texas. *Proceedings of the National Academy of Sciences*, 117(51):32340–32347.
- Louis-Sidois, C. (2025). Both judge and party? investigating the political unbiasedness of fact-checkers. *Journal of the European Economic Association*, 23(6):2137–2164.
- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.
- National Academies of Sciences, Engineering, and Medicine (2017). *The Economic and Fiscal Consequences of Immigration*. The National Academies Press, Washington, DC.
- Peterson, E. and Iyengar, S. (2021). Partisan gaps in political information and information-seeking behavior: Motivated reasoning or cheerleading? *American Journal of Political Science*, 65(1):133–147.
- Shin, J. and Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2):233–255.
- Skeptic Research Center (2021). How informed are americans about race and policing? Technical Report CUPES-007, Skeptic Research Center. Research report (survey evidence on beliefs about unarmed Black men killed by police in 2019).
- Thaler, M. (2021). The supply of motivated beliefs. *arXiv preprint arXiv:2111.06062*, 385.
- The Washington Post (2026). Fatal force: Police shootings database. The Washington Post. Database of fatal police shootings in the United States (coverage includes Jan. 1, 2015–Dec. 31, 2024).

## A Supplementary Tables and Figures

### A.1 Sample Characteristics

Table A.1: Sample characteristics and treatment balance

	Population (1)	Sample (2)	BASE (3)	A (4)	I (5)	A×I (6)
<i>Gender</i>						
Female	0.509	0.502	0.507	0.490	0.503	0.510
Male	0.491	0.478	0.470	0.477	0.477	0.487
<i>Ethnicity</i>						
White	0.578	0.587	0.577	0.590	0.587	0.593
Black	0.121	0.129	0.137	0.120	0.140	0.120
Hispanic	0.187	0.182	0.180	0.183	0.177	0.187
Asian	0.059	0.058	0.063	0.063	0.057	0.050
Other	0.055	0.044	0.043	0.043	0.040	0.050
<i>Political Affiliation</i>						
Democrat	0.386	0.388	0.377	0.393	0.393	0.387
Independent	0.310	0.308	0.323	0.303	0.303	0.300
Republican	0.304	0.298	0.297	0.297	0.300	0.300

Notes: Population benchmark are based on the 2020 US Census.

### A.2 Additional Summary Statistics

Table A.2: Prior knowledge of the four immigration statistics

	Share of Responses		
	Correct	False	Don't know
Crime	0.700	0.223	0.076
Labor	0.794	0.141	0.065
Border	0.612	0.317	0.071
Fiscal	0.304	0.553	0.143

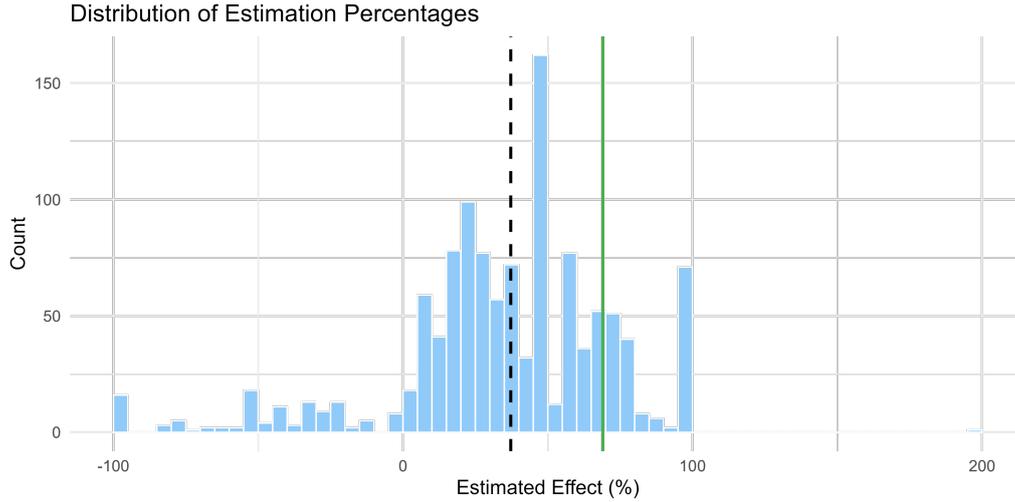


Figure A.1: Senders’ beliefs about the persuasive effect of information (pre-treatment). *Notes:* The green line marks the true effect of 69% from [Haaland and Roth \(2020\)](#).

Table A.3: Perceived Credibility of Statistics by Political Affiliation

	Percentage Perceiving Statistic as Credible			
	All	Democrats	Independents	Republicans
Crime	0.962	0.996	0.949	0.933
Labor	0.978	0.998	0.965	0.964
Border	0.978	0.985	0.976	0.972
Fiscal	0.968	0.981	0.962	0.958

### A.3 Robustness of Main Results

**Logistic Regression.** Table A.6 re-estimates specification (3) using logistic regression, which is the correctly specified model if the latent propensity to disclose follows a logistic distribution—as implied by the cumulative logistic  $F$  assumed in Section 2.2. All results replicate: the  $\text{Align} \times \text{ACTION}$  interaction is positive and significant across all specifications ( $p < 0.05$ ), and the  $\text{Align} \times \text{INFO}$  interaction remains null. I report the linear probability model in the main text for interpretability of coefficients as percentage-point effects.

Table A.4: Expected effects of information on political preferences with controls

Statistic	Expect Effect of Information on Political Preferences			
	Crime (1)	Labor (2)	Border (3)	Fiscal (4)
<i>Panel (A): INFO</i>				
Constant	0.236*** (0.067)	0.033 (0.071)	-0.336*** (0.073)	-0.259*** (0.080)
INFO	0.055*** (0.019)	0.035* (0.019)	-0.034 (0.021)	0.028 (0.023)
Estimated Effect Size	0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Belief Accuracy	0.116*** (0.023)	0.133*** (0.027)	0.003 (0.024)	0.129*** (0.026)
<i>Panel (B): Receiver Belief State</i>				
Constant	0.249*** (0.068)	0.043 (0.071)	-0.347*** (0.074)	-0.255*** (0.080)
Receiver Belief False	0.014 (0.019)	0.006 (0.019)	-0.000 (0.021)	0.018 (0.023)
Estimated Effect Size	0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Belief Accuracy	0.116*** (0.023)	0.134*** (0.027)	0.001 (0.024)	0.129*** (0.026)
Clustered SE	No	No	No	No
Controls	Yes	Yes	Yes	Yes
<i>Panel (A): R<sup>2</sup></i>	0.049	0.054	0.047	0.076
<i>Panel (B): R<sup>2</sup></i>	0.042	0.051	0.044	0.075
<i>Panel (A) Adj. R<sup>2</sup></i>	0.031	0.036	0.029	0.058
<i>Panel (B) Adj. R<sup>2</sup></i>	0.025	0.033	0.027	0.058
Num. obs.	1200	1200	1200	1200

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ 

**Alternative Information Provision Specification.** Table A.7 allows asymmetric INFO effects by statistic direction to accommodate the asymmetric belief-shifting effect documented in Table 2. The additional terms  $\text{INFO}_i \times \text{PROSTAT}_j$  and  $\text{Align}_{i,j} \times \text{INFO}_i \times \text{PROSTAT}_j$  are not statistically significant. The core result— $\text{Align} \times \text{ACTION}$ —is unchanged.

Table A.5: Treatment Effects on Alignment Bias

	Alignment Bias	
	(1)	(2)
Intercept	0.067*** (0.021)	-0.026 (0.071)
<i>Treatment: A</i>	0.065** (0.029)	0.066** (0.029)
<i>Treatment: I</i>	0.008 (0.029)	-0.000 (0.029)
<i>Treatment: A × I</i>	0.029 (0.029)	0.024 (0.029)
Controls	No	Yes
R <sup>2</sup>	0.005	0.048
Adj. R <sup>2</sup>	0.003	0.029
Num. obs.	1114	1114

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ 

## A.4 Heterogeneous Treatment Effects

Table A.8: Sender-level heterogeneous treatment effects, pooling across INFO conditions

Heterogeneity	Alignment Bias					
	<i>Partisanship</i>		<i>Political Party</i>		<i>Issue Polarization</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.003 (0.027)	-0.054 (0.068)	0.114*** (0.023)	0.062 (0.069)	0.008 (0.033)	-0.054 (0.070)
ACTION	0.073* (0.037)	0.070* (0.038)	0.027 (0.033)	0.030 (0.033)	-0.017 (0.049)	-0.018 (0.049)
Partisanship	0.137*** (0.041)	0.139*** (0.041)				
ACTION × Partisanship	-0.055 (0.058)	-0.046 (0.058)				
Independent			-0.086** (0.035)	-0.081** (0.035)		
Republican			-0.052 (0.036)	-0.039 (0.036)		
ACTION × Independent			0.088* (0.049)	0.085* (0.049)		
ACTION × Republican			-0.040 (0.050)	-0.044 (0.050)		
Alignment Intensity					0.093** (0.043)	0.107** (0.043)
ACTION × Alignment Intensity					0.075 (0.062)	0.078 (0.062)
Controls	No	Yes	No	Yes	No	Yes
R <sup>2</sup>	0.017	0.042	0.017	0.038	0.020	0.047
Adj. R <sup>2</sup>	0.015	0.026	0.013	0.021	0.018	0.031
Num. obs.	1114	1114	1114	1114	1114	1114

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ 

## A.5 Mechanism-Consistent Evidence

Table A.6: Treatment Effects on Information Provision Logistic

	Probability of Statistic Provision		
	(1)	(2)	(3)
Intercept	1.967*** (0.148)	1.800*** (0.163)	2.120*** (0.502)
Align	0.460*** (0.133)	0.416*** (0.134)	0.430*** (0.137)
INFO	-0.126 (0.183)	-0.126 (0.184)	-0.126 (0.188)
ACTION	0.173 (0.195)	0.172 (0.196)	0.167 (0.200)
Belief Accuracy	0.399*** (0.103)	0.308*** (0.114)	0.317*** (0.113)
Align × INFO	0.005 (0.177)	-0.009 (0.177)	-0.010 (0.183)
Align × ACTION	0.535** (0.210)	0.538** (0.209)	0.553*** (0.214)
INFO × ACTION	-0.234 (0.259)	-0.233 (0.260)	-0.241 (0.268)
Align × INFO × ACTION	-0.434 (0.270)	-0.425 (0.270)	-0.433 (0.277)
Clustered SE	Subject	Subject	Subject
Statistics FE	No	Yes	Yes
Controls	No	No	Yes
No. of Senders	1200	1200	1200
AIC	3258.054	3247.312	3248.924
BIC	3316.341	3325.029	3449.691
Log Likelihood	-1620.027	-1611.656	-1593.462
Deviance	3240.054	3223.312	3186.924
Num. obs.	4800	4800	4800

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

## B Experimental Design Details

### B.1 Statistic Selection and Validation Survey

I selected statistics for the main experiment from a larger pool of candidates using a pre-registered validation survey ( $N = 100$ , AsPredicted No. 232064) conducted on Prolific between June 10 and June 21, 2025. The survey evaluated eight candidate statistics—four hypothesized as pro-immigration and four as anti-immigration—against three survey-measured criteria (perceived credibility, policy relevance, and perceived political direction) and one researcher-verified criterion (veracity). The selection rule was pre-specified: within each direction, select the two statistics with the highest policy relevance among those with adequate perceived credibility, subject to the constraint that the final set contains two pro-immigration and two anti-immigration statistics.

**Candidate Statistics and Selection Criteria.** Table B.12 reports the eight candidate statistics with their survey-measured scores. The four criteria were operationalized as follows:

*Veracity* was established by the researcher through source verification. Each selected statistic is supported by multiple independent, authoritative sources: peer-reviewed academic publications, federal statistical agencies, or nonpartisan research institutions. Source documentation is provided below.

*Policy relevance* was measured by asking respondents: “How relevant do others consider the statistic when forming immigration policy positions?” (1 = completely irrelevant, 5 = extremely relevant).

*Perceived credibility* was measured by asking respondents: “How credible is the statistic?” (1 = not at all, 5 = entirely).

Table A.7: Treatment Effects on Information Provision

	Probability of Statistic Provision		
	(1)	(2)	(3)
Intercept	0.871*** (0.015)	0.852*** (0.018)	0.879*** (0.046)
Align	0.041*** (0.012)	0.037*** (0.013)	0.037*** (0.013)
INFO	-0.034* (0.019)	-0.018 (0.021)	-0.018 (0.021)
ACTION	-0.002 (0.016)	-0.002 (0.016)	-0.003 (0.017)
Accuracy Belief	0.038*** (0.012)	0.035*** (0.012)	0.034*** (0.012)
Align × INFO	-0.015 (0.021)	-0.009 (0.022)	-0.011 (0.022)
Align × ACTION	0.041** (0.018)	0.041** (0.018)	0.041** (0.018)
INFO × ACTION	-0.007 (0.024)	-0.007 (0.024)	-0.008 (0.024)
Align × INFO × ACTION	0.038*** (0.015)	0.005 (0.020)	0.005 (0.020)
INFO × PROSTAT	-0.027 (0.026)	-0.027 (0.026)	-0.027 (0.026)
Align × INFO × PROSTAT	0.025 (0.023)	0.025 (0.023)	0.029 (0.025)
Clustered SE	Subject	Subject	Subject
Statistics FE	No	Yes	Yes
Controls	No	No	Yes
No. of Senders	1200	1200	1200
R <sup>2</sup>	0.032	0.034	0.041
Adj. R <sup>2</sup>	0.030	0.031	0.034
Num. obs.	4800	4800	4800

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

*Perceived political direction* was measured through two items: *usability* (“For whom would the statistic be most useful in a policy debate?”; 1 = exclusively supporting expansive immigration, 5 = exclusively supporting restrictive immigration) and *opinionability* (“How does learning about the statistic change someone’s immigration policy views?”; 1 = convince to support expansion, 5 = convince to support limitation). These directional measures serve the classification of statistics as pro- or anti-immigration and the balance criterion.

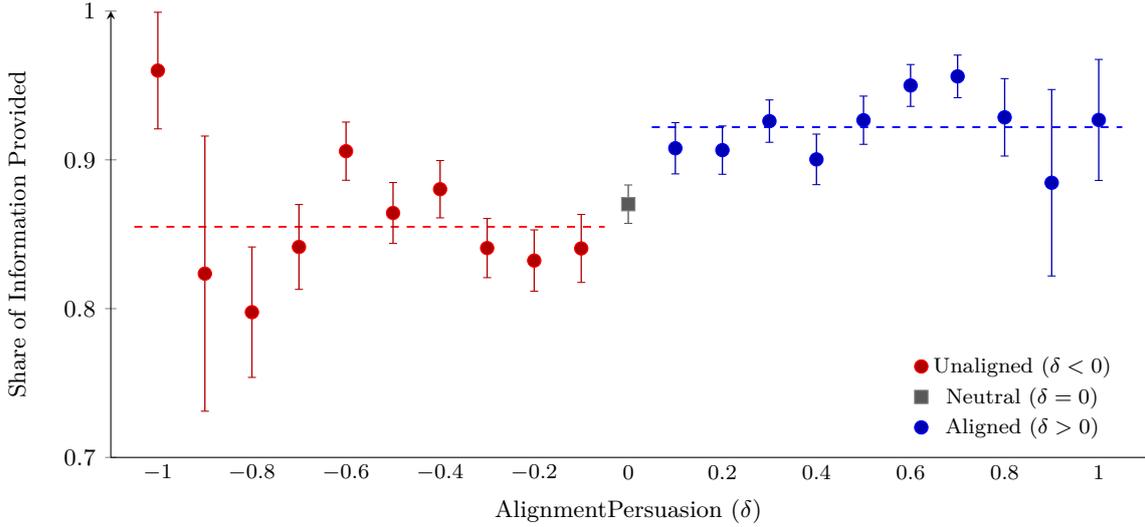


Figure A.2: Provision Rates by Perceived Effect on Receiver’s Preferences

Table B.12: Candidate Statistics: Survey-Measured Indicators

Statistic	Policy	Perceived	Perceived Direction	
	Relevance	Credibility	Usability	Opinionability
<i>Pro-immigration</i>				
Crime Rate <sup>†</sup>	3.74 (1.18)	3.50 (0.97)	2.23 (1.06)	2.36 (1.11)
Labor <sup>†</sup>	3.64 (1.04)	3.81 (1.02)	2.40 (1.09)	2.57 (1.23)
Entrepreneurship	3.12 (1.08)	3.54 (1.00)	2.09 (1.05)	2.18 (1.15)
Healthcare	3.28 (1.03)	3.59 (1.07)	2.03 (1.05)	2.19 (1.15)
<i>Anti-immigration</i>				
Border Crossings <sup>†</sup>	3.85 (1.14)	3.97 (0.90)	3.94 (1.14)	4.08 (1.11)
Fiscal Impact <sup>†</sup>	3.79 (1.01)	3.55 (0.93)	3.68 (1.15)	3.73 (1.01)
Poverty Rate	3.41 (1.01)	3.93 (0.93)	3.40 (1.13)	3.52 (1.11)
Wage Inequality	3.35 (0.95)	3.53 (0.90)	3.41 (1.06)	3.43 (0.99)

Notes:  $N = 100$ . Cells report means with standard deviations in parentheses. Selected statistics with <sup>†</sup>. Policy relevance: 1 = completely irrelevant, 5 = extremely relevant. Perceived credibility: 1 = not at all credible, 5 = entirely credible. Usability and opinionability: 1 = supports expansive immigration, 5 = supports restrictive immigration. One-sample  $t$ -tests against the neutral midpoint of 3 reject the null for all eight statistics on both directional measures ( $p < 0.001$  for all selected statistics; see text).

**Selection.** The selection rule applied within each direction independently. Among the four pro-immigration candidates, Crime Rate (3.74) and Labor Force Participation (3.64) had the highest policy relevance scores, both with adequate perceived credibility ( $\geq 3.50$ ). Among the four anti-immigration candidates, Border Crossings (3.85) and Fiscal Impact (3.79) had the highest policy relevance, again with adequate credibility. This rule uniquely determines the four selected statistics reported in the main text (Section 3.1.2). The one rejected statistic with notably high credibility—Poverty Rate (3.93)—scored lower on policy relevance (3.41) than both selected anti-immigration statistics, confirming that the selection was not driven by credibility alone.

Perceived directional scores did not vary significantly by respondents’ political affiliation (Democrat, Republican, Independent), indicating that the directional classification reflects consensus perception rather than partisan projection.

**Direction Validation.** I classify Crime Rate and Labor Force Participation as pro-immigration and Border Crossings and Fiscal Impact as anti-immigration. This classification is validated by the survey’s directional measures. For usability and opinionability respectively, the pro-immigration statistics score

below the neutral midpoint of 3 (Crime: 2.23, 2.36; Labor: 2.40, 2.57) and the anti-immigration statistics score above it (Border: 3.94, 4.08; Fiscal: 3.68, 3.73). One-sample  $t$ -tests against the midpoint of 3 reject the null of no directional content for all four selected statistics on both measures: Crime ( $t = -7.25$  and  $t = -5.74$ ), Labor ( $t = -5.50$  and  $t = -3.49$ ), Border Crossings ( $t = 8.21$  and  $t = 9.75$ ), and Fiscal Impact ( $t = 5.89$  and  $t = 7.20$ ), all  $p < 0.001$ .

This classification is independently confirmed by the main experiment. In the expected attitude shift elicitation (Section 4.2, Table 2), senders in the main experiment ( $N = 1,200$ ) expect Crime and Labor to shift receiver preferences toward pro-immigration positions (25 and 17 pp, both  $p < 0.001$ ) and Border and Fiscal to shift preferences toward anti-immigration positions ( $-34$  and  $-32$  pp, both  $p < 0.001$ ). The convergence of two independent samples using two independent methods—survey-based directional perception in the validation study and incentivized attitude-shift predictions in the main experiment—makes the directional classification robust.

**Source Documentation.** Veracity was established through source verification rather than survey measurement. Each selected statistic is supported by multiple independent, authoritative sources. I report the specific values shown to participants, the primary source for each, and corroborating evidence. All sources were available at the time of the experiment (July–August 2025).

*Crime Rate* (immigrants have lower crime rates than U.S.-born Americans). Participants were shown violent crime *arrest* rates per 100,000 residents in Texas (2012–2018): 213.0 for U.S.-born citizens, 185.3 for legal immigrants (13% lower), and 96.2 for undocumented immigrants (55% lower). These figures are from Light et al. (2020), who analyze the universe of felony arrests recorded in the Texas Computerized Criminal History database cross-referenced with Department of Homeland Security immigration status checks.<sup>49</sup> The finding is independently corroborated by ?, who uses Texas Department of Public Safety conviction data over a longer period (2013–2022) and reports homicide conviction rates of 3.0 per 100,000 for native-born Americans, 2.2 for unauthorized immigrants (26% lower), and 1.2 for legal immigrants (61% lower).<sup>50</sup> At the national level, Abramitzky et al. (2024) analyze census-linked incarceration records from 1870–2020 and find that immigrants have been consistently less likely to be incarcerated than the U.S.-born population throughout this period.<sup>51</sup> Texas is the only U.S. state that systematically records the immigration status of all arrestees, making it the sole source of population-level crime-rate comparisons by immigration status. The national-level incarceration evidence in Abramitzky et al. (2024) suggests the Texas pattern is not anomalous.

*Labor Force Participation* (immigrants have higher labor force participation than U.S.-born Americans). Participants were shown 2023 annual average labor force participation rates: 66.6% for the foreign-born population and 61.8% for the native-born population (4.8 percentage points higher). These figures are from Table 1 of the Bureau of Labor Statistics annual report *Foreign-Born Workers: Labor Force Characteristics—2023*.<sup>52</sup> The same data are available through the Federal Reserve Bank of St. Louis FRED database, Current Population Survey Table A-7.<sup>53</sup> The foreign-born participation rate has exceeded the native-born rate in every year of the BLS series.

*Border Crossings* (unauthorized border crossings have been higher in the past five years relative to historical averages). Participants were shown two figures computed from U.S. Customs and Border Protection (CBP) enforcement statistics:<sup>54</sup> a historical average of 890,435 apprehensions per year over fiscal years 1970–2019, and a recent five-year average of 1,584,573 encounters per year over fiscal years 2020–2024 (78% higher). The comparison requires combining two CBP reporting series: “apprehensions” (reported through FY2019) and “encounters” (reported from FY2020 onward, which combine apprehensions and expulsions under Title 42 and Title 8 processing).<sup>55</sup> The statistic reflects data available at experiment time (July–August 2025). Subsequent enforcement changes under the second Trump administration reduced encounters to record lows in FY2025, as documented by Pew Research Center analysis of

<sup>49</sup>Published in *Proceedings of the National Academy of Sciences*; data and replication materials available at <https://www.openicpsr.org/openicpsr/project/124923>.

<sup>50</sup>Cato Institute Policy Analysis no. 977; available at <https://www.cato.org/policy-analysis/illegal-immigrant-murderers-texas-2013-2022>.

<sup>51</sup>Published in *American Economic Review: Insights*.

<sup>52</sup>The annual report is available at <https://www.bls.gov/news.release/pdf/forbrn.pdf>. At the time of writing, the 2024 edition (released May 2025) reports 66.5% versus 61.7%, confirming the same pattern.

<sup>53</sup><https://fred.stlouisfed.org/release/tables?rid=50&eid=1736>

<sup>54</sup>CBP historical data available at <https://www.cbp.gov/newsroom/stats/cbp-enforcement-statistics>.

<sup>55</sup>Beginning in March 2020, CBP shifted from reporting apprehensions to reporting “encounters,” a broader category. The historical series (1970–2019) uses the apprehension metric; the recent series (2020–2024) uses the encounter metric. Since encounters  $\geq$  apprehensions by construction, the 78% comparison is an upper bound—i.e., it may slightly overstate the increase by comparing a narrower historical measure to a broader recent measure.

CBP data.<sup>56</sup> The FY2025 decline does not affect the validity of the statistic as presented to participants, which concerned the five-year average through FY2024.

*Fiscal Impact* (net fiscal costs of first-generation immigrants are higher than U.S.-born Americans at the state and local level). Participants were shown that first-generation immigrant households generate \$57.4 billion more in annual state and local government costs than they contribute in taxes, based on 2011–2013 data. This figure is from the National Academies of Sciences, Engineering, and Medicine comprehensive report on the economic and fiscal consequences of immigration, which remains the most authoritative and comprehensive fiscal analysis of immigration in the United States.<sup>57</sup> The report notes that this fiscal gap reverses at the federal level and across generations: the children and grandchildren of immigrants are among the strongest net fiscal contributors. I present only the state-and-local first-generation finding because it represents the anti-immigration direction in my balanced design. The Cato Institute provides complementary analysis of immigrants’ fiscal effects over a longer period (1994–2023), confirming that the fiscal picture varies substantially by level of government and generational horizon.<sup>58</sup>

## B.2 Information Condition Implementation

Part 2 of the experiment proceeds in three stages: study presentation and elicitation (all senders), information feedback (INFO = 1 only), and expected attitude change elicitation (all senders). I describe each stage and reproduce the text shown to participants.

**Stage 1: Study Presentation and Elicitation (All Senders).** All senders read the following description of the [Haaland and Roth \(2020\)](#) study:

A recent study published in the prestigious academic journal, *Journal of Public Economics*, examined how people form political opinions.

The researchers surveyed more than 3,000 Americans that are representative of the US population.

People were told about a 1980 event when 55,000 Cuban immigrants unexpectedly arrived in Miami, increasing the city’s workforce by 8% almost overnight. They were asked how this affected the wages and employment of U.S.-born workers.

Then, half of the participants learned that research found this wave of immigration had virtually no effect on the wages or unemployment for U.S.-born workers, because the new immigrants increased demand for goods and services, which created more jobs.

The other half of the participants received no information.

Senders then make two incentivized predictions:

1. *Behavioral effect (continuous)*: “By what percentage did learning about the research findings change people’s petition signing rates?” Senders enter a percentage (positive or negative).
2. *Differential updating (categorical)*: “After learning about the research findings, who do you think changed their attitudes toward immigration more?” Senders choose among: (a) people who initially believed immigration harmed U.S.-born workers, (b) people who initially believed immigration did not harm U.S.-born workers, or (c) both groups changed about the same amount.

Each question is equally likely to be selected for the bonus payment. For the continuous estimate, the probability of receiving the bonus is 100% if the estimate equals the true value (69%) and decreases linearly to 0% at a deviation of 20 percentage points or more.<sup>59</sup> For the categorical estimate, the probability is 100% if correct and 0% otherwise. Senders are instructed not to search for the answers and tab-switching is monitored to enforce compliance; senders who switched tabs during this stage were excluded as pre-registered.

<sup>56</sup><https://www.pewresearch.org/short-reads/2026/02/02/migrant-encounters-at-the-us-mexico-border-are-at-their-lowest-level>

<sup>57</sup>National Academies of Sciences, Engineering, and Medicine (2017); the state and local fiscal impact analysis is in Chapter 8. Available at <https://www.nationalacademies.org/read/23550/chapter/14>.

<sup>58</sup><https://www.cato.org/white-paper/immigrants-recent-effects-government-budgets-1994-2023>

<sup>59</sup>This implements a piecewise-linear proper scoring rule that is incentive-compatible under risk neutrality and approximately so under moderate risk aversion.

**Stage 2: Information Feedback (Info = 1 Only).** Senders in the INFO condition receive personalized feedback comparing their predictions to the actual results. Senders in the no-INFO condition proceed directly to Stage 3. The feedback screen displays the following:

**Petition Signatures**

The percentage of people who signed the White House petition actually increased by **69%**.

You estimated that the percentage of people who signed the White House petition changed by *[participant's estimate]*%.

That means *you* *[underestimated/overestimated]* the change by *[participant-specific gap] percentage points*.

**Effect on Attitudes towards Immigration**

Actual Result: People who initially believed immigration harmed U.S.-born workers changed their attitudes towards immigration **significantly more** than those who already held correct beliefs.

Your Answer: *[participant's answer]*.

The feedback is accompanied by a bar chart that displays the sender’s estimate and the true value (69%) side by side, with bar heights scaled proportionally. The sender’s categorical prediction is marked as correct or incorrect with a color-coded indicator. A summary reiterates the two key findings: (i) receiving information strongly changed behavior and (ii) the effect was strongest for participants who initially held false beliefs.<sup>60</sup>

**Stage 3: Expected Attitude Change Elicitation (All Senders).** All senders then predict how receiving each of the four immigration statistics would shift the attitudes of hypothetical receivers. For each statistic  $j$  and each receiver belief state (false belief or “don’t know”), senders allocate 10 hypothetical participants among three categories: those who become *more supportive* of immigration, those who become *less supportive* of immigration, and those whose opinions *do not change*. The allocations must sum to 10. This procedure generates the expected attitude change measure  $\Delta\text{Attitude}_{i,j}$  defined in Section 4.2, which provides within-sender variation in perceived persuasiveness across statistics and receiver types.

You will estimate how ten participants who hold **false beliefs** or **do not know the answer** would change their attitudes about immigration if they **are provided with research findings**.

In particular, you will estimate how many of them...

- become less supportive of immigration.
- don’t change their opinion.
- become more supportive of immigration.

The interface displays 10 person icons to make the allocation task concrete. Senders complete this task eight times: once for each of the four statistics crossed with each of the two receiver belief states. The order of statistics is randomized across senders.

### B.3 Experimental Instructions and Interface

Available upon request.

---

<sup>60</sup>The gap between the participant’s estimate and the true value is computed individually and displayed dynamically in the experimental interface. Senders who estimated the effect exactly see a confirmation message rather than a deviation.

## B.4 Classification Procedure

I classify the 1,200 open-ended responses to the question “What factors or considerations influenced your decisions during the study?” into theory-derived motive categories using large language models. The primary classification uses OpenAI’s GPT-4o (accessed via the OpenAI API, temperature = 0). To assess robustness, I independently replicate the classification using Anthropic’s Claude Sonnet 4.6 (`claude-sonnet-4-6`, temperature = 0) with an identical prompt. Both models receive the same structured prompt containing category definitions, strict decision rules, and explicit boundary conditions.

The classification scheme comprises six categories, summarized in Table B.13. Categories are not mutually exclusive: multiple labels are assigned when a response contains distinct and separable justifications for each. No motives are inferred beyond the text. In addition to binary category indicators, I assign a primary motive corresponding to the most central justification.

The classification proceeds in five steps, enforced through the prompt structure. First, each response is checked for any reference to information disclosure (sharing, withholding, or a receiver). Responses without such reference are classified as *did not refer to disclosure* and receive no further coding. Second, only explicitly stated reasons are extracted. Third, responses are evaluated against category-specific criteria, including explicit boundary rules designed to distinguish closely related motivations. Fourth, multiple categories are assigned only when each is supported by a distinct reason. Fifth, a primary motive is selected.

The theoretically most important boundary is between *projection* and *instrumental*. Projection captures the sender’s preference for the receiver to hold aligned beliefs or attitudes—wanting the receiver to *think* a certain way. The instrumental category captures the sender’s concern for the receiver’s downstream behavior—wanting the receiver to *do* something specific (e.g., donate, vote, support an organization). This distinction maps directly onto the separation of  $\theta_P$  from  $\theta_A$  in the theoretical framework. The prompt enforces this boundary through explicit decision rules: references to influencing what others *think* or *believe* are classified as projection, whereas references to influencing what others *do* are classified as instrumental.

A second important boundary separates *projection* from *did not refer to disclosure*. Projection requires that the respondent’s political beliefs are explicitly linked to the sharing or withholding decision—not merely stated. Responses that express political opinions without connecting them to the disclosure decision (e.g., “My personal beliefs”) are classified as *did not refer to disclosure*. This boundary is enforced through a strict necessary condition in the prompt: projection requires (i) explicit reference to own beliefs, (ii) an explicit link between those beliefs and the sharing decision, and (iii) that the direction of the information relative to those beliefs matters.

The complete classification prompt, including all decision rules, boundary conditions, and worked examples, is available upon request.

## B.5 Validation

To assess the robustness of the classification, I replicate the coding using a second LLM (Claude Sonnet 4.6) with the identical prompt and compare the two independent classifications. This inter-model agreement test checks whether the category definitions and decision rules are sufficiently precise to produce consistent classifications across different model architectures.

**Overall Agreement.** Table B.14 reports Cohen’s  $\kappa$  and disagreement rates for all categories. Inter-model agreement is substantial to near-perfect across the categories central to the analysis.  $\kappa$  exceeds 0.75 for four of the six binary categories: veracity ( $\kappa = 0.887$ ), instrumental ( $\kappa = 0.788$ ), did not refer to disclosure ( $\kappa = 0.771$ ), and disbelief ( $\kappa = 0.708$ ). Agreement is somewhat lower for projection ( $\kappa = 0.670$ ) and miscellaneous ( $\kappa = 0.524$ ). The lower agreement on projection reflects the inherent difficulty of the projection-did-not-refer boundary: responses that state political opinions without clearly linking them to the sharing decision occupy a gray zone that different models resolve differently. The low agreement on miscellaneous is expected given its residual nature; this category enters no regression. For the composite primary motive classification,  $\kappa = 0.755$  with a disagreement rate of 16.3%.

**Precision, Recall, and F1.** Table B.15 reports category-level precision, recall, and F1-score, treating the GPT-4o classification as the positive reference for the purpose of computing directional metrics. For the three categories central to the main analysis—veracity, projection, and instrumental—performance is strong. Veracity achieves F1-scores of 0.96 (negative class) and 0.92 (positive class). Instrumental yields near-perfect agreement for the negative class (F1 = 1.00) and strong agreement for the positive class

(F1 = 0.79), consistent with its low base rate (2.3% of responses). Projection shows somewhat lower recall for positive cases (0.63), reflecting the boundary ambiguity noted above. Miscellaneous shows the weakest positive-class F1 (0.58), expected given its residual nature. This does not affect the main findings, which do not rely on this category.

**Confusion Matrices.** Table B.16 reports confusion matrices for each binary category. Disagreements are rare and largely symmetric for the categories central to the analysis. For veracity, 38 cases coded 0 by GPT-4o receive a 1 from Claude, and 22 cases show the reverse pattern—a small and roughly balanced discrepancy. Instrumental and disbelief show very few disagreements in absolute terms (11 and 19 cases respectively), consistent with their high  $\kappa$  values. The largest absolute disagreements occur in did not refer to disclosure (137 cases) and miscellaneous (122 cases), both of which exhibit higher base rates and broader definitional scope.

Table B.17 reports the confusion matrix for the primary motive classification. The diagonal entries confirm strong agreement for the most common categories: veracity (351 of 383 GPT-coded cases, 91.6%), did not refer to disclosure (492 of 584 cases, 84.2%), and miscellaneous (84 of 116 cases, 72.4%). The most frequent off-diagonal pattern is the confusion between did not refer to disclosure and miscellaneous (72 cases), which is theoretically expected: both categories capture responses that do not engage substantively with the disclosure decision, and their boundary is inherently ambiguous. Confusion between projection and other motivational categories remains limited, and instrumental is rarely misclassified, consistent with the high  $\kappa$  reported above.

## B.6 Coding Examples

Table B.18 presents representative examples of open-ended responses classified into each motive category, illustrating the classification boundaries.

## C Pre-Registration

This study was pre-registered on AsPredicted (No. 238343). The pre-registration specified four directional hypotheses (Hypotheses 0-3), a target sample size of  $N = 1,200$  (300 per treatment cell), and the decision-level linear probability model (3) as the primary specification.

A deviation from the pre-analysis plan is noted: the alignment measure was reparameterized from  $[0, 1]$  to  $[-1, 1]$  by centering at the indifference point ( $AIC_i = 250$ ). This is a one-to-one transformation that leaves all hypothesis tests invariant.

Table A.9: Heterogeneous Treatment Effects of Receiver Beliefs

	Probability of Statistic Provision	
	(1)	(2)
Intercept	0.863*** (0.018)	0.894*** (0.045)
Align	0.026* (0.014)	0.025* (0.014)
INFO	-0.034 (0.021)	-0.034 (0.022)
ACTION	-0.021 (0.020)	-0.022 (0.021)
Receiver Belief False	-0.025 (0.018)	-0.024 (0.018)
Accuracy Prior Belief	0.035*** (0.012)	0.034*** (0.012)
Align $\times$ INFO	0.025 (0.023)	0.025 (0.023)
Align $\times$ ACTION	0.053** (0.024)	0.054** (0.024)
INFO $\times$ ACTION	0.014 (0.030)	0.014 (0.031)
Align $\times$ Receiver Belief False	0.019 (0.022)	0.020 (0.022)
INFO $\times$ Receiver Belief False	0.043 (0.027)	0.043 (0.027)
ACTION $\times$ Receiver Belief False	0.034 (0.026)	0.034 (0.026)
Align $\times$ INFO $\times$ ACTION	-0.050 (0.035)	-0.051 (0.035)
Align $\times$ INFO $\times$ Receiver Belief False	-0.042 (0.032)	-0.042 (0.032)
Align $\times$ ACTION $\times$ Receiver Belief False	-0.020 (0.033)	-0.022 (0.033)
INFO $\times$ ACTION $\times$ Receiver Belief False	-0.041 (0.038)	-0.043 (0.038)
Align $\times$ INFO $\times$ ACTION $\times$ Receiver Belief False	0.047 (0.045)	0.048 (0.045)
Clustered SE	Subject	Subject
Statistics FE	Yes	Yes
Controls	No	Yes
No. of Senders	1200	1200
R <sup>2</sup>	0.034	0.041
Adj. R <sup>2</sup>	0.030	0.033
Num. obs.	4800	4800

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

Table A.10: Individual-level Heterogeneous Treatment Effects

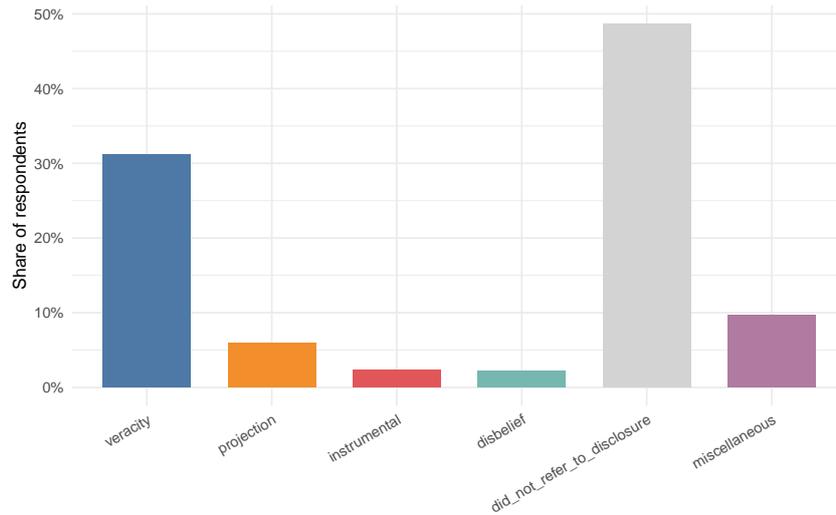
Heterogeneity	Probability of Statistic Provision			
	<i>Partisanship</i>		<i>Political Party</i>	
	(1)	(2)	(3)	(4)
Align	0.001 (0.016)	0.001 (0.016)	0.057*** (0.019)	0.057*** (0.019)
ACTION	0.026 (0.024)	0.026 (0.025)	0.018 (0.023)	0.014 (0.023)
Partisanship	-0.003 (0.026)	-0.001 (0.027)		
Accuracy Prior Belief	0.031** (0.012)	0.031** (0.012)	0.033*** (0.012)	0.034*** (0.012)
Align × ACTION	0.057** (0.027)	0.057** (0.027)	0.019 (0.025)	0.019 (0.025)
Align × Partisanship	0.065*** (0.023)	0.065*** (0.023)		
ACTION × Partisanship	-0.053 (0.036)	-0.054 (0.037)		
Align × ACTION × Partisanship	-0.029 (0.037)	-0.029 (0.037)		
Independent			0.009 (0.021)	0.005 (0.021)
Republican			0.005 (0.021)	0.001 (0.021)
Align × Independent			-0.043** (0.020)	-0.043** (0.020)
Align × Republican			-0.022 (0.023)	-0.022 (0.023)
ACTION × Independent			-0.034 (0.029)	-0.031 (0.029)
ACTION × Republican			-0.032 (0.029)	-0.027 (0.029)
Align × ACTION × Independent			0.061** (0.030)	0.061** (0.030)
Align × ACTION × Republican			0.009 (0.032)	0.009 (0.032)
Clustered SE	Subject	Subject	Subject	Subject
Statistics FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
R <sup>2</sup>	0.037	0.042	0.035	0.040
Adj. R <sup>2</sup>	0.034	0.036	0.032	0.033
Num. obs.	4800	4800	4800	4800

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

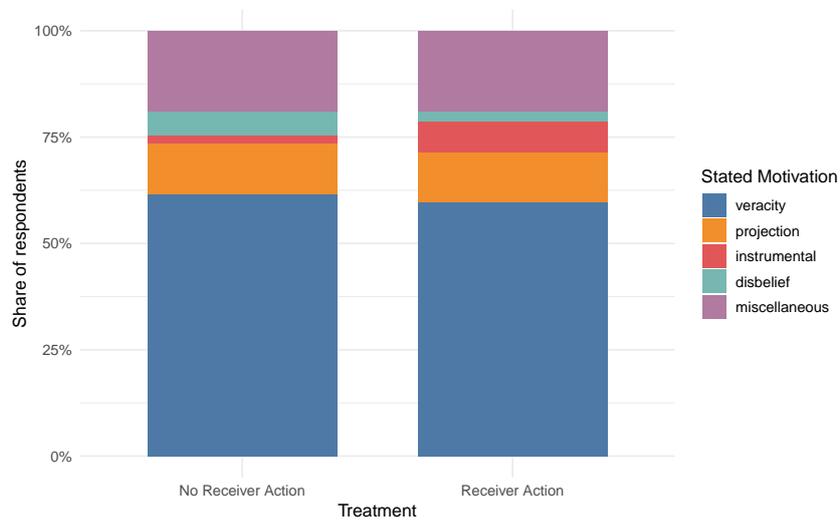
Table A.11: Treatment Effects on Information Provision

Statement	Probability of Statistic Provision			
	<i>False factual beliefs should be corrected</i> (1)	<i>Facts are important for forming opinions</i> (2)	<i>Social media should fact check</i> (3)	<i>Disagreement stems from facts not views</i> (4)
Intercept	0.538*** (0.056)	0.415*** (0.083)	0.734*** (0.033)	0.898*** (0.027)
Align	0.197*** (0.052)	0.213** (0.088)	0.109*** (0.028)	0.007 (0.021)
INFO	-0.009 (0.016)	-0.007 (0.016)	-0.011 (0.017)	-0.014 (0.017)
ACTION	-0.003 (0.016)	-0.005 (0.016)	-0.004 (0.016)	-0.003 (0.016)
Accuracy Belief	0.034*** (0.012)	0.032*** (0.012)	0.035*** (0.012)	0.034*** (0.012)
Statement_Agree	0.069*** (0.011)	0.091*** (0.016)	0.026*** (0.006)	-0.013** (0.006)
Align × INFO	0.002 (0.017)	0.001 (0.017)	0.002 (0.017)	0.004 (0.017)
Align × ACTION	0.039** (0.018)	0.039** (0.018)	0.043** (0.018)	0.041** (0.018)
INFO × ACTION	-0.011 (0.023)	-0.010 (0.023)	-0.010 (0.024)	-0.006 (0.024)
Align × Statement_Agree	-0.036*** (0.011)	-0.036** (0.018)	-0.018*** (0.006)	0.009 (0.006)
Align × INFO × ACTION	-0.024 (0.026)	-0.023 (0.026)	-0.024 (0.026)	-0.027 (0.026)
Clustered SE	Subject	Subject	Subject	Subject
Statistics FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
No. of Senders	1200	1200	1200	1200
R <sup>2</sup>	0.065	0.066	0.044	0.036
Adj. R <sup>2</sup>	0.063	0.063	0.042	0.033
Num. obs.	4800	4800	4800	4800

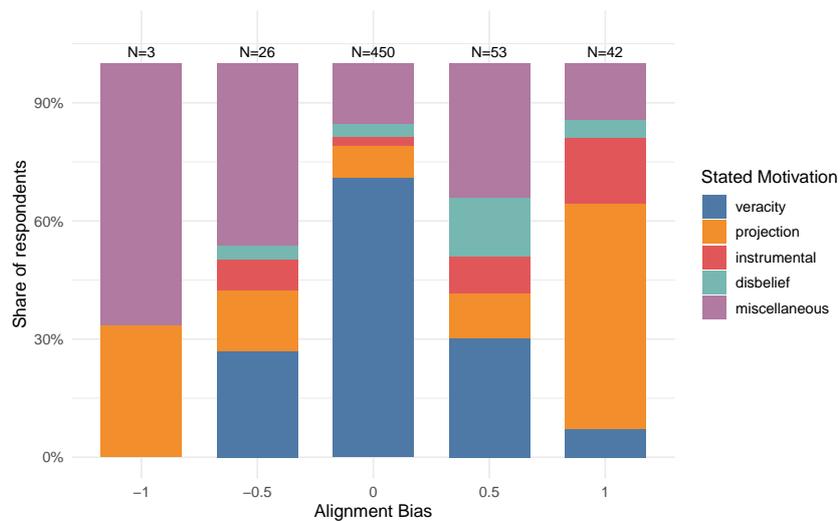
\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$



(a) Overall Distribution of Stated Motivations



(b) Stated Motivations by Receiver ACTION condition



(c) Stated Motivations by Alignment Bias

Figure A.3: Self-Reported Motivations for Information Provision. *Notes:* Panel (A) shows the overall distribution of primary stated motivations across all senders. Panel (B) shows the distribution by receiver ACTION condition. Panel (C) shows the distribution across values of the alignment bias.

Table B.13: Coding Scheme and Decision Rules

Category	Definition	Key Decision Rule / Boundary
Veracity	Truth-telling, honesty, fairness, or the receiver’s right to accurate information; non-directional motivation	Must not depend on political direction; sender shares all facts, even those contradicting own beliefs
Projection	Own beliefs or ideology explicitly linked to disclosure decision, including alignment of information	Requires explicit link between beliefs and sharing decision; mere opinions without disclosure reference → <i>did not refer to disclosure</i> ; if the sender references influencing behavior → <i>instrumental</i>
Instrumental	Intent to influence receiver’s actions (e.g., donating, voting, supporting an organization)	Requires explicit behavioral outcome; influencing beliefs only → <i>projection</i>
Disbelief	Doubts about accuracy or credibility used to justify withholding	Must explicitly refer to credibility concerns; politically motivated skepticism → <i>projection</i>
Did not refer to disclosure	No reference to sharing, withholding, or receiver	Triggered in first decision step; excludes further classification
Miscellaneous	Refers to decision but lacks interpretable reasoning	Residual category for non-classifiable responses

Table B.14: Inter-Model Reliability: Cohen’s  $\kappa$  and Disagreement Rates

Category	Cohen’s $\kappa$	Disagreement Rate
Veracity	0.887	5.0%
Projection	0.670	3.8%
Instrumental	0.788	0.9%
Disbelief	0.708	1.6%
Did Not Refer to Disclosure	0.771	11.4%
Miscellaneous	0.524	10.2%
Primary Motive (overall)	0.755	16.3%

Table B.15: Category-Level Precision, Recall, and F1-Score

Category	Class	Precision	Recall	F1	Support
Veracity	0	0.97	0.95	0.96	809
	1	0.91	0.94	0.92	390
Projection	0	0.97	0.99	0.98	1120
	1	0.76	0.63	0.69	79
Instrumental	0	0.99	1.00	1.00	1171
	1	0.84	0.75	0.79	28
Disbelief	0	1.00	0.99	0.99	1171
	1	0.62	0.86	0.72	28
Did Not Refer to Disclosure	0	0.86	0.93	0.89	615
	1	0.92	0.84	0.88	584
Miscellaneous	0	0.97	0.92	0.94	1083
	1	0.48	0.72	0.58	116

Table B.16: Confusion Matrices: Binary Categories (rows = GPT-4o, columns = Claude Sonnet)

Category		Claude = 0	Claude = 1
Veracity	GPT = 0	771	38
	GPT = 1	22	368
Projection	GPT = 0	1104	16
	GPT = 1	29	50
Instrumental	GPT = 0	1167	4
	GPT = 1	7	21
Disbelief	GPT = 0	1156	15
	GPT = 1	4	24
Did Not Refer to Disclosure	GPT = 0	570	45
	GPT = 1	92	492
Miscellaneous	GPT = 0	993	90
	GPT = 1	32	84

Table B.17: Confusion Matrix: Primary Motive (rows = GPT-4o, columns = Claude Sonnet)

GPT-4o	Claude Sonnet					
	DNRD	Disbelief	Misc.	Projection	Instr.	Veracity
Did Not Refer to Disclosure	<b>492</b>	3	72	5	0	12
Disbelief	2	<b>20</b>	2	0	0	2
Miscellaneous	17	2	<b>84</b>	2	0	11
Projection	8	1	11	<b>40</b>	3	9
Instrumental	1	0	0	2	<b>17</b>	8
Veracity	17	1	2	2	0	<b>351</b>

DNRD = Did Not Refer to Disclosure; Instr. = Instrumental.

Table B.18: Coding Examples

Motive	Example
Veracity	<p><i>“For the questions, I went off my knowledge and my gut feelings. For whether or not to tell people the correct answer to their missed question, I always picked for them to learn the correct answer as I want our society to be as educated as it can be whether I like the stats or not.”</i></p> <p><i>“I prefer facts even if they go against what I thought.”</i></p> <p><i>“People should always have the right information.”</i></p> <p><i>“I think it is important that people have all the facts.”</i></p> <p><i>“I believe that everyone should see their answer and the research. What influenced my decisions was the research (or the answers).”</i></p>
Projection	<p><i>“I think it’s important in this political environment that people advocate for immigrants regardless of statistics.”</i></p> <p><i>“Immigration is such a polarizing topic right now. People are literally being sent to concentration camps. I didn’t want to give anyone any information that would negatively impact their opinion against immigrants so I tried to give them positives and not share negatives.”</i></p> <p><i>“I know how harmless immigrants are, and I also know how much harder it’s been to get into this country. I want everyone else to know how harmless they are as well.”</i></p> <p><i>“Especially in this day and age where the general public strays away from fact-checking and relies heavily on AI generated search results and/or answers via apps which tends to spew out right-wing rhetoric, it is so important to fight against anti-intellectualism and continue to fight against fake news/misinformation/propaganda...”</i></p> <p><i>“Knowing how illegal immigrants change communities for the worse was my biggest consideration.”</i></p>
Instrumental	<p><i>“Wanting people to be more well informed, especially when making donation decisions towards a big cause.”</i></p> <p><i>“Some answers factored more into the organization I wanted donated more to. Other than that just using the information provided.”</i></p> <p><i>“I do not wish to have any funds going to the pro-immigration group. I only selected information to give to the ‘other participant’ that might discourage them from donating.”</i></p> <p><i>“I believe with the information, the participants will be able to make a more informed decision on how to allocate the \$500.”</i></p> <p><i>“Consideration of how that person’s opinions will directly affect a donation.”</i></p>
Miscellaneous	<p><i>“The facts.”</i></p> <p><i>“Money.”</i></p> <p><i>“Try to use common sense.”</i></p> <p><i>“Information provided and my own moral compass.”</i></p> <p><i>“Whether correcting or giving information would be beneficial and if they would be open to hearing it.”</i></p>
Did Not Refer to Disclosure	<p><i>“Research findings and the inhumane treatment that immigrants often receive.”</i></p> <p><i>“Estimating based on my knowledge of human nature but mostly instinctive.”</i></p> <p><i>“Mainly my own knowledge or opinions, partially the compensation and desire to get things correct.”</i></p> <p><i>“Being an empathetic person.”</i></p> <p><i>“Own beliefs and presented findings.”</i></p>
Disbelief	<p><i>“In terms of what influenced my decisions, I would say strength of evidence, the credibility of it, and if it had an intent to educate and not manipulate us.”</i></p> <p><i>“I think the premise of some of the questions is wrong. For instance, illegally entering the country is a crime; therefore 100% of illegal immigrants are criminals even if the government doesn’t universally prosecute them.”</i></p> <p><i>“Doesn’t really matter as significant portions of said study used effectively weaponized study propaganda.”</i></p> <p><i>“I don’t think people are going to change their mind on immigration based on outdated information from 15 years ago.”</i></p> <p><i>“How believable is the fact.”</i></p>