Political Research Quarterly
2026, Vol. 0(0) 1–14

DOI: 10.1177/10659129261454646
journals.sagepub.com/home/prq
Sage

Partisan Hearts, Polarized Minds? Revisiting the Relationship Between Partisanship and Polarization

Alexa Bankert[1] and **Joshua N. Zingher**[2]

Abstract

Existing scholarship has increasingly recognized that strong partisanship does not inevitably produce out-party hostility—yet the scope and consistency of this decoupling across multiple measures of affective polarization remains underexamined. In this study, we revisit the relationship between partisan strength and affective polarization using original national survey data across three established measures: feeling thermometers, trait attributions, and social distance. We provide systematic, multi-measure evidence that the empirical linkage between partisan strength and out-party hostility is weaker and more heterogeneous than much of the literature implies. Substantial variation in out-party hostility exists at every level of partisan strength—even among the strongest partisans. We further show that ideological and issue-attitude strength meaningfully predict which strong partisans become most hostile toward the out-party. These results extend prior work by demonstrating that the decoupling of partisanship and hostility is robust across measurement approaches, and they highlight the importance of distinguishing partisanship's constructive democratic functions from its corrosive forms.

Keywords

polarization, partisanship, out-party hostility, measurement

Partisanship is the lifeblood of democratic politics, structuring competition, organizing governance, and providing citizens with a sense of political identity (Rosenblum 2008). Yet, in the United States, partisanship appears to have taken on a more troubling form: one fueled not just by loyalty to one's own party but by deep-seated animosity toward the other. Scholars have termed this phenomenon *affective polarization*, describing a growing distrust, contempt, and even hatred between opposing partisans (Iyengar et al. 2012; Mason 2018). Prior research identifies strong partisan identity as a key correlate of affective polarization. Scholars have consistently shown that stronger party identification predicts greater hostility toward the out-party, with some arguing that partisanship has become the primary lens through which Americans evaluate political opponents (Iyengar et al. 2019; Iyengar and Krupenkin 2018). Yet recent work has begun to complicate this picture. Bankert (2021) demonstrated that positive and negative partisanship are empirically separable phenomena: many partisans exhibit deep in-party attachment without corresponding out-party animus. Similarly, Mason (2015, 2018) shows that partisan sorting—the alignment of partisan, ideological, and social identities—rather than partisan strength alone, is what amplifies out-party hostility; absent that alignment, even strong partisans tend to display more moderate attitudes toward the opposing party. Together, these contributions have meaningfully refined our understanding of when and why partisanship produces animosity.

What remains less well understood is how consistently this decoupling appears across different measurement approaches to affective polarization. Studies of affective polarization rely variously on feeling thermometers, trait

[1]Department of Political Science, School of Public and International Affairs, University of Georgia, Athens, GA, USA
[2]Department of Political Science and Geography, Old Dominion University, Norfolk, VA, USA

Corresponding author:
Joshua N. Zingher, Department of Political Science and Geography, Old Dominion University, 7000 Batten Arts & Letters, Norfolk, VA 23529, USA.
Email: jzingher@odu.edu

attribution batteries, and social distance scales—measures that, as we show, do not behave identically in their relationship to partisan strength. Whether the weak link between partisan strength and out-party hostility documented by Bankert and implied by Mason's sorting logic holds uniformly across these measures, and whether the individual-level predictors of hostility among strong partisans are similarly consistent, are questions that have not yet received systematic attention. This paper addresses that gap.

Using an original dataset with multiple measures of affective polarization and partisanship, we qualify and build on prior findings: first, we find that the relationship between partisanship and polarization is weaker than expected and it varies significantly across different measures of polarization; second, substantial variation in out-party hostility exists at every level of partisan strength—even among the strongest partisans; and third, ideological and issue-attitude strength meaningfully predict which strong partisans become most hostile toward the out-party. We conclude with suggestions for future research to extend and deepen our current understandings of affective polarization as well as the normative significance of partisanship in a democracy.

The (Conditional) Relationship Between Partisanship and Affective Polarization

The rise of affective polarization—the tendency to dislike, distrust, and even dehumanize the out-party—is a central concern in contemporary political science (see, e.g., Iyengar et al. 2012; Cassese 2021; Druckman et al. 2024, but also Broockman et al. 2023) as it has significant implications for democratic norms, including tolerance for opposing viewpoints, willingness to compromise, and trust in political institutions (Campos and Federico 2025; Iyengar et al. 2019; Kalmoe and Mason 2022; Mason 2018). While numerous studies highlight a wide range of factors contributing to the rise of affective polarization—including changes in the media environment (Lelkes et al. 2017; Levendusky 2013; Stroud 2011), partisan-ideological realignment (Abramowitz and Saunders 2008; Levendusky 2009; Mason 2015; Zingher 2022), and even growing income inequality (McCarty et al. 2006)—partisanship has reliably been identified as a marker and a driver of affective polarization. This pattern is consistent with the underlying nature of affective polarization. Unlike policy polarization, which arises from disagreement over specific issues, affective polarization is rooted in group dynamics and social identity processes (Tajfel and Turner 2004). As classic social identity research demonstrates, even minimal forms of group affiliation can evoke "both positive feelings toward the in-group and negative evaluations of the out-group" (Tajfel et al. 1971). Yet the relationship between in-group identification and out-group hostility is considerably more complex. In a foundational contribution, Brewer (1999) directly interrogates the assumption that loving one's in-group and hating the out-group are two sides of the same psychological coin. Drawing on decades of intergroup research, Brewer demonstrates that in-group favoritism—the tendency to prefer and positively evaluate one's own group—does not reliably or necessarily generate hostility toward out-groups. Instead, she argues that in-group love and out-group hate are empirically separable motivational orientations: the former is rooted in positive attachment, belonging, and identity affirmation, while the latter requires an additional psychological ingredient—typically perceived threat, status competition, or zero-sum framing of intergroup relations. In-group bias, Brewer contends, most commonly takes the form of preferential treatment of one's own group rather than active derogation of others; out-group hostility is the exception, not the automatic byproduct of group identification. Brewer (2007) extends this argument, showing that the human need to belong and affiliate does not inherently produce antagonism toward those outside the group.

This distinction is theoretically foundational for the present paper because it speaks to the conclusion of many prior studies that strong partisanship is coinciding with or even fueling affective polarization. For example, in their comprehensive review, Iyengar and colleagues (2019) argue that affective polarization is rooted in partisanship as a social identity rather than policy disagreement alone—partisan identity, not issue positions, is what leads citizens to evaluate political opponents in emotional, us-versus-them terms (see also Iyengar and Krupenkin 2018). These findings frame affective polarization as a largely partisan identity-based phenomenon. It is therefore unsurprising that scholars have likened modern partisans to fervent sports fans. Mason (2015) finds that strongly identified partisans derive emotional satisfaction from their party's victories and even pleasure from the opposing party's losses—reflecting a tribal, in-group versus out-group mentality. Extending this analogy beyond the U.S. context, Reiljan (2020) shows that partisans in many democracies display "extremely hostile" attitudes toward rival parties, driven by strong identifiers who behave like fans cheering for their "home team" and expressing animosity toward the opposing "team." In short, these studies show that the more fervently people identify with a political party, the more likely they are to exhibit in-party favoritism and out-party loathing, whether in the United States or in other democracies around the world.

Yet, consistent with Brewer's argument, the relationship between partisanship and affective polarization is not entirely straightforward. Research consistently shows that partisans' feelings toward their own party have remained relatively stable over time, while their feelings toward the opposing party have grown markedly more negative (Abramowitz and Webster 2016; Iyengar et al. 2019). This pattern suggests that out-party hostility can intensify independently of in-party attachment. This is precisely the distinction that makes the partisan case so theoretically interesting. In the American two-party system, the structural opposition between parties makes it unusually easy for in-party attachment to shade into out-party animosity—yet prior research underscores that attachment to one's own party and hostility toward the opposing party are analytically distinct phenomena. For example, Bankert (2024) shows that intense party loyalty can exist without a deep hatred of the other side, and vice versa. This decoupling means partisans can feel deeply attached to their own party without feeling equally deep disdain toward their opponents. Such evidence complicates the simple narrative that stronger partisanship inherently produces stronger animosity.

If strong partisanship alone does not inevitably produce out-party hostility, what drives the more toxic forms of partisan animus? Martherus and colleagues (2021) examine one especially troubling expression of partisan hostility: the dehumanization of political opponents, or viewing them as less than fully human. One might assume that the strongest partisans would be the most likely to dehumanize their rivals. However, their findings show that the standard measure of partisan strength has only a weak relationship with dehumanization, while affective polarization—the degree to which partisans dislike the out-party—is substantially more predictive of it. Indeed, the tendency to dehumanize appears to increase sharply only among the most extremely affectively polarized partisans. The key insight is that it is the intensity of out-party hostility, rather than partisan strength per se, that is most closely associated with the most extreme expressions of partisan animosity.

Mason (2015, 2018) further demonstrates that strong partisanship alone is insufficient to produce partisan animus. When a partisan's identities are cross-cutting or misaligned—for example, a strong partisan with a moderate or mixed ideology—their level of affective polarization is substantially lower. As Mason notes, "partisanship in the absence of a strongly aligned ideological identity is… a far less potent contributor to in-group bias than a highly sorted partisan and ideological identity" (Mason 2015, 136). In other words, out-party hostility intensifies when partisan and ideological identities align, reinforcing each other to create a fused "mega-identity." Without this alignment, even strong partisans tend to display more moderate and civil attitudes toward the opposing party.

Taken together, this body of work has already substantially qualified the view that partisan strength and out-party hostility are intrinsically linked. Our contribution is not to establish this decoupling anew, but to extend and sharpen it in three ways. First, we examine the relationship across three distinct measures of affective polarization simultaneously, allowing us to assess whether the partisan strength–hostility gap is consistent or measure-dependent. Second, we document within-group heterogeneity in hostility at every level of partisan strength, providing distributional evidence that goes beyond mean-level comparisons. Third, we identify the demographic and political variables—including ideological strength and issue-attitude intensity—that predict which strong partisans become hostile.

Hypotheses

Building on the theoretical framework outlined above, we develop three predictions about the relationship between partisan strength and affective polarization.

We first expect that *partisan strength and measures of affective polarization are only weakly correlated (H1)*.

If the correlation between partisanship and affective polarization is weak, it follows that individuals with the same level of partisan strength—from weak to strong—can differ substantially in their degree of affective polarization. We thus expect to find that *levels of affective polarization vary significantly within each level of partisan strength (H2)*.

Moreover, if strong partisanship can exist without affective polarization, then levels of polarization among strong partisans should be systematically related to individual characteristics and attitudes. This reasoning leads to our final hypothesis: *stronger expressions of affective polarization among strong partisans should be associated with distinct demographic and attitudinal profiles (H3)*.

Data & Measures

To test hypotheses 1–3, we utilize a sample of 1,003 American respondents. The sample, collected by Bovitz Inc. in August 2024, broadly reflects the U.S. population in terms of key demographics (see sample profile in Table A1 in the Appendix). 34 percent of our sample consists of Republicans ($N = 339$) and 51 percent consists of Democrats ($N = 512$). These numbers include partisan leaners. Pure independents make up 15 percent of the sample ($N = 152$) and are excluded from our analyses since they do not identify with either political party. To demonstrate the robustness of our results, we replicate

most of our analyses with a second sample, also collected by Bovitz Inc. in September 2025 (see sample profile in Table A2 in the Appendix).[1]

Our data set includes all three commonly used measures of polarization: the feeling thermometer scales to assess both parties, followed by the battery of trait assignments for both the in-party and the out-party as well as a scale to gauge social distance towards the out-party. All three measures share a core objective: assessing how individuals perceive and differentiate between members of their own party and those of the opposing party. Scholars differ, however, in how they operationalize affective polarization using these measures. Some calculate affective polarization as the difference between in-party and out-party evaluations (e.g., Iyengar et al. 2012), while others focus exclusively on out-party evaluations (e.g., Druckman et al. 2024). The difference-score approach combines in-party warmth and out-party hostility into a single metric, which can obscure their distinct contributions and mask potential asymmetries—for instance, when high polarization scores result from extreme out-party dislike even as in-party feelings remain moderate, or vice versa. In our analyses, we focus solely on out-party evaluations for two reasons: Changes in affective polarization are driven primarily by increasingly negative evaluations of the out-party rather than more positive assessments of the in-party (Abramowitz and Webster 2016). Because in-party attitudes remain relatively stable over time, focusing on out-party hostility is—in our opinion—the most suitable approach for explaining variation in affective polarization. Moreover, from a normative perspective, affective polarization stemming primarily from strong in-party favoritism and only moderate out-party dislike is less concerning than polarization driven by intense out-party animosity. Across all our analyses, out-party feeling thermometer ratings, trait attributions, and social distance serve as our measures of affective polarization.

Our second key variable is partisan strength, which we measure using two distinct indicators. The first one, taken from the *American National Election Studies (ANES)*, is the conventional 3-point strength measure that divides partisans into three categories: partisan leaners, weak partisans, and strong partisans. The second one is a measure of partisan identity strength (Bankert et al. 2017). This multi-item scale explicitly measures partisanship as an identity by assessing the centrality of the in-party to respondents' self-image.

Finally, for our observational analyses, we include a range of standard demographic variables. These include race (coded 1 for White respondents, 0 otherwise), ethnicity (1 for Hispanic respondents, 0 otherwise), and gender (1 for female, 0 for male). We also incorporate education, coded on a continuous scale from 0 to 1, where 0 represents no formal schooling and 1 represents a Ph.D. Religious attendance is similarly scaled from 0 to 1, with 0 indicating never attending services and 1 representing attendance more than once a week. Finally, we also include a measure of self-reported ideology—where 0 corresponds to self-identified moderates and 1 to respondents who identify as either extremely liberal or extremely conservative—alongside preference intensity on a range of political issues, including abortion, immigration, and general government involvement. These questions (and their phrasing) were taken directly from the ANES.

We conduct all analyses separately using both measures of partisan strength, as the conventional three-point partisanship strength item cannot distinguish between partisanship as policy preference and partisanship as social identity. Nevertheless, we include the conventional measure due to its widespread use and to examine potential measurement-specific effects. All survey items and response options are provided in the Appendix.

Analysis #1: Correlations Between Partisan Strength and Measures of Affective Polarization

To examine the relationship between partisan strength and affective polarization, we begin with a simple correlation analysis between the 3-point partisan strength item and the three affective polarization (AP) measures: feeling thermometer ratings, trait evaluations, and social distance toward the out-party, each of which has been recoded to range from low to high, with higher values indicating more negative sentiments toward out-partisans. Similarly, higher values on the partisan strength item indicate stronger party attachments. All variables are re-scaled to range from 0 to 1 for better comparability.

On average, partisans rate supporters of the out-party quite negatively, with a mean score of 0.76 (SD = 0.22) on the reversed feeling thermometer scale—equivalent to 0.24 on the original 0–100 scale. Out-partisans also receive generally unfavorable personality assessments, with an average score of 0.66 (SD = 0.22) on a 0–1 scale (e.g., perceived as untrustworthy or close-minded). Despite these negative evaluations, the social distance scale shows a relatively low mean value of 0.36 (SD = 0.23), suggesting that negative perceptions of out-partisans—whether in terms of affect or personality—do not necessarily translate into a strong desire to avoid social or professional interactions with them.

Indeed, all three measures correlate at varying levels (see Table 1): Feeling thermometer scores for the out-party correlate with out-party traits at 0.62, while the

social distance measure correlates with these two measures at less than 0.30—suggesting that despite reaching conventional levels of statistical significance, these measures share relatively little variance and may not gauge the same underlying construct. Yet, what is most remarkable is their correlation with partisan strength: Across all three measures—FTs, traits, and social distance—we find a surprisingly weak relationship to partisan strength ranging from 0.11 to 0.15. Although these coefficients are statistically significant, they account for no more than roughly 2 percent of the variance in any of the three measures.[2]

We obtain similar results when utilizing the multi-item partisan identity scale which—due to its continuous nature—shares more measurement properties with the commonly used measures of affective polarization (see Table 2).

Compared to the 3-point single-item strength measure, the partisan identity scale is more strongly related to the social distance measure but markedly more weakly related to feeling thermometer values for the out-party. The correlation between out-party FTs and partisan identity strength is 0.04—statistically indistinguishable from zero. One possibility is that this near-zero correlation is driven by inattentive respondents—a concern that is particularly acute in online, opt-in survey samples. To assess this, we conduct two robustness checks. First, we restrict the sample to respondents who report being at least "slightly interested" in government and politics, on the grounds that politically disengaged respondents may answer the partisan identity scale and the feeling thermometer inconsistently, not because the constructs are unrelated but because low-engagement respondents anchor their responses differently across batteries. Among this subsample, the correlation between partisan identity strength and out-party FTs remains almost unchanged at 0.05, suggesting that general political disengagement alone does not account for the null result. Second, we use response timing data to identify respondents who spent more time than the sample average (7.8 seconds) on an unrelated survey instrument, as a proxy for attentiveness. Among this subsample, the correlation rises more substantially, to 0.17—still modest, but no longer negligible, and broadly consistent with the correlations observed using the 3-point strength item. This pattern suggests that inattentive responding does partially attenuate the observed relationship. At the same time, we do not think measurement noise provides a complete explanation. Even among attentive respondents, the correlation between partisan identity and out-party FTs remains substantially weak. This persistent gap across two robustness checks points toward a theoretically meaningful finding that aligns with Brewer's theoretical framework. The partisan identity scale—which includes items such as whether criticism of one's party feels like a personal insult and whether one feels connected to fellow partisans—captures the self-definitional core of partisan identity. This reflects precisely the "in-group love" dimension that Brewer argues does not necessarily translate into out-group negativity, such as that measured by feeling thermometers.[3] Overall, these results align with our first prediction that partisan strength and measures of affective polarization are only weakly correlated. We find similarly weak correlations in our replication analyses (see Table A3 in the Appendix). Critics might be concerned that these patterns are mostly driven by our opt-in survey samples. To address this concern, we replicate our analyses one final time, using data from the *American National Election Study*. Reassuringly, our results remain robust. Figure A1, which plots the correlation between the three-point partisan strength item and out-party feeling thermometer scores using ANES cumulative data, yields substantively identical conclusions: across several decades of electoral cycles, partisan strength accounts for only a small share of the variation in out-party hostility, ranging from a correlation of 0.14 in 1980 to 0.31 in 2024. While a maximum correlation of 0.31 is stronger than what we observe in our original samples, the average correlation of 0.26 across the full ANES time series nonetheless confirms that partisan strength and affective polarization share only a weak empirical relationship.

Analysis #2: Affective Polarization Variation Within Levels of Partisan Strength

We next examine to what extent affective polarization varies within each level of partisan strength. This analysis

Table 1. Correlation between partisan strength and AP measures

	Partisan strength	Out-party FTs	Out-party traits	Out-party social distance
Partisan strength	(–)	(–)	(–)	(–)
Out-party FTs	0.11*	(–)	(–)	(–)
Out-party traits	0.15*	0.62*	(–)	(–)
Out-party social distance	0.15*	0.27*	0.29*	(–)

Note. Entries are correlation coefficients. All variables range from 0 to 1. * denotes statistical significance at $p < 0.05$. $N = 849$.

Table 2. Correlation between partisan identity strength and AP measures

	Partisan identity strength	Out-party FTs	Out-party traits	Out-party social distance
Partisan identity strength	(−)	(−)	(−)	(−)
Out-party FTs	0.04	(−)	(−)	(−)
Out-party traits	0.14*	0.62*	(−)	(−)
Out-party social distance	0.29*	0.27*	0.29*	(−)

Note. Entries are correlation coefficients. All variables range from 0 to 1. * denotes statistical significance at $p < 0.05$. $N = 839$.

allows us to answer two questions. First, even among equally strong partisans, do we see wide variation in affective polarization? And second, do we see similar variation among partisan leaners and weak partisans? In our data, less than half of all respondents (47 percent) consider themselves 'strong partisans' ($N = 400$), while 31 percent are weak partisans ($N = 262$) and 22 percent are partisan leaners ($N = 189$). In Figure 1, we plot the distribution of all three polarization measures within each level of partisan strength.

Across all three measures, strong partisans exhibit the highest levels of out-party negativity. They report significantly more negative thermometer scores for the out-party (M = 0.80) than weak partisans (M = 0.68, t (659) = 6.60, $p < 0.001$) and partisan leaners (M = 0.76, t (586) = 2.08, $p < 0.01$). They also assign significantly more negative personality traits to out-partisans (M = 0.71) than weak partisans (M = 0.58, t (658) = 7.55, $p < 0.001$) and partisan leaners (M = 0.66, t (586) = 3.06, $p < 0.001$), and they desire much higher levels of social distance from the out-party (M = 0.41) than weak partisans (M = 0.29, t (659) = 6.84, $p < 0.001$) and partisan leaners (M = 0.35, t (586) = 3.03, $p < 0.001$).

While we might expect a linear increase in negative sentiments toward the out-party as partisan strength increases, our results paint a more nuanced picture. For example, weak partisans report significantly more positive FT scores for the out-party than partisan leaners (t (449) = −3.61, $p < 0.001$). They also tend to assign the out-party significantly fewer negative personality traits than partisan leaners (t (448) = −3.57, $p < 0.001$) and report significantly lower levels of social distance to out-partisans than partisan leaners (t (449) = −3.03, $p < 0.001$). This lack of linearity indicates that increases in partisan strength do not translate into uniform changes in out-party negativity. One possible interpretation is that leaners, precisely because their partisan identity is more ambivalent, may be more attuned to the out-party as a negative reference point—relying on out-party dislike as a partial substitute for the in-party attachment they lack. However, we do not have the data to adjudicate this explanation directly, and we flag this non-linearity as a puzzle warranting further investigation.

Despite these clear differences in central tendencies,[4] the boxplots reveal substantial overlap in the distributions across partisan-strength categories, underscoring the considerable heterogeneity that exists within each group. For instance, even among strong partisans—who have the highest median levels of out-party negativity—the boxplots show wide dispersion: some strong identifiers rate the out-party extremely negatively, while others provide mid-range or even relatively moderate evaluations. This spread indicates that strong partisans are not uniformly hostile toward the out-party, despite having the highest average levels of negativity. Thus, partisan strength is not always a clear-cut predictor of how intensely someone feels about the out-party.

Finally, the amount of variation among strong partisans is broadly comparable to the variation observed among weak partisans and leaners. For out-party feeling thermometer scores, the standard deviations across the three partisan-strength groups are nearly identical (SD = 0.21–0.22). The same pattern appears for out-party trait ratings (SD = 0.21–0.23) and for social distance (SD = 0.22–0.24). These values indicate that while the three partisan-strength groups differ in their average levels of out-party negativity, the range of individual responses within each group is similarly wide—suggesting that substantial variation in out-party negativity is a feature of every level of partisan strength. To ensure, once again, that these results are not just an artifact of our opt-in sample, we replicate the within-group variation analysis using ANES cumulative data which yields broadly consistent results: substantial heterogeneity in out-party feeling thermometer scores persists at every level of partisan strength (see Figure A2). Finally, to confirm that our results are not driven by our measure of partisan strength, we replicate our analysis, using the partisan identity scale instead of the 3-point strength item. In our sample, partisan identity scores an average value of 0.41 on a 0–1 scale. We split the scale's distribution into quartiles and examine levels and variation of affective polarization measures within each quartile of partisan identity strength (Figure 2).

We find that the boundaries between identity-strength quartiles are even less distinct than those observed with the 3-point partisan-strength measure. For instance, partisans in the highest quartile report out-party FT scores that are almost identical to those in the third quartile (M =

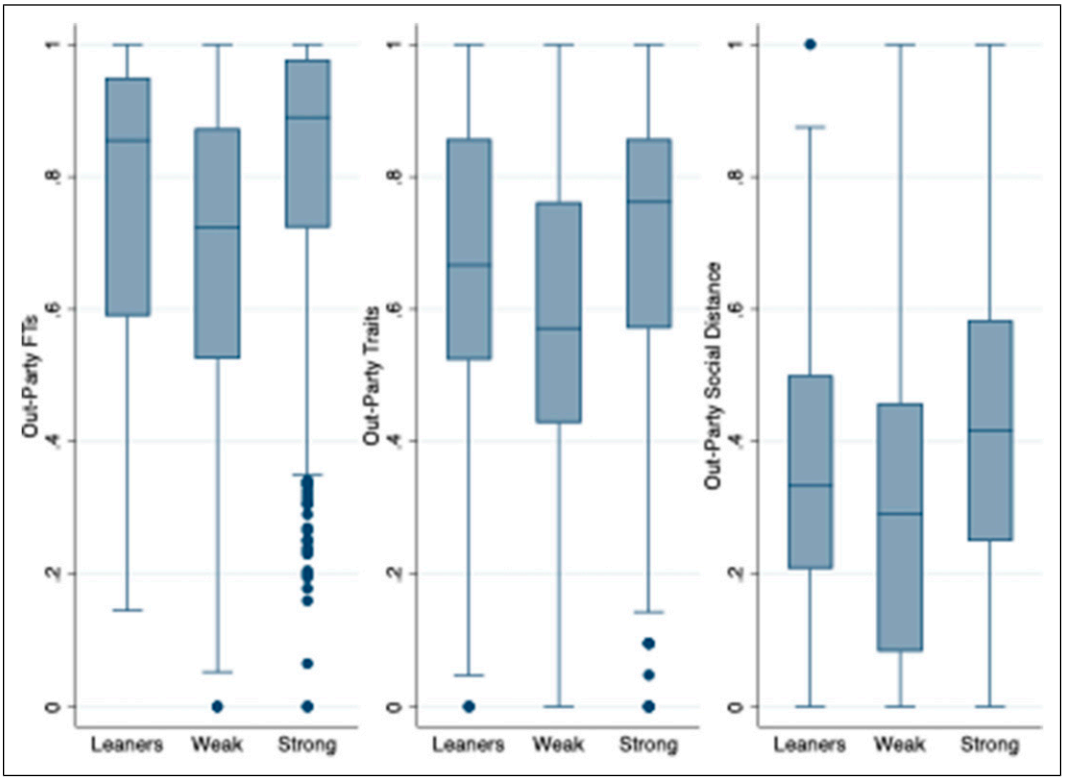

Figure 1. Variation in AP measures within levels of PID strength. *Note*. Box plots display the variation in three AP measures for each of the three levels of partisan strength. The boxes show the interquartile range (middle 50 percent of the data), with the median indicated by the line inside the box. The whiskers indicate the range, and the dots represent outliers. All variables range from 0 to 1. Higher values indicate more negative sentiments for the out-party.

0.78 vs. 0.76, t (276) = 0.59, p = 0.27), and the same pattern holds for out-party trait ratings (M = 0.69 vs. 0.70, t (276) = −0.66, p = 0.74). Only for social distance do we observe a more noticeable drop-off, with the strongest partisans reporting higher social distance than those in the third quartile (M = 0.49 vs. 0.42, t (276) = 2.33, p = 0.01). Overall, the relationship between partisan identity strength and out-party hostility remains weakly patterned and far from linear, echoing the results from the 3-point measure. Moreover, within-quartile variation is substantial—standard deviations range from 0.21 to 0.26 on a 0–1 scale across all three measures—and these within-group values are nearly identical to the overall standard deviation of each measure. This indicates that most of the variation in out-party hostility occurs within, rather than between, levels of partisan identity strength.

Analysis #3: Predictors of Affective Polarization Among Strong Partisans

The preceding analyses establish that strong partisans vary considerably in their levels of out-party hostility—but they do not explain why some strong partisans become hostile while others do not. Mason's (2015, 2018) identity alignment logic predicts that strong ideological and issue-based attitudes should amplify hostility among strong partisans—particularly on policy domains tightly bundled with partisan identity, such as abortion and government services. This leads to the expectation that, in our analyses, ideological and issue-attitude strength should meaningfully predict which strong partisans become hostile toward the out-party.

To test our prediction, we construct three dependent variables—one for each polarization measure—restricted to respondents who identify as strong partisans. For instance, the "Out-Party FTs" variable captures feeling thermometer ratings of the out-party among strong partisans. As independent variables, we include ideological strength and attitude strength on issues such as government services, abortion, and immigration. We also include a standard set of demographic controls, which we treat as exploratory given the absence of strong theoretical priors about which demographic characteristics should predict hostility among strong partisans.

We display our results in coefficient plots in Figure 3.

Across the regression models, a similar pattern emerges for the feeling-thermometer and trait-based

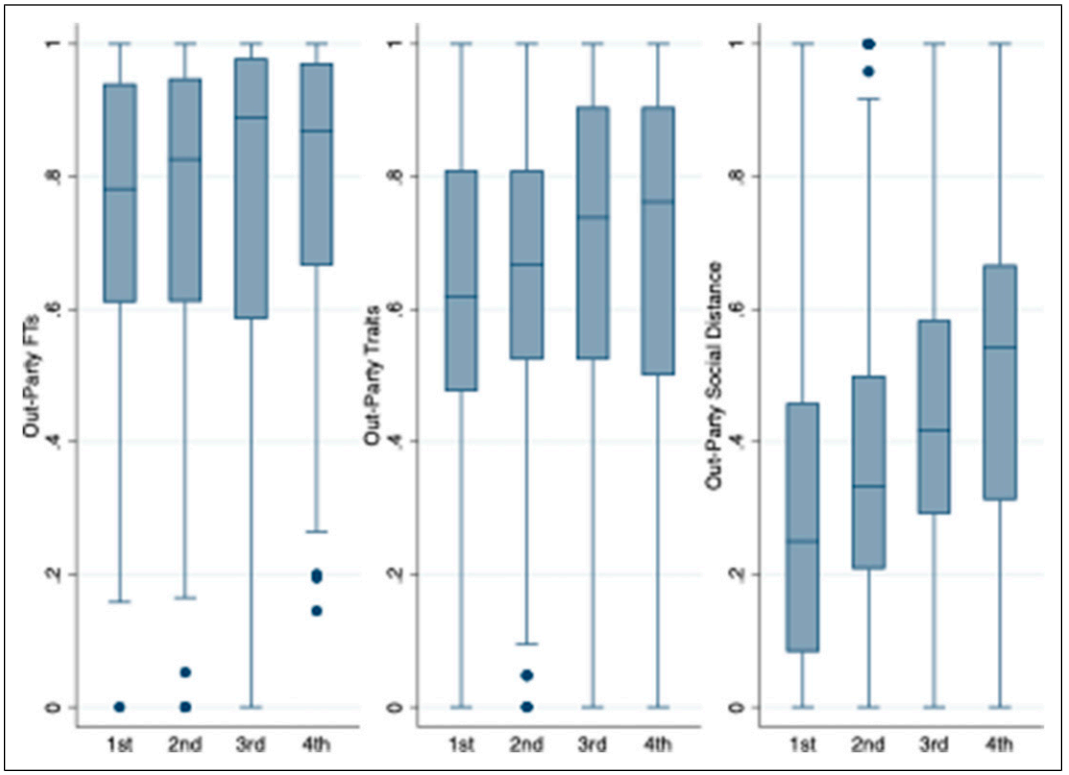

Figure 2. Variation in AP measures within levels of partisan identity strength. *Note*. Box plots display the variation in AP measures for each quartile of partisan identity strength. The boxes show the interquartile range (middle 50 percent of the data), with the median indicated by the line inside the box. The whiskers indicate the range, and the dots represent outliers. All variables range from 0 to 1. Higher values indicate more negative sentiments for the out-party.

polarization outcomes: In both models (Columns 1 and 2), ideological strength significantly increases the likelihood that strong partisans are polarized: Moving from the minimum to maximum of ideological strength predicts a 0.12 increase in out-party thermometer values and a 0.07 increase in negative trait assignment, holding other variables constant. Since the outcome measures range 0–1, this represents about a 12 and 7 percentage point shift across the full scale of out-party thermometer and trait ratings, respectively—modest but meaningful increases. At the same time, religious attendance exerts the largest effect in the model in absolute terms. Moving from never attending religious services to attending more than once a week predicts a 0.16 decrease in out-party thermometer scores and a 0.19 decrease in negative out-party traits—about a 16 and 19 percentage point shift, respectively. Ethnicity, gender, and higher education levels show meaningful effects only in the model predicting out-party thermometer values. Ideological strength also emerges as a significant predictor in the social distance model—comparable in magnitude to its effect in the trait model—in addition to race (Column 3). We replicate these models but replace ideological strength with strength of attitudes on abortion, immigration, and government services (see Appendix for question wording) to examine whether concrete issue strength—versus more diffused and potentially identity-related measures of ideological strength—pushes partisans into hostile territory.

Across the three outcomes, the inclusion of issue-specific strength produces a more fragmented pattern than the earlier model with ideological strength. Attitude strength on issues like government services and abortion predicts higher out-party negativity in the feeling-thermometer and trait-based models (Columns 1 and 2), although the effect size is modest, with a shift ranging between 4 percent and 6 percent across the two models. In contrast, immigration strength remains consistently insignificant. None of these issue-based predictors are significant in the social-distance model, where only race emerges as a marginally significant predictor.

Taken together, these results indicate that both ideological strength and certain forms of issue-specific attitude strength—particularly on abortion and government services—are meaningfully associated with higher levels of affective polarization among strong partisans. Demographic patterns somewhat diverge across models: Gender and religious attendance are negatively associated

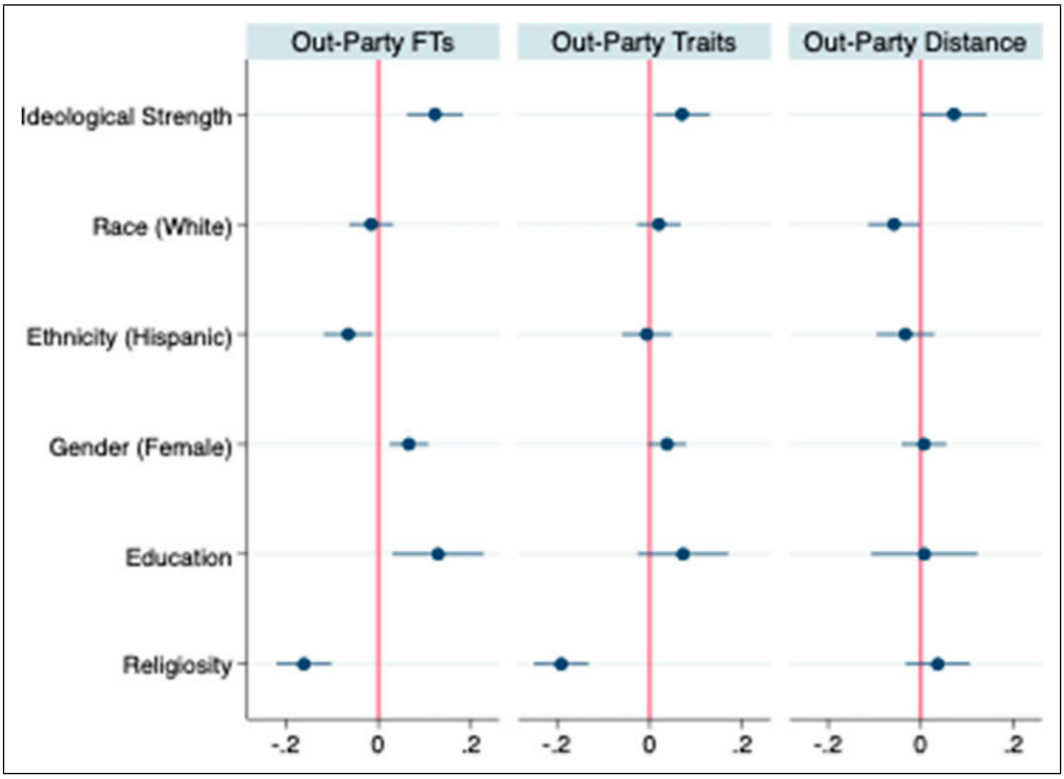

Figure 3. Predicting polarization among strong partisans (measured with 3-point strength item). *Note.* All variables are scaled to range from 0 to 1 for better comparability. $N = 395$. Models are estimated using OLS regression. Analyses are restricted to strong partisans only.

with out-party hostility on feeling-thermometer and trait-based measures—with men and more religiously observant respondents displaying lower levels—while being White is associated with lower social distance from the out-party.

We obtain similar results when we replace the strong partisans on the 3-point partisan strength item with strong partisans who fall in the fourth quartile on the partisan identity scale (see Figure 4). Ideological strength and religiosity remain strong and significant predictors of out-party negativity among strong partisans in both the thermometer and trait-based models (Columns 1 and 2), mirroring earlier models, while none of the included predictors explain variation in social distance preference among strong partisans (Column 3).

These patterns hold when we replace ideological strength with attitude strength on government services, abortion, and immigration (see Figure A4 in the Appendix). In these models, only the first two attitudes—government services and abortion—emerge as significant predictors of out-party negativity. Once again, higher levels of religious attendance consistently reduce the likelihood of being highly polarized in the thermometer- and trait-based models.[5]

To assuage concerns about the nature of our opt-in samples, we replicate the core models from Figures 3 and 5 using ANES cumulative data (2000–2024). As Figures A5 and A6 show, the results are broadly consistent with our original analyses: ideological strength, along with attitude strength on issues[6] such as abortion and government services, remain a significant predictor of out-party hostility among strong partisans across the full time series, as do demographic predictors—particularly religious attendance. That these patterns hold in a probability-based sample such as the ANES strengthens our confidence that our findings are not artifacts of the specific samples we use.

These findings collectively indicate that even among the strongest partisans, substantial variation in out-party negativity remains, and this variation is systematically associated with ideological and issue-specific attitude strength as well as demographic factors such as religious attendance. While ideological strength consistently pushes strong partisans into more hostile territory—a finding that we replicate in our second sample (see Figure A3 in the Appendix)—it is important to note that within the group of strong partisans, only about 28 percent–30 percent identify as extremely liberal or conservative across both of our samples. About half

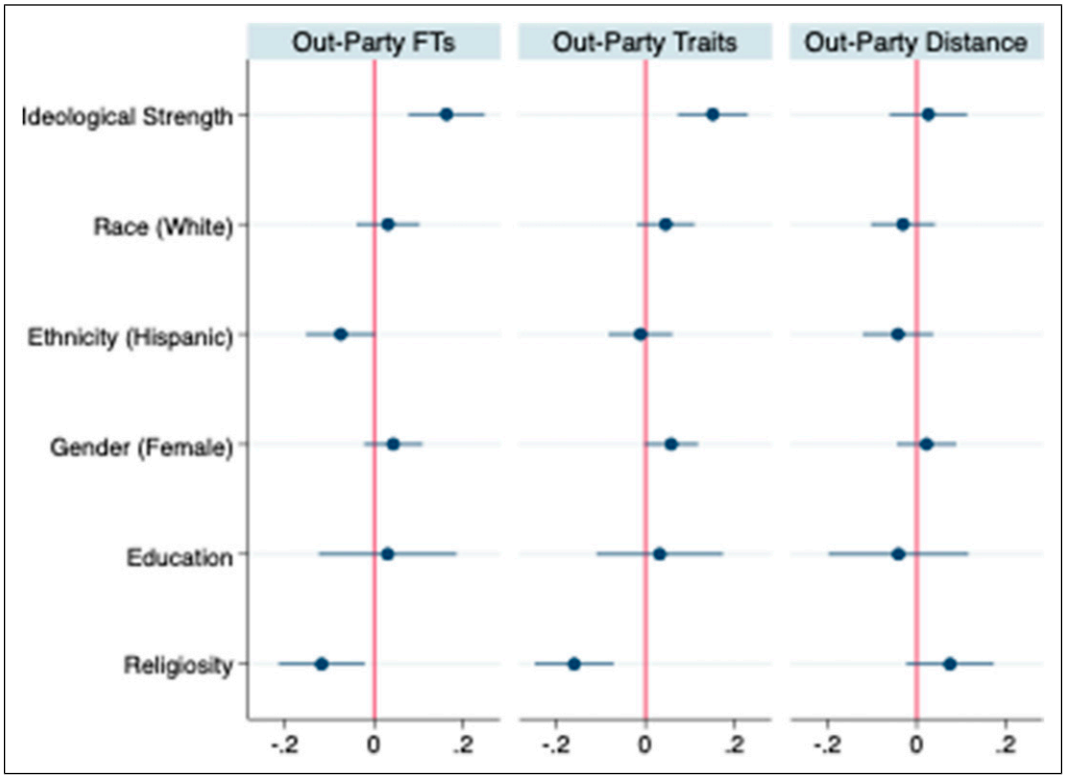

Figure 4. Predicting Polarization among strong partisans (measured with partisan identity strength). *Note.* All variables are scaled to range from 0 to 1 for better comparability. *N* = 223. Models are estimated using OLS regression. Analyses are restricted to strong partisans only.

identify at a more moderate level—simply as liberal or conservative. In the 2024 ANES, the share of strong partisans who place themselves in one of the most ideologically extreme categories is even lower—just 17.2 percent. In contrast, 45 percent identify with the more moderate labels of "liberal" or "conservative." In other words, the attitudinal profile most conducive to out-party hostility characterizes only a minority of strong partisans. The majority, despite their strong partisan commitment, lack the ideological intensity that our analyses identify as a key driver of hostility. This further underscores our central finding: strong partisanship and out-party animosity are not synonymous, and most strong partisans do not carry the ideological fuel that tends to ignite hostile sentiment.

Discussion & Implications

While this study offers important insights into the relationship between partisanship and affective polarization, we acknowledge several limitations.

First, although our dataset is diverse and draws on a representative sample of the U.S. population, the observational design limits our ability to identify what transforms strong partisans into hostile strong partisans. Only experimental designs can provide an answer to that question by manipulating the conditions under which strong partisan identity escalates into active out-party hostility—such as electoral competitiveness, media environments, elite rhetoric, and partisan norms. Without such designs, the present analysis is best understood as documenting robust correlations rather than adjudicating among competing causal explanations. We hope that future experimental research will help uncover the mechanisms underlying the substantial heterogeneity in out-party hostility that we observe even among strong partisans.

A second limitation concerns the timing of our two original surveys. Sample 1 was collected in August 2024—in the immediate aftermath of President Biden's withdrawal from the presidential race, a period of unusually high partisan uncertainty among Democrats. Sample 2 was collected in September 2025, less than a year after the conclusion of a highly consequential presidential election. Both moments are characterized by elevated partisan salience, heightened emotional engagement, and potentially unusual distributions of partisan affect. It is therefore possible that the levels of affective polarization we observe—and the strength of the relationship between partisan identity and out-party

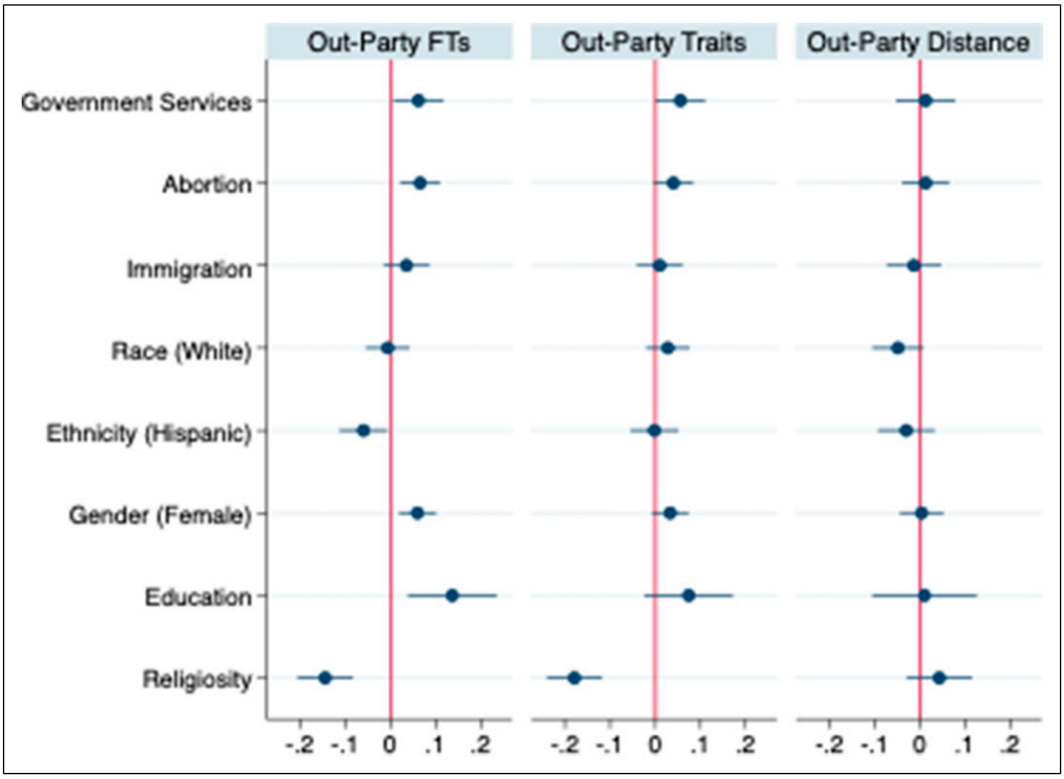

Figure 5. Predicting polarization among strong partisans (including issue strength). *Note.* All variables are scaled to range from 0 to 1 for better comparability. *N* = 395. Models are estimated using OLS regression. Analyses are restricted to strong partisans only.

hostility—reflect the specific political context of these measurement windows rather than stable underlying patterns. The ANES replication analyses mitigate this concern, demonstrating that across more than two decades of electoral cycles, the correlations between partisan strength and out-party hostility remain consistently weak, and substantial within-group heterogeneity in thermometer scores persists at every level of partisan attachment regardless of electoral context (see Figures A1–A2). That the decoupling we observe is not unique to the politically charged moments captured in our surveys strengthens our confidence that it reflects a structural feature of the partisanship–polarization relationship rather than a snapshot artifact. Nonetheless, future work using panel data would be better positioned to assess whether and how the strength of that relationship fluctuates with the broader political environment.

A deeper limitation of the present study—and of the affective polarization literature more broadly—concerns the absence of any agreed-upon standard for what level of partisan hostility is democratically consequential. Existing work has demonstrated that affective polarization correlates with a range of concerning outcomes, including support for democratic norm violations, dehumanization of out-partisans, and willingness to discriminate in everyday social settings (Iyengar et al. 2019; Kalmoe and Mason 2022; Settle 2018). Yet these findings tell us that more hostility predicts worse outcomes—not that any particular level of hostility crosses a threshold beyond which democratic harm becomes likely. Indeed, some recent work has begun to question even the directional link, suggesting that affective polarization may not reliably translate into anti-democratic behavior or reduced accountability at all (Broockman et al. 2023). Whether, and at what level, partisan hostility produces concrete democratic harm thus remains an open empirical question—and one that matters directly for interpreting our results. If even the strongest partisans in our data display meaningful within-group heterogeneity in out-party hostility, the normative stakes of that heterogeneity depend on whether such differences in hostility correspond to differences in democratically consequential behavior. Establishing such a correspondence would require, at minimum, linking survey-based polarization measures to behavioral indicators with concrete democratic stakes—voter intimidation, refusal to accept election outcomes, support for political violence—and examining whether the relationship is linear, stepped, or characterized by a tipping point. Some recent work on pernicious polarization (McCoy and Somer 2019) and radical partisanship (Kalmoe and Mason 2022) has begun to move in this direction, but the field has not yet

converged on a framework for translating levels of affective polarization into assessments of democratic risk. Until it does, claims about whether observed levels of partisan hostility are alarming, tolerable, or benign remain difficult to adjudicate empirically—a limitation this study shares with the broader literature it contributes to.

Despite these limitations, our findings significantly extend a growing body of scholarship—anchored by Bankert (2021) and Mason (2015, 2018)—documenting that partisan strength and out-party hostility are empirically separable phenomena. Viewed through the lens of Brewer's (1999) foundational framework, this decoupling is theoretically coherent rather than surprising: strong in-group identification does not automatically generate out-group hostility, and the partisan domain appears to be no exception. What our study adds is systematic, multi-measure evidence that this decoupling is robust—consistent across feeling thermometers, trait attributions, and social distance scales, across two distinct operationalizations of partisan strength, and across opt-in and probability-based samples spanning more than two decades of American electoral politics. Consistent with Mason's sorting framework, ideological strength emerges as the most reliable predictor of which strong partisans become hostile, underscoring that it is the alignment of partisan and ideological identities—not partisan strength per se—that activates the more corrosive forms of out-party animosity.

These findings carry important implications for how we understand and evaluate the relationship between partisanship and polarization. At the most basic level, they caution against treating partisan strength as a proxy for affective polarization (e.g., Graham and Svolik 2020). But they also highlight that measurement choices matter: different operationalizations of affective polarization can meaningfully alter conclusions about how polarized the public actually is and who the most hostile partisans are. More broadly, they invite a reconceptualization of what strong partisanship means for democratic health. If the majority of strong partisans are neither ideologically extreme nor particularly hostile toward the out-party, then the democratic risks associated with rising partisan strength may be considerably more concentrated—and more tractable—than headline accounts of polarization imply.

For those concerned with the normative health of democracy, this reframing matters: the goal should not be to weaken partisan attachments, which serve important functions in organizing political competition and mobilizing citizens, but to understand and address the conditions under which partisanship tips into the kind of hostility that threatens democratic norms and institutions.

ORCID iDs

Alexa Bankert https://orcid.org/0000-0002-3672-8015
Joshua N. Zingher https://orcid.org/0000-0003-3928-4269

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

Replication data and code are available on the author's website at https://www.alexabankert.com/.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. The replication sample does not include the partisan identity scale nor the social distance battery. However, it does contain the 3-point partisan strength item and a multi-item scale assessing the extent to which partisans view violence against the out-party as justified. This measure is arguably an even stronger indicator of partisan hostility than social distance alone, and it reinforces our claim that partisan strength is only weakly related to different forms of out-party negativity. See Table A3 in the Appendix for details.
2. As a robustness check, we recomputed all correlations using Spearman's rank-order correlation, a nonparametric alternative that does not assume linearity, and obtained substantively identical results for both measures of partisan strength (see Tables A4 and A5).
3. Critical readers may then ask why the partisan identity scale correlates much more strongly with the social distance measure. Although we cannot fully resolve this question empirically here, we offer a theoretical explanation: social distance captures a behavioral disposition—such as willingness to avoid or exclude—whereas feeling thermometers primarily measure affective evaluations. It is plausible that partisan identity maps more readily onto avoidance than onto negative affect, because the former is a more direct expression of group boundary maintenance—exactly the function that Brewer (2007) identifies as the primary behavioral output of strong in-group identification. Put differently, one can be deeply invested in one's partisan identity, deriving a strong sense of self and belonging from it, without that investment translating into cold affect toward the out-party—but the same identity investment may more readily produce a preference for social separation, because separation preserves group distinctiveness rather than requiring active animus.
4. It is worth noting that measures of affective polarization are often non-normally distributed, typically skewed toward very high in-party warmth or strong out-party dislike.

Because the mean is sensitive to skewness and outliers, relying on it alone can distort or overstate differences across levels of partisan strength.

5. To address concerns regarding sample size, we re-estimated all models using 2,000 bootstrap replications. The bootstrapped standard errors and bias-corrected confidence intervals closely mirror the conventional OLS results. Substantive conclusions remain unchanged, suggesting that the findings are not driven by small-sample estimation assumptions (see Tables A6 and A7).
6. Notably, immigration attitude strength emerges as a significant predictor of out-party hostility among strong partisans in the ANES replication analyses, a result that does not appear in our original samples. One plausible interpretation is that immigration's salience as a driver of partisan hostility varies across electoral contexts—a pattern that, if anything, reinforces our broader argument that the relationship between partisanship and out-party animosity is conditional rather than fixed.

References

Abramowitz, Alan, and Kyle Saunders. 2008. "Is Polarization a Myth?" *The Journal of Politics* 70 (2): 542–55.

Abramowitz, Alan, and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of US Elections in the 21st Century." *Electoral Studies* 41: 12–22.

Bankert, Alexa. 2021. "Negative and Positive Partisanship in the 2016 Presidential Elections." *Political Behavior* 43: 1467–85.

Bankert, Alexa. 2024. *When Politics Becomes Personal: The Effect of Partisan Identity on Anti-Democratic Behavior*. Cambridge University Press.

Bankert, Alexa, Leonie Huddy, and Martin Rosema. 2017. "Measuring Partisanship as a Social Identity in Multi-Party Systems." *Political Behavior* 39 (1): 103–32.

Brewer, Marilynn. 1999. "The Psychology of Prejudice: Ingroup Love or Outgroup Hate?" *Journal of Social Issues* 55 (3): 429–44.

Brewer, Marilynn. 2007. "The Importance of Being We: Human Nature and Intergroup Relations." *American Psychologist* 62 (8): 728–38.

Broockman, David E., Joshua L. Kalla, and Sean J. Westwood. 2023. "Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not." *American Journal of Political Science* 67 (3): 808–28.

Campos, Nicolas, and Christopher Federico. 2025. "A New Measure of Affective Polarization." *American Political Science Review* 120 (1): 160–78.

Cassese, Erin. 2021. "Partisan Dehumanization in American Politics." *Political Behavior* 43 (1): 29–50.

Druckman, James N., Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2024. *Partisan Hostility and American Democracy: Explaining Political Divisions and when They Matter*. University of Chicago Press.

Graham, Matthew H., and Milan Svolik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States." *American Political Science Review* 114 (2): 392–409.

Iyengar, Shanto, and Masha Krupenkin. 2018. "The Strengthening of Partisan Affect." *Political Psychology* 39 (S1): 201–18.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, Not Ideology: a Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76 (3): 405–31.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22 (1): 129–46.

Kalmoe, Nathan, and Lilliana Mason. 2022. *Radical American Partisanship: Mapping Violent Hostility, Its Causes, and the Consequences for Democracy*. University of Chicago Press.

Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar. 2017. "The Hostile Audience: the Effect of Access to Broadband Internet on Partisan Affect." *American Journal of Political Science* 61 (1): 5–20.

Levendusky, Matthew. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. University of Chicago Press.

Levendusky, Matthew. 2013. *How Partisan Media Polarize America*. University of Chicago Press.

Martherus, James, Andres Martinez, Paul Piff, and Alexander G. Theodoridis. 2021. "Party Animals? Extreme Partisan Polarization and Dehumanization." *Political Behavior* 43: 517–40.

Mason, Lilliana. 2015. "'I Disrespectfully Agree': the Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59 (1): 128–45.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. University of Chicago Press.

McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press.

McCoy, Jennifer, and Murat Somer. 2019. "Toward a Theory of Pernicious Polarization and How It Harms Democracies." *The Annals of the American Academy of Political and Social Science* 681 (1): 234–71.

Reiljan, Andres. 2020. "'Fear and Loathing Across Party Lines' (Also) in Europe: Affective Polarisation in European Party Systems." *European Journal of Political Research* 59 (2): 376–96.

Rosenblum, Nancy. 2008. *On the Side of Angels: An Appreciation of Parties and Partisanship*. Princeton University Press.

Settle, Jaime E. 2018. *Frenemies: How Social Media Polarizes America*. Cambridge University Press.

Stroud, Natalie Jomini. 2011. *Niche News: The Politics of News Choice*. Oxford University Press.

Tajfel, Henri, and John Turner. 2004. "The Social Identity Theory of Intergroup Relations." In *Key Readings in Social Psychology: Political Psychology*, edited by John Jost and Jim Sidanius. 1st ed. Psychology Press.

Tajfel, Henri, Michael G. Billig, R. P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1 (2): 149–78.

Zingher, Joshua N. 2022. *Political Choice in a Polarized America: How Elite Polarization Shapes Mass Behavior*. Oxford University Press.