

# “Out of my network!”.

## The Effect of Misinformation and Partisan Political Messages on Social Network Pruning.

Natalia Aruguete\*   Paula Clerici†   Ernesto Calvo‡   Carlos Scartascini§  
Tiago Ventura¶

May 23, 2026

**Words: 8,038**

### Abstract

We study why social media users eliminate contacts who share dissonant content, and show that pruning operates through two distinct mechanisms depending on whether the content is partisan or factually false. We embed two image-based conjoint experiments in a nationally representative Argentine panel ( $N \approx 3,000$  per wave). When the dissonant content is partisan (sovereign debt), pruning intentions are frame-sensitive and directionally asymmetric across parties: *Frente de Todos* voters prune more in response to pro-Macri text and imagery, while *Juntos por el Cambio* voters respond in the opposite direction. When the dissonant content is misinformation (COVID-19 vaccine false claims), pruning rates are uniformly high and frame-insensitive. We call these mechanisms *partisan-directional pruning* and *epistemic-norm pruning* and show they coexist within the same population. We also show that partisan users expect to be pruned at rates exceeding both their own behavior and the population base rate, a self-reinforcing dynamic that could sustain network sorting independently of algorithmic curation.

---

\*Universidad Nacional de Quilmes - CONICET. Email: [nataliaaruguete@gmail.com](mailto:nataliaaruguete@gmail.com).

†Universidad Torcuato Di Tella - CONICET. Email: [pclerici@utdt.edu](mailto:pclerici@utdt.edu).

‡University of Maryland, Government and Politics, UMD. Email: [ecalvo@umd.edu](mailto:ecalvo@umd.edu).

§IADB. [CARLOSSC@iadb.org](mailto:CARLOSSC@iadb.org).

¶McCourt School of Public Policy, Georgetown University. Email: [tv186@georgetown.edu](mailto:tv186@georgetown.edu).

# 1 Introduction

Research on social media sorting describes echo chambers as the product of users following like-minded contacts and selecting ideologically congruent content (Stroud, 2010; Bakshy et al., 2015; Conover et al., 2011; Sunstein, 2018; Mosleh et al., 2021). A less studied but theoretically important mechanism is *pruning*, the active removal of friends or followers who post dissonant messages or spread misinformation (Kwak et al., 2011; Bode, 2016; Anspach, 2017). While research on content selection has advanced considerably, we know comparatively little about what drives users to eliminate existing network ties, and whether the same mechanisms operate when the dissonant content is politically partisan versus factually false.

This article examines pruning intentions using two image-based conjoint experiments fielded in Argentina, in the context of two issues that sharply divided voters during 2021–2022: the sovereign debt crisis and COVID-19 misinformation. The conjoints expose respondents to paired tweets varying four frame elements (author, text, image, and social support) and ask whether they would remove the friend who shared each tweet, and whether they expect their own friends to remove them. We use “pruning” as shorthand for “stated pruning intention”; the intention-behavior gap is treated as a limitation throughout.

Our central theoretical contribution is the distinction between two pruning mechanisms with different empirical signatures. *Partisan-directional pruning* is activated by out-group political content: the decision to prune is sensitive to ideological cues carried by the message’s frame elements, and varying those cues produces directionally asymmetric responses across partisan groups. *Epistemic-norm pruning* is activated by misinformation: the decision reflects the perceived violation of shared epistemic standards, not partisan identity, and occurs at high, broadly symmetric rates regardless of the observer’s partisan affiliation, with little sensitivity to the mes-

sage’s specific frame elements.

The empirical results support this account, with nuances we flag throughout. In the debt conjoint, frame elements drive pruning in opposing directions across parties for the text element. *Frente de Todos* (FdT) voters additionally respond directionally to image and header cues, while *Juntos por el Cambio* (JxC) voters show weaker or null responses to these elements, a response-intensity asymmetry that the partisan-directional mechanism in its strong form did not predict. In the COVID-19 conjoint, pruning rates are uniformly high and frame-insensitive for both parties, with the single exception of the abortion-fetus text claim, which raises pruning rates symmetrically across the partisan divide. In the misinformation conjoint, partisans prune at higher rates than nonpartisans; in the debt conjoint, this gap disappears, a wave-level asymmetry we treat as informative on its own terms rather than as confirmation of the partisan-elevation hypothesis. Finally, we find no evidence of the ideological asymmetry in misinformation tolerance documented in the United States: Argentine left- and right-leaning voters respond to equivalent misinformation signals at comparable rates.

The article makes four contributions. First, we provide a theoretical framework that disaggregates pruning into two mechanisms with distinct observable implications, clarifying why prior empirical work on unfriending has produced heterogeneous findings. Second, we provide experimental evidence from a non-U.S. context using a nationally representative panel with two pre-registered conjoints (one per wave) across two issue domains. Third, we contribute to the debate on ideological asymmetry in misinformation sharing (Osmundsen et al., 2021; Guess et al., 2021): our findings do not replicate the US conservative asymmetry, suggesting it may be context- or institution-specific. Fourth, by documenting that partisan users expect to be pruned at elevated rates (consistent with a friendship paradox mechanism in heterogeneous networks),

we identify a self-reinforcing feedback between pruning behavior and expectations that could sustain network sorting even in the absence of algorithmic amplification.

## 2 “What I Don’t Like” vs. “Who I Don’t Like”: Two Mechanisms of Network Pruning

Research on political polarization has long asked whether social media contributed to its rise (Stroud, 2010; Settle, 2018; Sunstein, 2018). One influential argument holds that social media generates echo chambers by allowing users to choose like-minded contacts and avoid counter-attitudinal content. Sunstein (2018) argued that user-side sorting, by which we mean choices about whom to follow and what to consume, increases homophily and reduces exposure to dissenting views, with polarization rising as cross-cutting contact declines (Mason, 2018).

More recent studies find less evidence of large-scale segregation. User networks documented in early work (Bakshy et al., 2015; Conover et al., 2011) are not substantially more segregated than offline media consumption (Gentzkow and Shapiro, 2011). Homophily is concentrated in politically oriented interactions (Barberá et al., 2015), and media diets in the United States are less polarized than often assumed, with centrist sources remaining prominent (Guess et al., 2021). Cross-ideological exposure is more common than early accounts suggested.

Sorting into partisan communities is documented on social media (Torres-Lugo et al., 2022), but the literature has focused primarily on self-selection into communities (Bakshy et al., 2015; Messing and Westwood, 2014) and content sharing (Aruguete et al., 2021). Less attention has been paid to the dissolution of existing ties: removing friends who post disliked content (Anspach, 2017; Sasahara et al., 2021), blocking users from news feeds (Kaiser et al., 2022), or unfollowing contacts for political reasons (Kwak et al., 2011; Bode, 2016). This article focuses

on this second mechanism, pruning, as a driver of network homogenization that is conceptually and empirically distinct from selective following.

The social media bubble literature thus rests on at least two distinct user-side processes: selecting like-minded contacts and removing dissonant ones. Unfriending (Sasahara et al., 2021) and blocking (Kaiser et al., 2022) contribute to network sorting by shifting the composition of content in users’ feeds and restructuring communication networks. Yet existing accounts treat pruning as a relatively uniform response to dissonance. We argue this obscures an important distinction. The motivation to prune a friend who shares a partisan message differs in kind, and not only in degree, from the motivation to prune a friend who spreads misinformation; the psychological and social mechanisms involved are different.

## Two Pruning Mechanisms

**Partisan-directional pruning: “who I don’t like.”** The first mechanism is activated by exposure to politically dissonant content, which we define as messages carrying ideological signals misaligned with the receiver’s partisan identity. The decision to prune is a response to *who* is speaking. A voter on the left who encounters pro-right content experiences cognitive dissonance (Festinger, 1957) and affective discomfort grounded in negative partisanship (Iyengar et al., 2019): the contact’s sharing behavior reveals an out-group identity, motivating network contraction. The mechanism is closely related to partisan motivated reasoning (Taber and Lodge, 2006), which leads partisans to attend more closely to ideological cues and react more strongly to out-group signals than nonpartisans (Osmundsen et al., 2021). The key empirical implication is frame sensitivity: variation in the author, text, image, and social support signals should produce directionally asymmetric differences in pruning rates across partisan groups. The theory does not require both parties to show responses of equal magnitude, only that the direction of the

response inverts across the partisan divide.

**Epistemic-norm pruning: “what I don’t like.”** The second mechanism is activated by misinformation, which is content that violates shared norms of factual accuracy regardless of ideological direction. The trigger is the perceived violation of an epistemic norm (Pennycook and Rand, 2021; Lewandowsky et al., 2012). Norms around accuracy and truth-telling are broadly shared across partisan lines (Marwick and Lewis, 2017; Roozenbeek et al., 2020): a contact who spreads misinformation violates these norms irrespective of political alignment, motivating pruning as social sanctioning (Centola, 2018; Watts and Dodds, 2007). The empirical implication is symmetry and frame insensitivity: the specific ideological packaging of the message should have little additional effect on pruning rates, and partisans on both sides should prune at broadly similar rates. The prediction stands in contrast to the US literature on ideological asymmetry in misinformation sharing (Osmundsen et al., 2021; Guess et al., 2021; Altay and Acerbi, 2023), which documents higher misinformation engagement among conservatives. If that asymmetry reflects differential exposure or motivated sharing, and not fundamentally different epistemic norms, the pruning response should remain symmetric even when sharing is not.

**Polarization as cross-cutting amplifier.** Both mechanisms should be amplified by affective polarization (Mason, 2018; Iyengar et al., 2019). Higher polarization should increase the aversiveness of out-group contact and the perceived social cost of association with a misinformation spreader, raising pruning rates across both mechanisms.

**Pruning expectations and the friendship paradox.** Pruning behavior has a social mirror: users form expectations about whether they themselves will be pruned. The friendship paradox (Feld, 1991; Jackson, 2019)—the tendency for individuals’ friends to be more popular, more extreme,

or more politically active than themselves on average—implies that users embedded in partially partisan networks will systematically overestimate the partisanship of their social environment (Lerman et al., 2016; Eom and Jo, 2015). Formally, if  $p$  denotes the population prevalence of partisans and  $\bar{k}_P > \bar{k}_{NP}$  are the average degrees of partisans and nonpartisans, a random user’s expected share of partisan friends is:

$$\mathbb{E}[\text{partisan friend share}] = \frac{p \cdot \bar{k}_P}{p \cdot \bar{k}_P + (1 - p) \cdot \bar{k}_{NP}} > p \quad (1)$$

A user whose network is overrepresented by partisan contacts will rationally anticipate being pruned at rates exceeding the population average, generating a self-reinforcing dynamic in which expecting to be pruned increases the propensity to prune preemptively (Sasahara et al., 2021; Tokita et al., 2021). Social projection, risk aversion, and prior personal experience could generate the same expectation pattern; we treat the friendship paradox as one consistent account rather than an established mechanism.

### 3 Research Design

The conjoint design connects changes in network topology to partisanship and content type. Following Hainmueller et al. (2014), conjoint experiments present respondents with two profiles varying randomized attributes; applications span immigration (Hainmueller and Hopkins, 2015), corruption (Mares and Visconti, 2020; Klačnjaja et al., 2021), and vote choice (Franchino and Zucchini, 2015; Kirkland and Coppock, 2018). Unlike traditional conjoints that force a choice between profiles, our design allows “both” and “neither” responses, a non-forced design following Aruguete et al. (2021), which avoids conflating the threshold decision to prune with a relative preference between two tweets.

Each respondent is exposed to a pair of tweets that vary four features: the author (header), the text content, the embedded image, and the level of social support (likes and retweets).

Table 1 summarizes the frame elements for each wave; Figure 1 presents example tweet pairs.

**Table 1** Conjoint Frame Elements by Wave

<b>Feature</b>	<b>N</b>	<b>Wave 1: COVID-19 Misinformation</b>	<b>Wave 2: Sovereign Debt</b>
<i>Author</i>	2	<i>Médicos por la Verdad</i> (MxV): doctors and virologists known for opposing vaccines. <i>Viviana Canosa</i> (VC): TV presenter who promoted vaccine denialism.	<i>Página 12</i> : left-leaning newspaper. <i>La Nación</i> : right-leaning newspaper.
<i>Text</i>	2	Claim that vaccines were developed using cells from aborted fetuses. Claim that a nurse died after receiving the Sputnik V vaccine.	Text blaming former President Macri for Argentina’s debt crisis. Text defending Macri’s record.
<i>Image</i>	3	Neutral (virus photo); fetal ultrasound; nurse photo. <i>Same image cannot appear twice in a pair.</i>	Neutral (bar chart); Macri photo; Fernández photo. <i>Same image cannot appear twice in a pair.</i>
<i>Social Support</i>	2	High or low volume of likes and retweets. Four within-pair combinations: High–High, High–Low, Low–High, Low–Low.	

Each wave yields 48 distinct tweet profiles. The image-repetition constraint reduces the set of valid pairs shown to any respondent.

Head

Medicos por la Verdad @medicosporlaverdad

¡Vacuna Sputnik V! Usan células de fetos abortados para hacer la vacuna y después quieren que se las demos a nuestros hijos! #MedicosPorLaVerdad

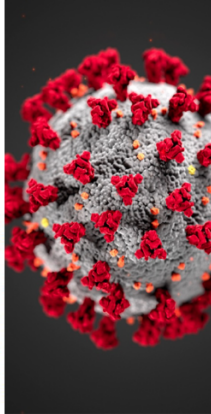
Medicos por la Verdad @medicosporlaverdad

#EnfermeraZapala Muere enfermera cabo primero Daniela Dominguez en Zapala después de recibir la vacuna Sputnik V contra el covid 19.

Viviana Canosa @vivianacansa

¡Vacuna Sputnik V! Usan células de fetos abortados para hacer la vacuna y después quieren que se las demos a nuestros hijos! #MedicosPorLaVerdad

Text



Image

217 Retweets 121 Likes

4 Retweets 3 Likes

4 Retweets 3 Likes

(a) Wave 1 (COVID-19 Misinformation Conjoint).

Head

Página 12 @página12

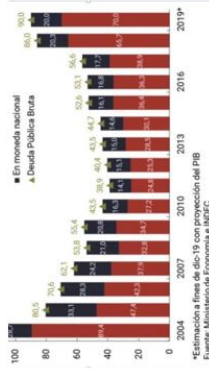
La filosa ironía del jefe de gabinete Bonaerense, Carlos Blanco, contra Mauricio Macri: "Es como que Drácula se preocupe por transfusiones"

"Que Macri se preocupe por el endeudamiento es como que Drácula se preocupe por las transfusiones de sangre", Carlos Blanco.

LA NACIÓN @LANACION

Mauricio Macri en conferencia de prensa: "Tenemos que honrar nuestras deudas, respetar las instituciones"

Text



Image

2 Retweets 1 Likes

2 Retweets 3 Likes

202 Retweets 123 Likes

(b) Wave 2 (Sovereign Debt Conjoint).

Figure 1 Example Tweet Pairs from the COVID-19 Misinformation (a) and Sovereign Debt (b) Conjoint Experiments.

## Survey, Variables, and Estimation

Both conjoint experiments were embedded in a nationally representative online survey of Argentine voters conducted by Netquest. Wave 1 was fielded between November 4 and December 7, 2021, spanning Argentina’s November 14 midterm general elections, with the bulk of responses collected in the week before the election and a second push between late November and early December (approximately 3,000 respondents). Wave 2 was fielded between February 7 and 15, 2022, during active renegotiations with the International Monetary Fund over Argentina’s sovereign debt (approximately 2,400 respondents). After reading each pair of tweets, respondents answer three questions: *Q1 (Prune)*: Would you remove from your friends list whoever shares this tweet? *Q2 (Forced choice)*: If you deleted only one, which? *Q3 (Expectation)*: Would you expect your friends to delete YOU for sharing one of these tweets? Q1 (own pruning intention) and Q3 (counterfactual self-referential expectation) are the primary dependent variables; each is coded 1 if the respondent indicates pruning or being-pruned for that tweet side.

*Partisanship* is measured via self-reported identification with FdT or JxC and via intended vote choice with a blank-vote option; respondents with neither party nor a non-blank vote are classified as nonpartisans. *Polarization* is the absolute distance between the respondent’s seven-point ideological self-placement and the scale midpoint (0 = centrist, 3 = extreme). We report average marginal component effects (AMCEs) and marginal means (Hainmueller et al., 2014; Leeper et al., 2020), estimated via OLS with standard errors clustered by respondent. Predicted probabilities (Figures 3 and 5) are generated from logistic regression with clustered standard errors to avoid out-of-range predictions. Results are robust to sociodemographic controls (SIF, Tables A12–A14).

## Pre-registration Status

Each wave was pre-registered separately on OSF prior to fielding: the Wave 1 PAP (COVID-19 conjoint) and the Wave 2 PAP (debt conjoint, deposited one day before Wave 2 fielding). The two PAPs together cover the conjoint designs, sample sizes, dependent-variable coding, AMCE and marginal-means estimation, and the within-wave content of  $H_{1a}$ ,  $H_{1b}$ ,  $H_{2a}$ ,  $H_{2b}$ ,  $H_{3a}$ ,  $H_{3b}$ ,  $H_{3c}$ ,  $H_{4a}$ ,  $H_{4b}$ ,  $H_5$ , and  $H_6$ .

Three components are exploratory and not described in the pre-registration: (i) the two-mechanism framework integrating both waves, developed after observing the cross-wave contrast; (ii) the cross-conjoint component of  $H_{3c}$  (steeper polarization gradient predicted for partisan than misinformation content); and (iii) the cross-conjoint component of  $H_{4b}$  (larger directional shifts in expectations for the debt conjoint). The framework is a structured interpretation of pre-registered findings rather than a confirmed mechanism; replication with the framework pre-specified would strengthen its status. Both PAPs are reproduced in the SIF.

## Hypotheses

We test six families of hypotheses.

**Signal strength ( $H_1$ ).** Within the COVID-19 conjoint, the frame elements vary in how explicitly they depart from accepted scientific knowledge. The author feature contrasts a nationally prominent anti-vaccine commentator (Canosa) with the less recognized organization *Médicos por la Verdad*; the text ranges from a nurse death claim to the more inflammatory abortion-fetus link; and the image ranges from a neutral virus photograph to a fetal ultrasound. If pruning responds to the magnitude of an epistemic violation rather than to partisan cues, more explicit misinformation signals should produce higher pruning rates ( $H_{1a}$ ) and higher expectations of

being pruned ( $H_{1b}$ ), symmetrically across partisan groups.

**Partisan elevation ( $H_2$ ).** A substantial literature documents that partisans are unconditionally more motivated than nonpartisans to engage with political content on social media (Mason, 2018; Barberá, 2020; Slothuus and De Vreese, 2010; Guess et al., 2021; Törnberg, 2018). The elevated motivation extends to misinformation: Osmundsen et al. (2021) shows that partisans share false news at higher rates, and Bode (2016) and Kaiser et al. (2022) find that partisan users are more likely to block or remove dissenting contacts. We expect the same unconditional elevation to hold for pruning behavior: partisan users will prune at higher rates than nonpartisans across both political and misinformation content ( $H_{2a}$ ), and will expect their friends to prune at higher rates ( $H_{2b}$ ).

**Frame sensitivity and polarization ( $H_3$ ).** The partisan-directional mechanism predicts that frame elements carrying ideological cues should trigger cognitive dissonance (Festinger, 1957) and motivated reasoning (Taber and Lodge, 2006) among partisans encountering out-group content, raising the probability of pruning. The same frame elements should reduce pruning among co-partisans. The misinformation mechanism, by contrast, predicts no such directional sensitivity. Together: in the partisan conjoint, frame elements signaling the out-group will increase pruning rates for co-partisan respondents and decrease (or fail to increase) them for opposing partisans, producing directionally asymmetric AMCEs across parties ( $H_{3a}$ ), with parallel effects on pruning expectations ( $H_{3b}$ ). More polarized partisan users will prune at higher rates ( $H_{3c}$ ); the two-mechanism framework also predicts the polarization gradient should be more pronounced for partisan-directional content than for misinformation, though we treat the cross-

conjoint component of  $H_{3c}$  as an exploratory extension.<sup>1</sup>

**Friendship paradox ( $H_4$ ).** If partisan users are embedded in networks disproportionately populated by politically active contacts, they should expect their friends to prune political content sharers at higher rates than nonpartisan users ( $H_{4a}$ ). The directional sensitivity of pruning expectations to out-group frame elements should be more pronounced in the partisan conjoint than in the misinformation conjoint ( $H_{4b}$ ).<sup>2</sup>

**Ideological asymmetry ( $H_5$ ).** A well-documented finding in the United States is that conservative voters are more likely than liberals to share, consume, and believe online misinformation (Guess et al., 2021; Osmundsen et al., 2021; Altay and Acerbi, 2023). Whether this asymmetry reflects structural features of the US media environment, properties of the Republican party’s rhetorical style, or a more fundamental difference in epistemic dispositions remains debated. The epistemic-norm mechanism offers a way to adjudicate. If misinformation triggers pruning through a shared norm of factual accuracy, left and right partisans should respond symmetrically, and the US asymmetry should not replicate. The relevant test is whether AMCEs for misinformation frame elements differ across parties (i.e., differential *frame sensitivity*), not whether baseline pruning rates differ (which is addressed in  $H_2$ ). If misinformation tolerance is ideologically asymmetric, JxC voters will show lower frame sensitivity to misinformation signals than FdT voters.

**Attention as moderator ( $H_6$ ).** Survey-based conjoints face the general challenge that not all respondents attend equally to the treatment (Berinsky et al., 2014); response latency provides

---

<sup>1</sup>The within-wave polarization moderator was pre-registered in the Wave 2 PAP; the cross-conjoint directional prediction (steeper gradient for partisan than misinformation content) is exploratory.

<sup>2</sup>The within-wave directional expectations prediction was pre-registered in the Wave 2 PAP; the cross-conjoint comparison with the misinformation conjoint is exploratory.

a behavioral proxy for engagement, and [Banks et al. \(2021\)](#) demonstrate that latency amplifies the effects of social media framing on polarization. We therefore expect that the partisan frame effects under  $H_3$  will be stronger among respondents who engaged carefully with the stimuli, while misinformation pruning effects under  $H_1$  may be less attention-sensitive because the norm violation is salient even under low engagement.<sup>3</sup>

## 4 Results

We organize results around the two-mechanism framework. All effects refer to stated pruning intentions; the intention-behavior gap is discussed in the limitations. Regression tables are in the SIF.

### Signal Strength and Epistemic-Norm Pruning ( $H_{1a}$ , $H_{1b}$ )

Figure 2 presents AMCEs for each frame element within the COVID-19 conjoint. The pattern has three components.

The text feature produces the only substantively and statistically significant effect. The abortion-fetus claim raises the probability of pruning by approximately 0.19–0.22 relative to the nurse death claim, a large effect by conjoint standards, and this effect is statistically indistinguishable between FdT and JxC voters.

The author, image, and social support features are uniformly null across all groups. The null author result deserves comment: Viviana Canosa is a nationally prominent anti-vaccine commentator and the theory predicted she would function as a stronger epistemic signal than MxV.

A plausible reading is that both authors are sufficiently associated with vaccine denialism that

---

<sup>3</sup>The within-wave attention prediction was pre-registered in the Wave 2 PAP; its cross-wave divergence (debt vs. COVID-19) is interpretive rather than pre-registered.

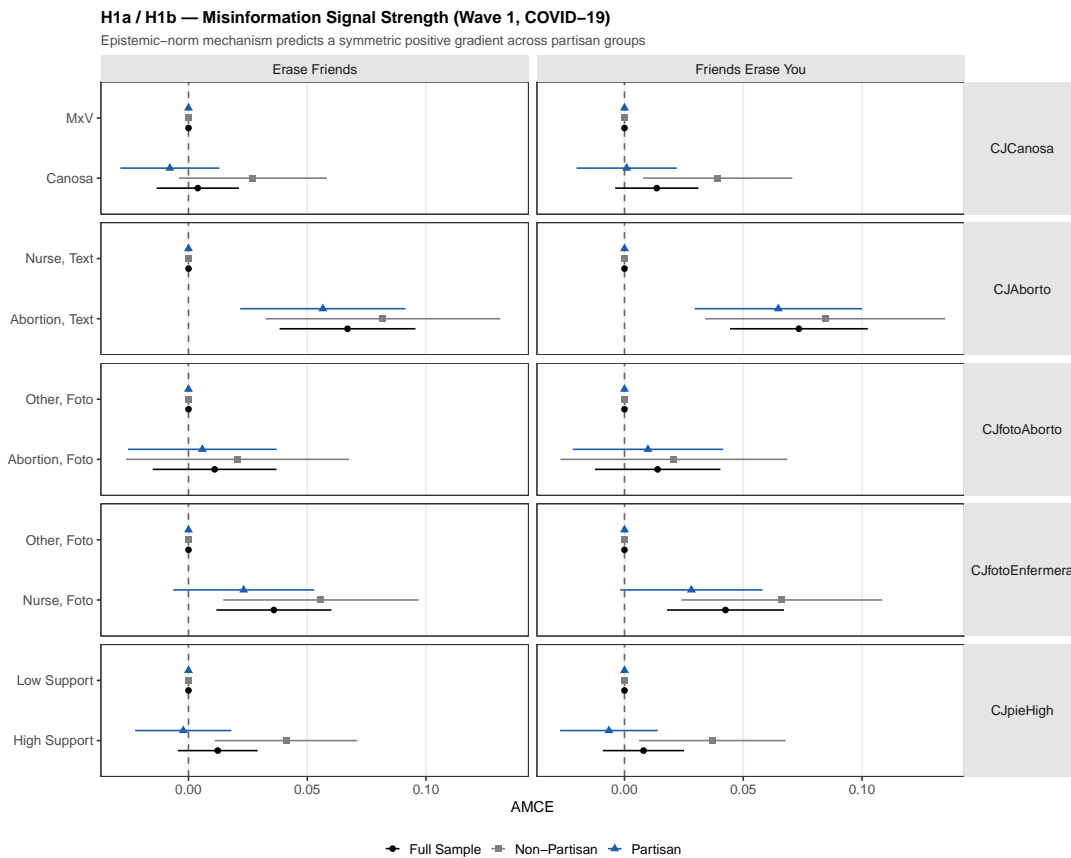
they trigger equivalent pruning responses, a signal-ceiling interpretation that we acknowledge is post hoc. An alternative reading is simply that author and image cues are weaker than text cues for activating the epistemic-norm response; the attention analysis ( $H_6$ ) provides additional evidence on this point. Social support is null in both waves, suggesting that peer endorsement does not modulate pruning once the content is salient.

The signal strength gradient holds for both dependent variables: the abortion-fetus text raises both the probability of pruning ( $H_{1a}$ ) and the expectation of being pruned ( $H_{1b}$ ), symmetrically across partisan groups.  $H_{1a}$  and  $H_{1b}$  are partially supported: of the three frame elements expected to carry signal strength (author, text, image), only text produced the predicted gradient. This is the first empirical signature of the epistemic-norm mechanism: one strong content-level driver, frame-element insensitivity elsewhere, and symmetric effects across the partisan divide.

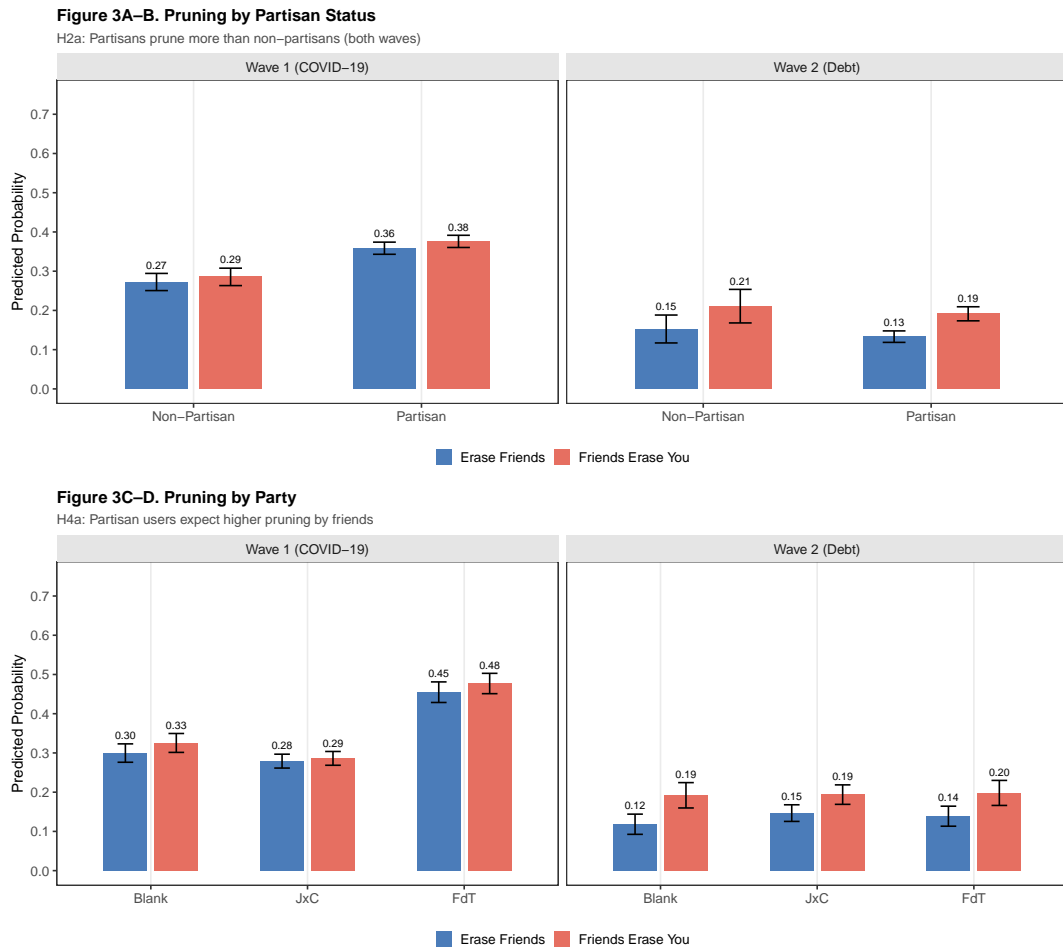
### **Partisanship and Unconditional Pruning ( $H_{2a}$ , $H_{2b}$ )**

Figure 3 reveals a striking wave-level asymmetry. In the COVID-19 conjoint, partisans prune at substantially higher rates than nonpartisans (0.36 vs. 0.27 for self-pruning; 0.38 vs. 0.29 for expectations;  $p < .001$  for both); FdT voters lead at 0.45, JxC at 0.28, blank voters at 0.30. Both  $H_{2a}$  and  $H_{2b}$  are supported in Wave 1.

*In the sovereign debt conjoint,  $H_{2a}$  and  $H_{2b}$  are not supported.* Predicted self-pruning rates are 0.12–0.15 across all three groups, with overlapping CIs; predicted expectation rates are 0.19–0.21, also with overlapping CIs. The pre-registered hypothesis (that partisans prune more than nonpartisans across both political and misinformation content) is contradicted by Wave 2. The null is not a power failure: the debt conjoint produces strong and precisely estimated frame effects (below), so the null gap reflects a genuine absence of unconditional partisan elevation in this domain.



**Figure 2** Signal Strength and Epistemic-Norm Pruning ( $H_{1a}$ ,  $H_{1b}$ ). AMCEs for frame elements in the COVID-19 misinformation conjoint, by partisanship. The abortion-fetus text is the only element that systematically raises pruning rates, and does so symmetrically across partisan groups. OLS with clustered standard errors; 95% CIs.



**Figure 3** Unconditional Pruning Rates by Partisan Status and Party ( $H_{2a}$ ,  $H_{2b}$ ,  $H_{4a}$ ). Partisans prune significantly more than nonpartisans in Wave 1 but not in Wave 2. Respondents expect their friends to prune them at higher rates than they themselves report pruning others. Logistic regression with sociodemographic controls; 95% CIs.

We read the contrast as suggestive evidence, exploratory and not predicted, that partisan identity raises the baseline pruning rate specifically when the content activates a moral-norm response, but does not unconditionally elevate pruning across all political content. Partisan political content appears to be processed through ideological cue-matching, a process that produces directional frame sensitivity without raising baseline rates.

### Frame Sensitivity and the Partisan-Directional Mechanism ( $H_{3a}-H_{3c}$ )

The central empirical contrast of the paper is between the frame-insensitive epistemic-norm signature documented above and the frame-sensitive, directional signature predicted for the partisan conjoint. Figure 4 presents AMCEs for both conjoints, separately by party.

In the debt conjoint, the pro-Macri text produces a clean directional inversion: it raises FdT pruning by 0.11 ( $p < .05$ ) and lowers JxC pruning by 0.09 ( $p < .05$ ), for a net directional gap of 0.20 on the self-pruning outcome. This inversion is the defining signature of partisan-directional pruning: the same content, carrying an ideological cue, activates pruning among those for whom it is dissonant and suppresses it among those for whom it is congruent.

The image and header features show the directional logic for FdT voters but not for JxC voters. The Macri photograph produces a positive AMCE for FdT voters (approximately +0.064 on expectations, +0.053 on self-pruning); for JxC voters, the effect is near zero rather than negative.<sup>4</sup> The Fernández photograph shows a similar pattern (FdT: +0.038 self-pruning, +0.058 expectations; JxC: near-zero). The *Página 12* header reduces FdT self-pruning modestly ( $-0.05$ ,  $p < .05$ ), consistent with source congruence for left-leaning voters, and is null for JxC. Social support is null for both parties.

Different from the partisan-directional mechanism that predicted symmetric mirror-image responses across parties, the actual pattern is asymmetric: FdT voters show strong directional sensitivity across multiple frame elements, while JxC voters show clean directional sensitivity only for the text frame. The directional logic is preserved where both parties respond, but the JxC response intensity is muted. One candidate explanation is the political context. Wave 2 was fielded in February 2022, when the FdT base was actively mobilized against the government's

---

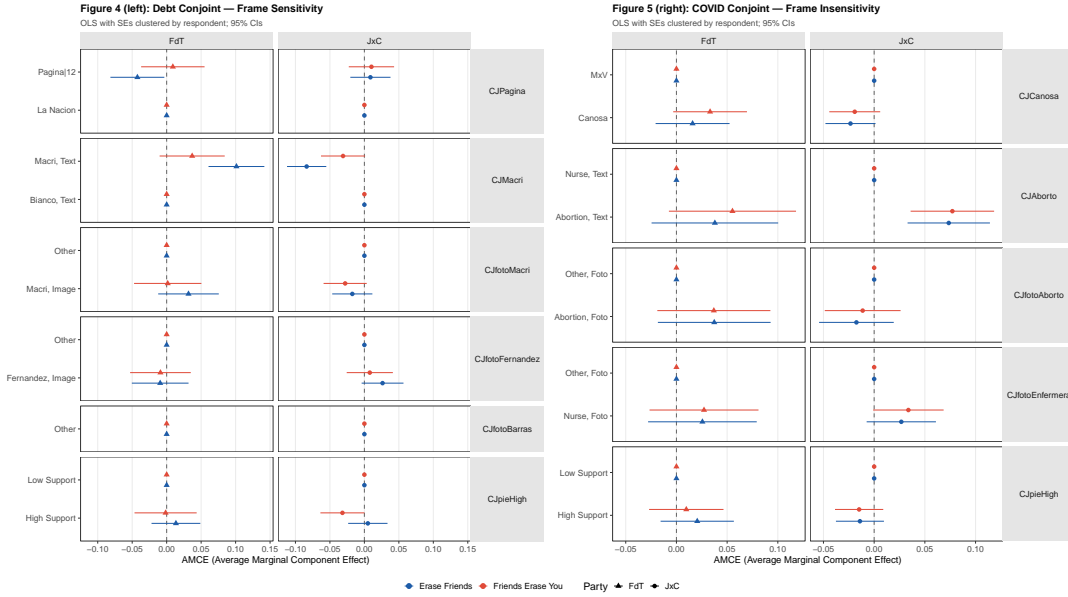
<sup>4</sup>Image AMCEs are reported on the same probability-shift scale as the rest of the AMCE reporting in the paper, i.e., shifts of approximately 3–6 percentage points in pruning probability.

debt negotiations while the JxC base was in a defensive opposition posture, making FdT voters more reactive to debt-related frame variation. A second candidate is selection: JxC voters engaged with political content during this period may have been a less reactive subsample. Our design cannot adjudicate; we flag the asymmetry as a finding to be replicated rather than a fully theorized feature.

In the COVID-19 conjoint, no frame element produces a statistically significant directional AMCE for either party, with the single exception of the abortion-fetus text, which raises pruning symmetrically for FdT (+0.22) and JxC (+0.19) voters. The symmetric, content-driven effect contrasts sharply with the directional effects in the debt conjoint. The misinformation conjoint's null results are tight enough to rule out effects of the magnitude observed in Wave 2; this is not a power issue. Together, the results provide support for  $H_{3a}$  and  $H_{3b}$ , with the caveat that the cross-party symmetry assumed by the original framing is not borne out for all frame elements.

**Polarization gradient: not as predicted by  $H_{3c}$ .** Figure 5 presents predicted pruning probabilities across the polarization gradient. Polarization is positively associated with pruning rates in both waves and for both parties, supporting the base  $H_{3c}$  prediction. *The cross-conjoint directional component of  $H_{3c}$  is not supported.* We predicted the gradient would be steeper for partisan-directional content (debt) than for misinformation, since the theory positions affective polarization as primarily an amplifier of identity-based responses. The data show the opposite: the gradient is steeper and more consistent in the COVID-19 conjoint than in the debt conjoint, with the FdT gradient particularly pronounced in Wave 1.

Two readings are plausible: polarization as we operationalize it (ideological extremity) may capture engagement and moral reactivity more than the partisan-identity amplification the theory had in mind, or floor effects in Wave 2 (pruning rates 0.12–0.15) may leave less room for

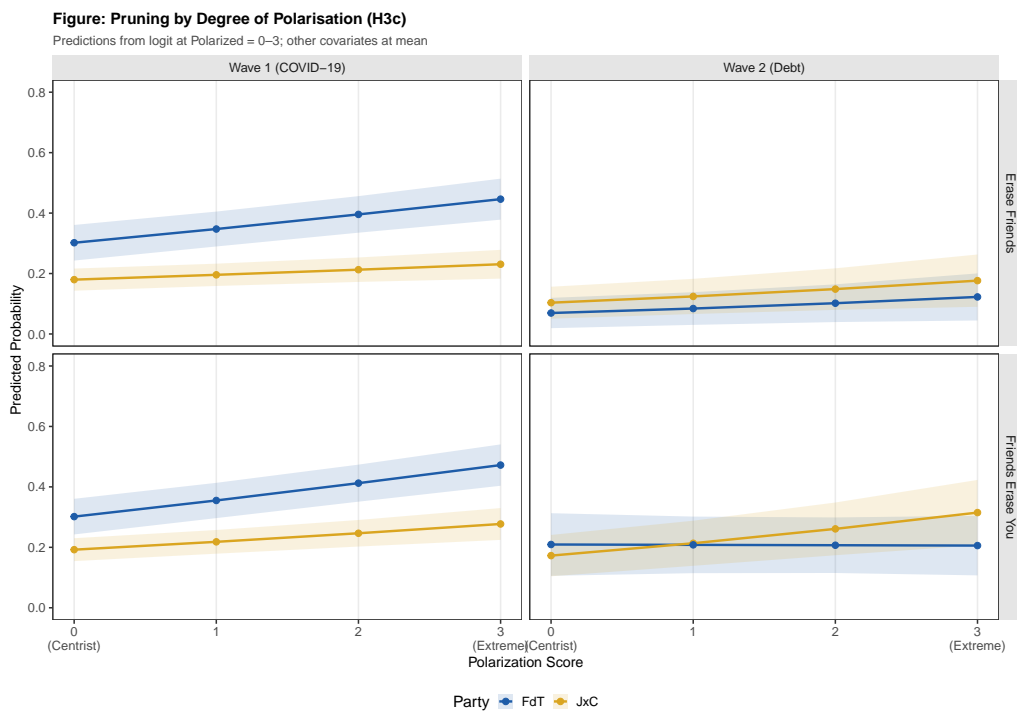


**Figure 4** Frame Sensitivity by Conjoint Wave ( $H_{3a}$ ,  $H_{3b}$ ). AMCEs for frame elements in the Sovereign Debt conjoint (left) and COVID-19 conjoint (right), by party. In the debt conjoint, the pro-Macri text frame shifts pruning in opposite directions across parties; image and header frames shift pruning directionally for FdT voters but produce near-null effects for JxC. In the COVID-19 conjoint, frame elements are uniformly null except for the abortion-fetus text, which raises pruning symmetrically. OLS with clustered standard errors; 95% CIs.

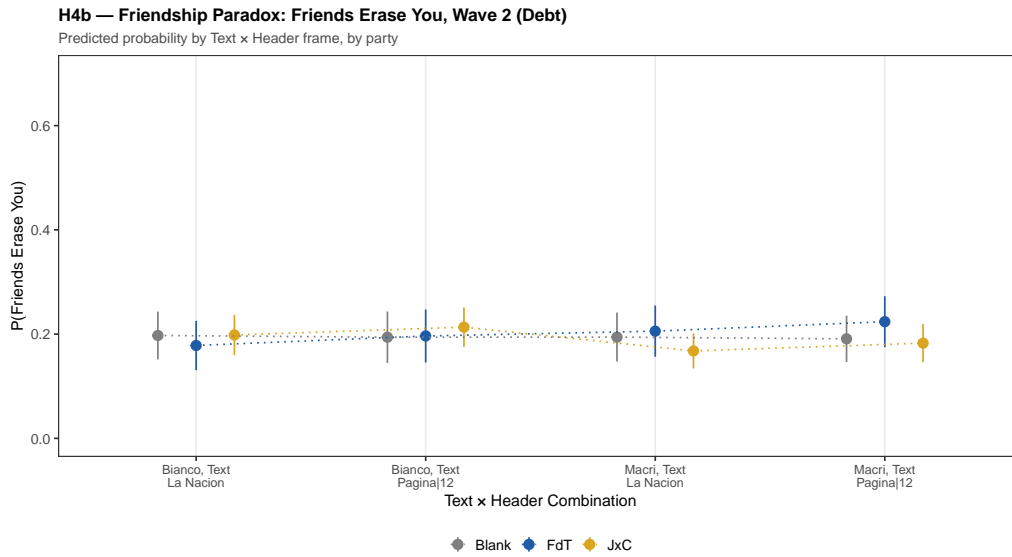
polarization to push rates upward. The cross-conjoint directional prediction of  $H_{3c}$  is not supported by these data.

### Friendship Paradox and Pruning Expectations ( $H_{4a}$ , $H_{4b}$ )

The unconditional results in Figure 3 provide mixed support for  $H_{4a}$ . In Wave 1, respondents expect their friends to prune them at higher rates than they themselves report pruning others, and the partisan-nonpartisan gap is positive: expected pruning is 0.38 for partisans vs. 0.29 for nonpartisans. *In Wave 2,  $H_{4a}$  is not supported.* Expected pruning is 0.19 for partisans and 0.21 for nonpartisans, a small gap in the opposite direction of the pre-registered prediction. The Wave 1 pattern is consistent with the friendship paradox account; the Wave 2 reversal parallels the wave-level asymmetry documented for  $H_{2a}/H_{2b}$  and suggests that the friendship-



**Figure 5** Pruning by Degree of Polarization ( $H_{3c}$ ). More polarized respondents prune at higher rates within each wave. Contrary to the cross-conjoint prediction of  $H_{3c}$ , the gradient is steeper in the COVID-19 conjoint than in the debt conjoint. Logistic regression with other covariates at means; shaded bands are 95% CIs.



**Figure 6** Directional Sensitivity in Pruning Expectations ( $H_{4a}$ ,  $H_{4b}$ ). Partisan frame elements (pro-Macri text and *Página 12* header) shift expectations directionally across parties, with the response substantially stronger for FdT voters. OLS with clustered standard errors; 95% CIs.

paradox mechanism, like the partisan-elevation mechanism, may operate selectively in moral-norm-violation contexts rather than across all political content.

The test of  $H_{4b}$  is equivocal. Figure 6 shows a directional pattern in expectations for FdT voters in the debt conjoint: they expect to be pruned more for pro-Macri text or the *Página 12* header. JxC voters show a weaker mirror pattern. However, the absolute range of predicted expectations is larger in the COVID-19 conjoint than in the debt conjoint (Wave 1 FdT range: 0.101; Wave 2 FdT range: 0.046), driven by the abortion-fetus text rather than by partisan cues. The COVID-19 conjoint produces larger but symmetric expectation shifts; the debt conjoint produces smaller but directional shifts. Whether this counts as support for  $H_{4b}$  depends on which feature of the prediction one privileges; we report both.

### Ideological Asymmetry in Misinformation Tolerance ( $H_5$ )

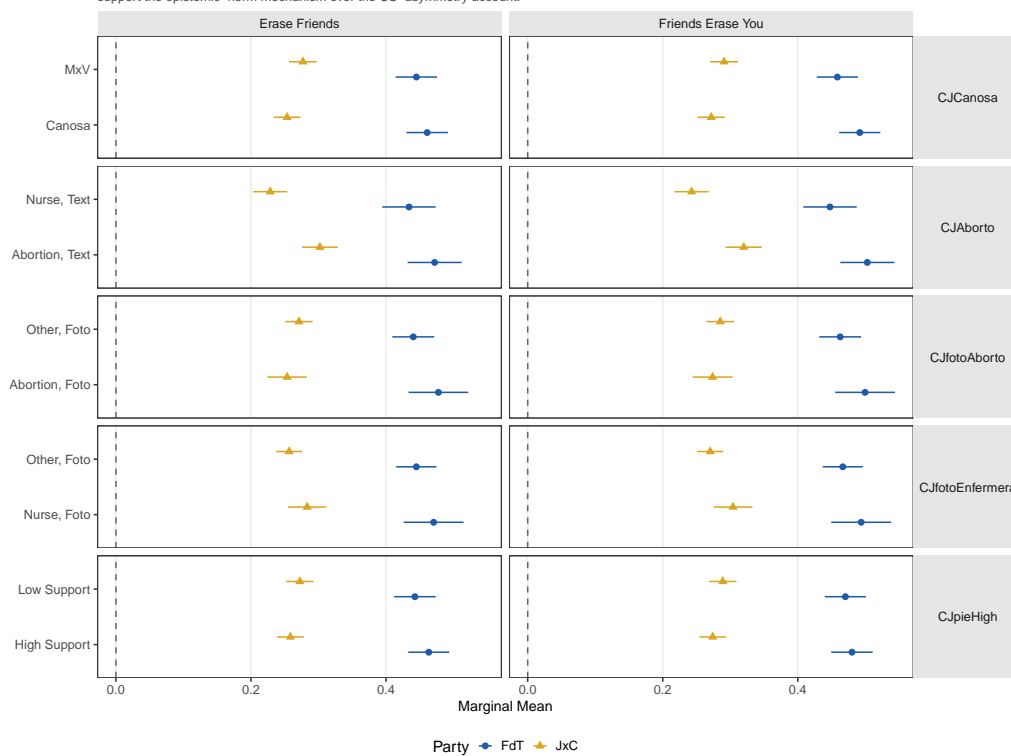
Figure 7 presents marginal means for each frame element in the COVID-19 conjoint, separately by party. Two findings need to be distinguished.

JxC voters do have a lower unconditional pruning rate than FdT voters in Wave 1 (0.28 vs. 0.45). The direction matches the Wave 1 PAP prediction that conservatives would prune less than liberals, but the gap is confounded with the partisan elevation effect documented in  $H_2$ : FdT voters prune more unconditionally across all content types in the misinformation conjoint, including for the nurse-death claim that the asymmetry account would not predict to distinguish parties. The conceptual claim under  $H_5$  is about differential *sensitivity* to the epistemic signal, not about overall pruning levels, so we treat the JxC  $\times$  frame interaction as the more informative test.

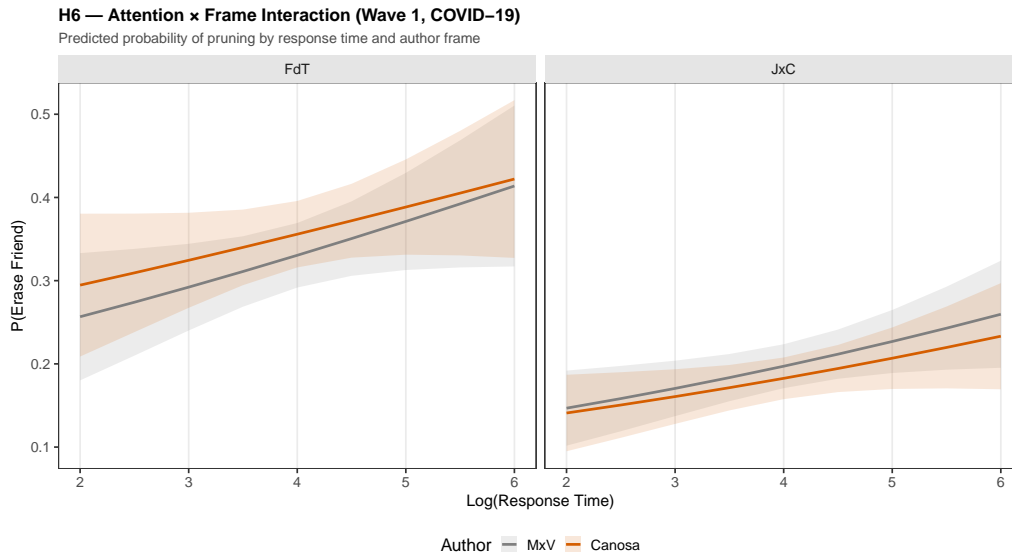
On that test, results are unambiguous. AMCEs are statistically indistinguishable across parties for every frame element. For the abortion-fetus text, which is the strongest epistemic signal, the AMCEs are +0.22 for FdT and +0.19 for JxC (SIF, Table A8), a substantively negligible gap. The interaction between JxC and the abortion-fetus text, which is the formal test of differential sensitivity, is +0.03 on self-pruning (95% CI: [-0.05, 0.11]) and +0.02 on expectations (95% CI: [-0.06, 0.10]) (SIF, Table A9). Argentine conservatives and liberals show equivalent frame sensitivity to equivalent misinformation signals. The result is consistent with the epistemic-norm mechanism and provides no support for the US asymmetry account in this context. A single-country comparison cannot adjudicate among the three explanations we offered (US media environment, Republican party rhetorical style, general property of conservative cognition); the Argentine result rules out the strongest form of the third but is consistent with the first two.

**H5 — Ideological Asymmetry in Misinformation Tolerance**

Wave 1 (COVID-19). Symmetric pruning rates across JxC and FdT support the epistemic-norm mechanism over the US-asymmetry account.



**Figure 7** Ideological Asymmetry in Misinformation Tolerance ( $H_5$ ). Marginal means for frame elements in the COVID-19 conjoint, by party. CIs overlap substantially across all frame elements and both outcomes. The null between-party difference in frame sensitivity is consistent with symmetric epistemic-norm pruning and inconsistent with the US conservative asymmetry account.



**Figure 8** Attention as a Moderator ( $H_6$ ). Predicted pruning probability as a function of log response time, by party and tweet author in the COVID-19 conjoint. More attentive respondents prune at higher rates, and the interaction with the Canosa author condition is positive and significant. 95% CIs.

### Attention as a Moderator ( $H_6$ )

Figure 8 shows that in the COVID-19 conjoint, response latency is positively and significantly associated with pruning for both FdT and JxC voters ( $p < .05$ ). The interaction between log response time and the Canosa author condition is positive and significant: respondents who spend more time reading the tweet are more likely to prune when the author is Canosa than when it is MxV, even though the unconditional author AMCE was near zero. We frame this as a heterogeneity finding given that the average AMCE for Canosa is null, and the interaction effect emerges only among more attentive respondents.

In the debt conjoint, the response latency coefficient is smaller and statistically non-significant for both parties. The contrast is substantively interpretable and not a power issue. In the debt conjoint, the partisan signals (Macri’s photograph and the pro-debt text) are salient enough to trigger pruning even among respondents who engage briefly: the signal is legible at low attention. In the misinformation conjoint, the epistemic content of the false claim requires more careful

processing to register as a norm violation, which is why latency amplifies the response. The pattern is consistent with dual-process accounts of misinformation processing (Pennycook and Rand, 2021), under which analytic engagement matters more for detecting epistemic violations than for processing identity-based cues.

## 5 Discussion

Pruning intentions on social media reflect two distinct mechanisms with different empirical signatures. *Partisan-directional pruning*, activated by politically dissonant content, is frame-sensitive and directionally asymmetric across partisan groups, though the intensity of that directional response is itself asymmetric in our 2022 Argentine sample, stronger for FdT than for JxC voters. *Epistemic-norm pruning*, activated by misinformation, is frame-insensitive, symmetric across the partisan divide, and driven by the severity of the factual violation rather than by the ideological identity of the sender. Both produce more homogeneous networks, but through different pathways that call for different interventions.

We provide five findings. First, the frame-sensitivity contrast between the two conjoints documents the central mechanism distinction: directional response to partisan content, symmetric response to misinformation. Second, partisan elevation in baseline pruning rates holds in the misinformation conjoint but not in the debt conjoint, suggesting that partisan identity raises baseline pruning specifically when content activates a moral-norm response. Third, Argentine conservatives and liberals show equivalent frame sensitivity to equivalent misinformation signals, which does not replicate the US conservative asymmetry and suggests the asymmetry is institutional rather than psychological. Fourth, partisan users in the misinformation conjoint expect to be pruned at rates exceeding both their own pruning behavior and the population base

rate, consistent with the friendship paradox account; in the debt conjoint, this expectation gap reverses, indicating that the friendship-paradox mechanism operates selectively in moral-norm-violation contexts. Fifth, attention moderates the misinformation response but not the partisan response, consistent with dual-process accounts in which detecting epistemic violations requires more analytic engagement than processing identity-based cues.

Two findings did not match predictions and warrant explicit acknowledgment. The polarization gradient was steeper in the misinformation conjoint than in the debt conjoint, the opposite of the cross-wave component of  $H_{3c}$ , suggesting that our extremity measure may capture moral reactivity more than identity-based amplification. And the directional response in the debt conjoint is asymmetric across parties in intensity, rather than symmetric, with FdT voters more reactive than JxC voters during the period we surveyed. We treat both as findings to be replicated and refined.

These findings extend the literature on social media sorting in four directions. First, they provide a direct experimental contrast between content-driven and norm-driven pruning within a single population, using a within-panel design that holds population heterogeneity constant. Prior work on unfriending and blocking has documented that partisanship and political content predict network removals (Bode, 2016; Kaiser et al., 2022; Sasahara et al., 2021) but has not distinguished whether the operative mechanism is ideological cue-matching or epistemic norm enforcement. Our two-conjoint design identifies the signatures of each mechanism independently and shows they coexist within the same population.

Second, the absence of asymmetry in the Argentine context suggests the conservative asymmetry documented in the United States (Guess et al., 2021; Osmundsen et al., 2021; Altay and Acerbi, 2023) may be specific to the US media environment or to the Republican party's dis-

tinctive relationship with misinformation during the study period, and not a general property of conservative cognition. If norm enforcement is symmetric in Argentina but asymmetric in the United States, the asymmetry is likely institutional rather than psychological.

Third, the attention moderation results provide micro-level evidence for dual-process accounts of misinformation processing (Pennycook and Rand, 2021). Response latency amplifies epistemic-norm pruning but not partisan-directional pruning, consistent with the view that detecting factual violations requires analytic engagement while processing identity-based cues is more automatic. Future research could extend the finding using eye-tracking or behavioral measures of engagement that do not rely on response time as a proxy.

Fourth, the friendship paradox formalization connects individual pruning decisions to network-level dynamics. In the misinformation conjoint, partisan users prune more, expect to be pruned more, and may therefore prune preemptively, generating a positive feedback that could sustain network sorting independently of algorithmic curation. The point has direct implications for platform design: reducing algorithmic amplification of misinformation may be insufficient to reverse homophily if user-side pruning expectations are themselves endogenous to the network’s partisan composition.

Our study is, of course, not without limitations. First, our survey design by construction captures what respondents say they would do when exposed to the experimental stimuli (stated intention), not what they actually do on their social media platforms (actual behavior). The relationship between pruning intentions and pruning behavior is poorly understood, and there is evidence from related domains (content sharing and political participation) that stated preferences overestimate behavioral rates (Mosleh et al., 2021). Future work linking stated-preference conjoint data to behavioral records from digital trace studies would provide a stronger test of

the mechanisms identified here and allow researchers to calibrate the magnitude of the intention-behavior gap in the pruning domain. All claims in this paper should be read as claims about pruning intentions.

Second, the post-hoc status of the two-mechanism framework should be kept in mind. Each individual hypothesis was pre-registered or explicitly flagged as a post-registration extension, but the framework that organizes them was developed after we observed the cross-wave contrast. The framework is a structured interpretation of pre-registered findings rather than a confirmed mechanism; replication in a new sample with the framework pre-specified would strengthen its status. Relatedly, the two conjoints were fielded in the same panel two months apart, and we cannot rule out learning or spillover effects between waves. Our main results depend on within-wave comparisons rather than between-wave comparisons of the same respondents, which limits the scope of potential spillover bias.

Third, the nonpartisan comparison group is compositionally heterogeneous. We define non-partisans as respondents with no stated party preference or blank vote intention, a category that combines genuinely apolitical citizens, ideological moderates, and strategic nonpartisans who hold political views but decline to express them. These subgroups may respond to political and misinformation content in systematically different ways, and our design does not allow us to decompose them.

Fourth, the Argentine context offers a clean two-bloc party system and high issue salience for both treatment domains, conditions that may favor sharp empirical separation of the two mechanisms. In multiparty systems, where ideological cues are more ambiguous and partisan identities less consolidating, the partisan-directional mechanism may produce weaker frame effects. In lower-salience political contexts, the epistemic-norm mechanism may generate weaker

unconditional pruning rates. On platforms with different information architectures, especially those that make content authorship less visible, the frame elements we manipulate may carry less signal. Cross-national replication holding the conjoint stimuli constant while varying the political context would help adjudicate context-specific from universal components of the framework.

Networks formed through partisan-directional pruning and epistemic-norm pruning differ in structure, arise through different processes, and require different policy interventions. Partisan-directional pruning produces ideologically sorted networks in which the ideological signal of a contact's shared content determines whether the tie survives; addressing it requires reducing the salience of ideological cues in political content or increasing tolerance for cross-cutting exposure. Epistemic-norm pruning produces networks from which misinformation-spreading contacts are broadly removed across partisan lines; addressing it requires reducing the prevalence of norm-violating content through debunking, inoculation, and platform-level misinformation governance. Conflating the two, by treating all pruning as a symptom of partisan polarization or all misinformation-driven unfriending as a manifestation of partisan motivated reasoning, will misdiagnose the problem and misallocate the intervention.

## **Data Availability**

The data and replication code will be made available at the Harvard Dataverse prior to publication.

## **Competing Interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## **Disclaimer**

Generative AI tools (Claude Opus 4.5) were used to assist with R code troubleshooting, LaTeX formatting, and copy-editing. The authors did not use generative AI to produce the substantive analysis, theoretical argument, or empirical conclusions of this article. The authors reviewed and verified all AI-assisted output and take full responsibility for the content.

## **References**

- Altay, S. and Acerbi, A. (2023). Misinformation on social media: Homogeneous and heterogeneous effects. *Current Opinion in Psychology*, 54:101711.
- Anspach, N. M. (2017). The new personal influence: How our facebook friends influence the news we read. *Political communication*, 34(4):590–606.
- Aruguete, N., Calvo, E., and Ventura, T. (2021). Network activated frames: Content sharing and perceived polarization in social media. *Working Paper*.

- Bakshy, E., Messing, S., and Adamic, L. (2015). Exposure to ideologically diverse news and opinion on facebook. *science*, 348(6239):1130–1132.
- Banks, A., Calvo, E., Karol, D., and Telhami, S. (2021). # polarizedfeeds: Three experiments on polarization, framing, and social media. *The International Journal of Press/Politics*, 26(3):609–634.
- Barberá, P. (2020). Social media, echo chambers, and political polarization. *Social Media and Democracy: The State of the Field, Prospects for Reform*, page 34.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542.
- Berinsky, A. J., Margolis, M. F., and Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, 58(3):739–753.
- Bode, L. (2016). Pruning the news feed: Unfriending and unfollowing political content on social media. *Research & Politics*, 3(3):2053168016661873.
- Centola, D. (2018). *How Behavior Spreads: The Science of Complex Contagions*. Princeton University Press, Princeton, NJ.
- Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., and Flammini, A. (2011). Political polarization on twitter. In *Fifth international AAAI conference on weblogs and social media*.

- Eom, Y.-H. and Jo, H.-H. (2015). Tail-scope: Using friends to estimate heavy tails of degree distributions in large-scale complex networks. *Scientific Reports*, 5:09752.
- Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, 96(6):1464–1477.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, CA.
- Franchino, F. and Zucchini, F. (2015). Voting in a multi-dimensional space: A conjoint analysis employing valence and ideology attributes of candidates. *Political Science Research and Methods*, 3(2):221.
- Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4):1799–1839.
- Guess, A. M., Barberá, P., Munzert, S., and Yang, J. (2021). The consequences of online partisan media. *Proceedings of the National Academy of Sciences*, 118(14).
- Hainmueller, J. and Hopkins, D. J. (2015). The hidden american immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, 59(3):529–548.
- Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political analysis*, 22(1):1–30.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. (2019). The origins

- and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22:129–146.
- Jackson, M. O. (2019). *The Human Network: How Your Social Position Determines Your Power, Beliefs, and Behaviors*. Pantheon Books, New York.
- Kaiser, J., Vaccari, C., and Chadwick, A. (2022). Partisan blocking: Biased responses to shared misinformation contribute to network polarization on social media. *Journal of Communication*, 72(2):214–240.
- Kirkland, P. A. and Coppock, A. (2018). Candidate choice without party labels. *Political Behavior*, 40(3):571–591.
- Klašnja, M., Lupu, N., and Tucker, J. A. (2021). When do voters sanction corrupt politicians?—corrigendum. *Journal of Experimental Political Science*, pages 1–2.
- Kwak, H., Chun, H., and Moon, S. (2011). Fragile online relationship: a first look at unfollow dynamics in twitter. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1091–1100.
- Leeper, T. J., Hobolt, S. B., and Tilley, J. (2020). Measuring subgroup preferences in conjoint experiments. *Political Analysis*, 28(2):207–221.
- Lerman, K., Yan, X., and Wu, X.-Z. (2016). The” majority illusion” in social networks. *PloS one*, 11(2):e0147617.
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., and Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3):106–131.

- Mares, I. and Visconti, G. (2020). Voting for the lesser evil: evidence from a conjoint experiment in romania. *Political Science Research and Methods*, 8(2):315–328.
- Marwick, A. and Lewis, R. (2017). Media manipulation and disinformation online. Technical report, Data & Society Research Institute.
- Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.
- Messing, S. and Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication research*, 41(8):1042–1063.
- Mosleh, M., Martel, C., Eckles, D., and Rand, D. G. (2021). Shared partisanship dramatically increases social tie formation in a twitter field experiment. *Proceedings of the National Academy of Sciences*, 118(7).
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., and Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind political fake news sharing on twitter. *American Political Science Review*, pages 1–17.
- Pennycook, G. and Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5):388–402.
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., and van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10):201199.
- Sasahara, K., Chen, W., Peng, H., Ciampaglia, G. L., Flammini, A., and Menczer, F. (2021).

- Social influence and unfollowing accelerate the emergence of echo chambers. *Journal of Computational Social Science*, 4(1):381–402.
- Settle, J. E. (2018). *Frenemies: How social media polarizes America*. Cambridge University Press.
- Slothuus, R. and De Vreese, C. H. (2010). Political parties, motivated reasoning, and issue framing effects. *The Journal of Politics*, 72(3):630–645.
- Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of communication*, 60(3):556–576.
- Sunstein, C. R. (2018). *# Republic*. Princeton university press.
- Taber, C. S. and Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3):755–769.
- Tokita, C. K., Guess, A. M., and Tarnita, C. E. (2021). Polarized information ecosystems can reorganize social networks via influence operations. *Proceedings of the National Academy of Sciences*, 118(50):e2102147118.
- Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLoS one*, 13(9):e0203958.
- Torres-Lugo, C., Yang, K.-C., and Menczer, F. (2022). The manufacture of partisan echo chambers by follow train abuse on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 1017–1028.
- Watts, D. J. and Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4):441–458.