Partisanship increasingly factors into the behavior of Americans in both political and nonpolitical situations, yet the bounds of partisan prejudice are largely unknown. In this paper, we systematically evaluate the limits of partisan prejudice using a series of five studies situated within a typology of prejudice. We find that partisan prejudice predicts suppression of hostile rhetoric toward one’s own party, avoidance of members of the opposition, and a desire for preferential treatment for one’s own party. While these behaviors may cause incidental or indirect harm to the opposition, we find that even the most affectively polarized — those with the strongest disdain for the opposition — are no more likely to intentionally harm the opposition than those with minimal levels of affective polarization.

Partisanship, by most accounts, divides the American public. Disdain for the political opposition — referred to as affective polarization (Iyengar, Sood, and Lelkes 2012), social polarization (Mason 2014), or partyism (Brooks 2014; Sunstein 2014) — has significantly grown over the past 30 years.1 At the same time, evidence shows that partisan affiliation not only increasingly affects political behavior (Bafumi and Shapiro 2009) but also appears to change the behavior of partisans in apolitical environments as diverse as mate selection (Alford et al. 2011; Huber and Malhotra 2012; Klofstad, McDermott, and Hatemi 2012), home buying (Bishop 2009), and scholarship allocations (Iyengar and Westwood 2015; Munro, Lasane, and Leary 2010; for a review, see Brandt et al. [2014]). Recent evidence shows that partisanship is a strong social identity (Huddy, Mason, and Aaroe 2015; Iyengar et al. 2012; Iyengar and Westwood 2015) that fundamentally drives America’s culture war (Jacoby 2014). These results are compelling, but just how bad has partisan prejudice become in America? Does partisan anger and animosity become Americans to forgo democratic norms?

Partisanship in the American context is highly conflictual and is inherently based on the formation and competition of parties and party members. The literature suggests that partisanship is important in many aspects of American life, but the magnitude of the effects of partisan affect on behavior is unclear. The extensive study of ideological polarization (Abramowitz and Saunders 1998; Brewer 2005; Fiorina and Levendusky 2006; Hetherington 2001; Iyengar et al. 2012; Iyengar and Westwood 2015; Jacobson 2003; Levendusky 2009; Wolfe 1998) — and the emerging work on affective polarization (e.g., Carlin and Love 2013; Levendusky and Malhotra 2013; Mason 2014) — does not explain how partisan affect changes the behavior of partisans, nor does it assess just how damaging partisan prejudice is within American society.

We present a rigorous assessment of the limits of partisan prejudice using a framework of prejudice developed by Gordon Allport (1954) that is widely used in both political science and the social sciences. One need not take Allport’s framework as the final word on prejudice to recognize its utility in demarcating increasing levels of prejudice. Prejudice, as Allport and many others argue (e.g., Green and Seher 2003), is not a binary construct but a continuum. And as prejudice increases, its manifestations evolve. The Allport framework is useful as it defines differing levels of
prejudice and formalizes criteria for discrimination.2 We augment this framework by also considering that interparty prejudice may not cause behavior toward the out-group that is reciprocally related toward behavior toward the in-group. For instance, prejudiced partisans may be willing to work to help an in-group member but not be willing to hurt an out-group member.

We offer five studies that investigate how partisan affective polarization changes behavior toward co-partisans and opposing partisans. Our first three experiments test the first two levels of Allport’s framework (promotion of negative rhetoric and avoidance) and show that partisan affect does relate to prejudicial behavior at these levels in Allport’s framework. Our first experiment shows that as partisan polarization increases, participants are more willing to suppress rhetoric that is hostile toward their own party but are not more willing to promote rhetoric that criticizes the out-party. Our second and third experiments show that higher levels of affective polarization increase the avoidance of opposition partisans. Affective polarization and the first levels on Allport’s framework of prejudice are related, which is consistent with existing work on partisan preferences. At these levels of prejudice, partisans both favor co-partisans and punish opposing partisans, but many posit that these low-level forms of prejudice have caused or will cause partisans to actively discriminate against the opposition (see Bishop 2009; Iyengar et al. 2012; Iyengar and Westwood 2015). Our last two experiments consistently show that affectively polarized partisans favor co-partisans even when such favoritism violates democratic norms. However, these co-partisan preferences are not mirrored by opposing partisan biases.

We show that, despite growing partisan animosity and co-partisan favoritism—a clear willingness to denigrate the opposition and a general preference to avoid the opposition—those at the highest levels of affective polarization are no more willing to intentionally harm (discriminate against) the opposition than those at the lowest levels of affective polarization. Even though citizens appear to endorse the hostility that dominates modern political campaigns (e.g., Geer 2008), partisan news (Prior 2007; Sobieraj and Berry 2011), and governance (Grimmer and King 2011), they endorse discriminatory behavior only by favoring co-partisans. Elites supporting or engaging in discriminatory behavior against the political opposition (e.g., politically motivated investiga-

gations, cover-ups, gerrymandering, and campaign finance abuse), our results suggest, are out of step with even the most partisan citizens.

Further, this paper addresses lingering puzzles in our understanding of partisan bias not engaged in earlier work (i.e., Carlin and Love 2013; Iyengar and Westwood 2015; Munro, Lasane, and Leary 2010). First, these studies test for differences in the behavior of identifiers of one party toward identifiers of another; they do not determine the degree to which partisan prejudice underpins this behavior. Second, existing models test partisan prejudice in limited contexts and largely assume that co-partisan (in-group) favoritism and opposing partisan (out-group) animosity are zero-sum and reciprocally related. Finally, past research treats prejudice as a binary outcome—it either exists or it does not—and does not differentiate between different manifestations of prejudice. We address all three of these points directly.

Our work has broad implications for the study of American politics. Our findings demonstrate the need to revisit assumptions on the symmetry and nature of the divisions of Americans along party lines. Partisan hostility among the electorate runs high, but the actual effects of affective polarization are not as pernicious as has been suggested in various academic and nonacademic sources. Partisan prejudice is real, but it is largely limited to partisan bluster, avoidance, and co-partisan favoritism. As we discuss later, this has implications for democratic politics and representation.

THE EXTENT OF AFFECTIVE POLARIZATION IN AMERICA

Ordinary Americans are divided, although not necessarily due to any sort of ideological or policy-based difference between parties (Mason 2014). While there is much disagreement over the existence and extent of ideological polarization in the electorate (Abramowitz 2007, 2010; Abramowitz and Saunders 1998, 2005, 2008; Abramowitz and Stone 2006; Brewer 2005; Fiorina and Abrams 2008; Fiorina, Abrams, and Pope 2005; Fiorina and Levendusky 2006; Hetherington 2001; Jacobson 2003, 2004, 2005; Jacobson and Abramowitz 2006; Levendusky 2009; Wolfe 1998), there is general agreement that partisans of different stripes do not seem to like each other very much.

Since the 1980s, the percentage of Americans who strongly dislike the political opposition has risen fairly dramatically. Iyengar et al. (2012) find that partisans today are much more likely to believe that members of the out-party are mean, hypocritical, and selfish. Additionally, the gap between in-party and out-party warmth registered on

2. Allport also includes violence and extermination as the most serious levels of prejudice. However, apart from scattered cases of vandalism (Associated Press 2004), these levels are not yet concerns in American politics.
the American National Election Studies’ feeling thermometer scales has increased by a third over the past 30 years (Iyengar et al. 2012), and Americans register increasing anger at the other side’s political candidate (Mason 2013). Citizens are also more dubious of the motivations and ulterior motives of politicians from the opposing party than for co-partisans (Munro, Weih, and Tsai 2010). Partisanship is now a strong social identity that significantly changes attitudes and behaviors (see Huddy et al. 2015; Iyengar and Westwood 2015). This line of work indicates that America is increasingly divided on affective, albeit not necessarily ideological, grounds (see also Hetherington and Weiler 2009; Jacoby 2014; Mason 2013, 2014).

A thorough exposition of the roots of affective polarization is beyond the scope of this paper. But, in brief, some place the blame on ideological polarization at the elite level (Haidt and Hetherington 2012; Iyengar et al. 2012; Rogowsky and Sutherland 2015) coupled with an influx of negative campaigning (Iyengar et al. 2012) and the growth of partisan media (Lelkes, Sood, and Iyengar, forthcoming). According to this view, vitriol among political elites is exacerbating out-party animosity. Fiorina (2013) and Mason (2014) take the view that affective polarization is due to issue-based sorting of the electorate. Here we focus on documenting the effects of partisan animosity.

Regardless of source, partisan animosity is unique among social divides. Strong social norms constrain overt bias along racial, ethnic, and gender divides, while there are no comparable norms to constrict partisan bias. The nature of politics is also inherently competitive and is based on conflicts between groups for the division of real and consequential goods. Partisan animosity is both inherent in the structure of democratic systems and encouraged by political officials, parties, and candidates. The result is that partisan identity is imbued with a constant sense of threat or fear of loss to rivals. By fostering group representation and identity, democratic systems also construct a strong in-group and out-group divide that, ironically, creates incentives to discriminate based on partisan affiliation.

DEFINING THE BOUNDS OF PARTISAN PREJUDICE

As a benchmark to which we can compare partisan prejudice, we start with Allport’s (1954) framework of prejudice, which is utilized to study prejudice in a variety of domains (Dovidio, Glick, and Rudman 2005), including racism (Feagin 1992), ageism (Fraboni, Saltstone, and Hughes 1990), and sexism (MacInnis and Hodson 2012). We focus on the first three levels of the framework: the promotion of negative speech, the avoidance of out-group members, and actual discrimination, which comes about only when a group actively and intentionally causes harm to other groups (Allport 1954). Although milder levels of prejudice captured in the first two levels of Allport’s typology are certainly politically important—bashing the opposition and avoiding the opposition are far from normative models of citizen behavior—discrimination is particularly concerning, as it is incompatible with doctrines of equal justice and egalitarianism defined in the American ethos (McClosky and Zaller 1984).

There is evidence that we treat co-partisans differently than we treat out-party members even in nonpolitical domains. For example, previous studies where partisans must allocate some punishment or reward, that is, make a zero-sum decision to either allocate a scholarship between two people (Iyengar and Westwood 2015) or to admit one of two students into college (Munro, Lasane, and Leary 2010), show that partisanship drives prejudicial decision making. In these studies, one of the hypothetical applicants is typically a Republican, and the other is typically a Democrat. While these studies clearly show an asymmetry in outcomes, the nature of these interactions pits one group against another, which makes it impossible to untangle “in-group love” from “out-group hate.” As Brewer (1999, 431) notes, “Whenever the structure of resources or opportunities really is a zero-sum situation, any preferential treatment of in-group members will be achieved at the detriment of out-group members, but this does not mean that attitudes are similarly zero-sum.”

Feelings toward the in-group are not necessarily reciprocally related to feelings toward the out-group (Brewer 1999; Crawford, Modri, and Motyl 2013; Halevy, Weisel, and Bornstein 2012; Weisel and Böhm 2015). While the original “minimal-group” experiments (Rabie and Horwitz 1969; Tajfel et al. 1971) demonstrate that the arbitrary categorization of participants into “in-groups” and “out-groups” leads to an allocation of resources that benefits the in-group, later studies demonstrate that the in-group/out-group effect dissipates when the resource represents a punishment rather than a reward (Otten and Mummendey 2000; Otten, Mummendey, and Blanz 1996). Although Tajfel et al.’s (1971) finding of in-group favoritism has been replicated numerous times when it comes to the distribution of money or other positive outcomes in the minimal group paradigm, extensions to the original design used by Tajfel show that out-groups are not punished more than in-groups when it comes to the distribution of unpleasant tasks or unpleasant experiences (Mummendey et al. 1992; Otten et al. 1996). Despite this work in psychology, the majority of the political science literature suggests, although does not formally test, that partisanship is a special case where punitive
behavior is especially prevalent and acceptable (e.g., see Bishop 2009; Iyengar et al. 2012; Iyengar and Westwood 2015).

We augment Allport’s framework to distinguish between in-group favoritism and out-group hostility within each domain. If partisan prejudice predicts quashing negative speech directed toward one’s own party, it is indicative of in-group favoritism within the domain of negative speech. If partisan prejudice is associated with promoting negative speech, we see it as indicative of out-group hostility. Furthermore, quashing of offensive speech directed toward one’s own group is a more benign manifestation of prejudice in this domain than promoting negative speech against the opposition. Similarly, if partisan prejudice predicts a preference for the company of co-partisans, it is a less severe manifestation in the second domain of Allport’s framework than is the active avoidance of out-partisans.

In the next section, we systemically evaluate the nature of partisan prejudice through a series of experiments designed to test the relationship between affective polarization and each level of prejudicial behavior. Throughout these experiments, we disentangle the effects of in-group and out-group favoritism. In supplemental analyses, we also examine whether Republicans and Democrats differ substantively in their reactions, as we might expect those on the right to be more punitive than those on the left (e.g., Hibbing, Smith, and Alford 2014; Zavala et al. 2010).

**LEVEL 1 (ANTILOCUTION): AFFECTIVE POLARIZATION AND THE PROMOTION OF HOSTILE RHETORIC**

The lowest level of prejudice in the Allport framework is a willingness to speak badly of, or at least promote negative speech toward, an out-group. Anecdotal evidence and prior research show that partisan interactions are often rancorous (e.g., Sobieraj and Berry 2011). This open willingness to derogate partisan opponents is very close to the traditional conceptualization of prejudice as “free-ranging, hostile feelings or unmitigated, derogatory stereotypes” (Jackman 2005, 96), but when citizens debate politics—especially online—they do so in spaces that are often dominated by highly motivated citizens and political elites. For example, the comments sections on FoxNews.com or HuffingtonPost.com may not be representative of what political activity or hostility looks like for average Americans. The exponential increase in name-calling, character assassinations, mockery, and other forms of “political outrage” shown in prior work (see Sobieraj and Berry 2011) are dramatic, but are less politically interested Americans willing to engage in these kinds of prejudicial behaviors? Furthermore, is political prejudice more associated with acts of hostility (speaking badly of their opponents) or with in-group solidarity (making sure no one else speaks badly of the group)?

We use an experiment to assess the degree to which affective polarization moderates the willingness to spread vitriolic political speech (operationalized as news content). This allows us to document the behavior for a representative group of citizens and to show the influence of partisan hatred on this behavior. Participants \( (N = 556) \) from a sample drawn from a Survey Sampling International (SSI) panel completed our experiment under the guise of helping a news organization decide what information it should feature on its website. Participants were therefore given the opportunity to promote or suppress negative news content in a realistic setting.

The study’s participants were told that an opinion piece had been submitted by a reader and that the news website needed input on the quality of the article. In this sense, the design replicates the content aggregation and assessment processes present in many online news outlets. Participants were not told that the research project was created by political scientists or that the content they would evaluate would be political.

Under this pretense, participants were randomly assigned to read one of two news opinion articles: a column from Fox News blaming congressional gridlock on Democrats or an MSNBC column blaming Republicans for the gridlock. The two stories were identical, with only the source and the name of the party altered. The text of the manipulations appear in table 1. The text goes beyond mere criticism and includes name-calling and hyperbole. After reading the article, participants were asked if they would endorse the article for inclusion on the news organization’s website. This design made it possible to measure the extent to which partisans are likely to share and endorse news content that criticizes the opposition and the extent to which partisans are likely to recommend against sharing news content that criticizes their own party.

Our measure of affective polarization here—and throughout this manuscript—is the difference in the feeling thermometer score for a respondent’s own party and the feeling thermometer score for the opposition party. This value

---

3. We use an opinion story because it serves as a conservative test of how willing partisans are to denigrate the opposition. Profane language and ad homonym attacks evoke a stronger affective response in pretests, but these kinds of messages are taken less seriously by respondents.

4. Following past work, we include respondents who rated their in-party lower than the out-party. Neither omitting these respondents from the analysis nor recoding their score to the neutral point substantively change the results in any of our analyses.
was then rescaled to range between 0 and 1. While prejudice has both cognitive and affective components, the affective component is most strongly associated with attitudes and behavior toward outside groups (e.g., Stangor, Sullivan, and Ford 1991). This measure is the most used measure of affective polarization (Haidt and Hetherington 2012; Hetherington and Weiler 2009; Iyengar and Westwood 2015; Mason 2014) and consistently yields results similar to other operationalizations (e.g., social distance measures). Relying on difference scores, rather than raw feeling thermometer scores, helps ameliorate the differential item functioning of feeling thermometers (Wilcox, Sigelman, and Cook 1989).

In this experiment, we find clear evidence of prejudice in how citizens evaluate critical news content. People generally supported sharing content that is critical of the opposition and did not support sharing content that criticizes their own party. Participants who read the article that criticized the opposing party recommended inclusion 64.74% of the time, while only 25.41% of participants recommended sharing the article criticizing their own party. At first glance, out-group hate handily overwhelms in-group love, but it is unclear how much influence partisan animosity actually had on these recommendations.

To assess whether affective polarization affects the probability that a partisan recommends an article, we regressed, using a binary logit model, a person’s decision on his or her affective polarization score. Figure 1 shows the effect of affective polarization on the probability of recommending each of the two articles. Higher levels of affective polarization are strongly related to quashing an article hostile to the in-party ($b = -3.51$, 95% confidence interval $[-5.00, -2.02]$), but not to promoting an article hostile to the out-party ($b = 1.31$, 95% confidence interval $[-0.39, 3.01]$).

The difference in how minimally and maximally polarized participants assess content critical of their party is massive (fig. 1, left panel). The probability of a minimally polarized participant recommending content critical of his or her party is .87, while the probability of a maximally polarized participant recommending the article is only .15. Although the relationship is not as stark, and the confidence interval around the estimate overlaps with zero, more affectively polarized participants are also more likely to promote speech critical of the other side. The likelihood of the most polarized participant recommending such an article is .85, while the likelihood of the least polarized participant recommending the article is .54. Hence, the effect of affective polarization on recommending was twice as large when the article bashed the in-party as when it attacked the out-party.

Those who are affectively polarized respond to negative partisan rhetoric even when evaluating a fairly benign opinion piece. Partisan affective polarization is clearly associated with the promotion of negative speech against one’s opponents. For the politically prejudiced, protecting the in-group from negative rhetoric is more important than attacking the out-group. Thus, partisan antipathy reaches the first level of Allport’s prejudice framework, but it does so in an asymmetric way.

**LEVEL 2 (AVOIDANCE): INTENTIONAL AVOIDANCE AND ISOLATION OF OPPOSING PARTISANS**

In many apolitical situations, people avoid those who do not share their partisanship. Partisans sort themselves into

6. An additional sample of 418 drawn from SSI completed this task, but they were randomly assigned the Republican-bashing story, the Democrat-bashing story, or a control article on shark attacks. Results from the two partisan conditions replicate results reported in the manuscript. The endorsement rate for the shark attack article is nearly 8% higher than for the partisan stories. Unlike the partisan articles, support for the control article is consistently high for both Democrats and Republicans.

7. We include the regression models with and without controls for all results in the appendix.
romantic relationships with others who share their partisan identity. Indeed, partisan spousal preferences are actually stronger than physical (e.g., body shape) or personality attributes (Alford et al. 2011). Partisan information is often strategically suppressed on online dating sites (Klofstad et al. 2012), but when it is present, shared partisanship predicts reciprocal communication between men and women seeking potential dates (Huber and Malhotra 2012). Partisans are increasingly uncomfortable with their offspring marrying a member of the opposing party (Iyengar et al. 2012), and actual marriage between opposing partisans is an infrequent occurrence (Stoker and Jennings 1995). This increase in social distance may be related to the actual geographical distance that has grown between partisans (Bishop 2009; Tam Cho, Gimpel, and Hui 2013). Partisans appear to prefer to date, marry, and live with co-partisans, but it is unclear if this is because partisans desire to be with similar individuals, if they dislike being with dissimilar individuals, or a combination of both.

Past research also does not directly test the second level of Allport’s framework. To meet the threshold set by Allport, it is not sufficient for partisans to prefer co-partisans or even to avoid opposing partisans. This avoidance must be intentional, clearly visible to the avoided group, and the avoidance must cause feelings of isolation among those avoided. When making a selection on dating websites, picking a house, or selecting a mate, it is usually not clear that a person or neighborhood was avoided or that this avoidance had anything to do with partisanship. We use a pair of novel experiments to show that, when forming groups, partisans will avoid opposing partisans, that partisanship is identified as the primary reason for exclusion, and that those excluded are both aware that they were excluded because of their partisan affiliation and feel isolated as a result.

**Study 2.1: Partisan prejudice in group selection**

We test avoidance using a team formation task. As part of the task, participants choose between ostensibly real participants that varied on a variety of attributes. This experiment allows us to determine whether sharing a party identity is more important than having a skilled teammate and, as a result, the impact that affective polarization has on this relationship.

Participants drawn from the SSI panel (N = 611) were told that they would complete a series of puzzles (simplified word completion tasks based on crossword puzzle clues) in teams. The tasks and all clues had nothing to do with politics. Participants were asked to complete three trial rounds of the word completion task in order to increase the validity of the design, to increase comprehension of what the group would do, and to demonstrate the importance of intelligence and education in the tasks. To enhance the believability of the task, each participant was placed in a “waiting room” and told that it would take a moment for a sufficiently large group of other people completing the survey to gather.

After a short amount of time, participants were told that they were randomly selected by our software to continue as

---

8. See “Supporting Information” in the appendix for the stimuli and the exact instructions.
one of two team leaders. The other leader, they were told, had already picked her team, and the respondent would need to select three players from a list of four for the second team. To incentivize the task, participants were told that their success/score would hinge on both their abilities and the abilities of their team. Participants were explicitly told that the person not picked as a team member would be excluded from participating in the remainder of the study. The player profiles were constructed so that the least academically qualified player was an Independent. Two partisans (one Democrat and one Republican) were essentially the same, both listed with college educations. An additional Independent had qualifications that were lower than those of the partisans but higher than the high school-educated Independent (see table 2). Participants could therefore select the three players with the highest level of education (tolerating a member of the political out-party) or make a selection that excluded the political out-party in favor of the less academically qualified Independents. The order of each player was randomized, participants only completed one round of team selection, and all other traits were fixed.

Partisanship had strong effects on team selection (see fig. 2). Roughly 54% of players excluded the opposing partisan player and opted for the less educated Independent, while 94% of players selected the co-partisan to join their team. Partisans clearly avoided the other side, even at the cost of having a less educated team member working on tasks that require education and verbal acumen. The effects of affective polarization are even more dramatic. While affective polarization is related to both favoritism and avoidance, the relationship is again asymmetric. As seen in the left panel of figure 2, as levels of affective polarization increased, the predicted probability of selecting the co-partisan to join a team grew 12 points, from .86 among the least affectively polarized to .98 among the most affectively polarized ($b = 1.89, 95\% \text{ confidence interval } [1.12, 3.65]$). Compare this to the relationship between affective polarization and avoiding the other side ($b = -2.77, 95\% \text{ confidence interval } [-3.66, -1.87]$). The predicted probability of selecting a teammate from the opposing party among the least affectively polarized is .89, and it drops to .34 among the most affectively polarized. To put this another way, the relationship between affective polarization and out-group avoidance is more than six times as large as the relationship between affective polarization and in-group attraction. Of course, the probability of selecting an in-group member is quite high even among the least polarized, so the asymmetry is, at least, in part due to ceiling effects.

Participants did not hide their prejudices, and many openly explained that they constructed their teams so as to avoid members of the political opposition. After team selection, participants were shown the profile of the person they did not pick. Participants, in an open response field, were asked to explain their reasons (if any) for excluding the person. In these responses 27.80% explicitly stated the decision was because of the player’s partisanship, with 0% of those in the bottom quartile of affective polarization and 47.37% of those in the top quartile of affective polarization listing partisanship. By comparison, 21.79% listed education as the reason for exclusion.

Participants citing partisanship as a reason for exclusion were also less moved to try to include the excluded player than those who listed some other reason for team selection. After explaining their selection, participants were told, “In earlier versions of this study many people who were not picked to participate in a group reported feeling left out and excluded. Knowing this, would you change your group se-

9. To ensure that those participating in this task viewed education and intelligence as highly relevant and political affiliation as irrelevant to task performance, we recruited an additional sample of 418 from SSI. We asked these individuals to read the task instructions, complete the example questions, and read the group-formation instructions. However, instead of showing the participants’ profiles, we asked individuals to rank a series of seven traits (randomly ordered: age, education, gender, race, income, political affiliation, and intelligence) by order of importance to successfully completing the task as a team member. Participants placed education as first- or second-most important 68% of the time. Similarly, participants placed intelligence as first or second 79% of the time. Political affiliation is in the top two spots only 6.6% of the time and was identified as the least important trait 24% of the time. The majority of those exposed to the task viewed political affiliation as unimportant to task performance.

10. This task asks participants to exclude individuals from a team and consequently excludes those individuals from possible payouts. In the Allport framework, the primary consideration is the motivation for exclusion. In this study (and detailed in study 2.2), participants excluded people out of self-interest (to form better teams) and not a desire to avoid contact. This exclusion does create incidental harm, but participants did not identify a desire to hurt the a supporter of the political opposition as the reason for exclusion.

11. We choose to exclude race from the profiles for several reasons. First, although race is a very important factor in group selection tasks, prior work shows that the effects of partisanship trump the effects of race in behavioral tasks (see Iyengar and Westwood 2015). Second, crossing race with partisanship would generate many pairings that are rare in the general public, which would reduce the validity and believability of this design.

12. Party identification does not moderate this relationship. The relationship between affective polarization and the probability of selecting a co-partisan and the relationship between affective polarization and the probability of not selecting an opposing partisan were the same for both Republicans and Democrats (see Appendix 3.4 in the appendix).

13. Partisan-based motivations and education-based motivations were coded using pattern matching. A human verified the classifications.
lection or keep it the same?” Participants who reported that their decision was based on partisanship were approximately three times less likely to want to change their group than those who made their decision for other reasons (3.70% vs. 10.52%). High levels of affective polarization are related to decreased selection rates for members of the political opposition, and participants freely admitted to engaging in prejudicial behavior. Moreover, those who selected based on partisanship were more confident in their choices. Hence, partisan avoidance is in no way unconscious or incidental among the affectively polarized.

Study 2.2: Perceived partisan bias among the excluded
Study 2.1 demonstrates that participants are more than willing to avoid opposing partisans, but to fully test avoidance in Allport’s framework, we must show that those who are excluded are aware that their exclusion is due to their partisanship and not due to some other cause. To test this component of avoidance, we constructed a second study to assess how likely people who are ostensibly not picked to play as part of a team attribute their exclusion to their partisanship. A separate sample of 153 participants was drawn from the SSI panel for this study. After completing the three example word completion tasks, these participants were told: “Two group leaders were randomly selected by our software. Both leaders reviewed basic (anonymous) demographic information about you. You were not picked and are therefore excluded from further participation.” The participants were then told, “The leaders saw your age, level of education, political party and marital status” and were asked “Do you think any of these caused you to not be picked (select as many as you think apply)?” Indeed, the most commonly selected trait is partisanship, with 43.3% of participants attributing their exclusion to their party affiliation (table 3). It is key to note that at this point in the study neither partisanship nor politics were mentioned or referenced (other than in three demographic questions within a set of 15). The percent of individuals reporting their partisanship as a reason for exclu-

<table>
<thead>
<tr>
<th>Table 2. Study 2.1 Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Marital status</td>
</tr>
<tr>
<td>Political affiliation</td>
</tr>
</tbody>
</table>

![Figure 2. Likelihood of selecting a co-partisan and opposing partisan team member](image-url)
sion is relatively stable across levels of affective polarization and is similar for those at the lowest level of affective polarization (42.9%) and those at the highest level of affective polarization (48.5%). Although many people identified that partisanship had an effect on their exclusion, those at the lowest levels of polarization were also likely to attribute their exclusion to age and, to a lesser extent, their level of education and martial status. For those at the highest level of affective polarization, these other factors were much less likely to be selected. The highly affectively polarized were no more likely to identify partisanship as a reason for exclusion, but they were much more likely to only identify partisanship as the reason for exclusion.

In sum, partisan prejudice is clearly related to avoiding the other side. Partisans would rather work with members of their own party than members of the other party. However, the more affectively polarized would rather work with a less qualified Independent than a more qualified out-party member. Additionally, both the avoiders and the avoided identified partisanship as the primary reason for exclusion. Surprisingly, the affectively polarized made no apologies about their choice.

LEVEL 3 (DISCRIMINATION): DISCRIMINATORY BEHAVIOR AND AMERICAN POLITICS

Partisan biases are strongly related to supporting hostile rhetoric about the opposition and avoiding opposing partisans, the first two levels of Allport’s framework. We now move on to evaluate the third level of prejudice: discrimination. Discrimination can manifest in many ways, from allocating different punishments for different groups, or, conversely, withholding rights or benefits from groups. Importantly, this level requires intentional actions designed to harm the opposing group. While study 1, study 2.1, and study 2.2 show evidence of behavior that might incidentally inflict harm, this is not the primary motivation (as we show in study 2.2). To explore the relationship between partisan affective polarization and the willingness of participants to make decisions that discriminate against the political opposition, we conduct two experiments in differing contexts. We construct our experiments to detect discriminatory preferences in situations where discrimination would violate core democratic norms. Regardless of the situation, the outcome remains the same: affective polarization is related to preferentially treating co-partisans but is not related to inflicting harm on the opposition.14 The least affectively polarized partisans generally treat the opposition no differently than the most affectively polarized. However, affective polarization is strongly related to helping or rewarding co-partisans.

Study 3.1: Tolerance for the suppression of political action

The right to protest is a fundamental part of American politics and is a key mechanism used to combat racial, ethnic, and gender discrimination (e.g., Etzioni 1970; Lipsky 1968; Shingles 1981). While protests can advance the interests of oppressed minorities, they can also galvanize negative views among those opposed to granting minority groups equality. We begin by testing how partisan biases affect responses to the suppression of political demonstrations—an example of where discriminatory behavior would be particularly troubling for the health of American democracy.

Participants (N = 395) read a newspaper story titled “Police Use Tear Gas on Peaceful Young [Democrat/Republican] Protest.”15 The story described a situation where police officers broke up a peaceful student protest. (See table 4.) Participants were asked two target questions. First, participants reported whether they “agree or disagree with the decision to use tear gas on the protesters.” Next participants were told, “The cost of the police response is unknown at this point, but the city can

### Table 3. Respondent-Identified Trait(s) Responsible for Exclusion

<table>
<thead>
<tr>
<th>Reason</th>
<th>Overall (%)</th>
<th>Q1 of Affective Polarization (%)</th>
<th>Q4 of Affective Polarization (%)</th>
<th>Significance (Q1 – Q4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>33.1</td>
<td>45.7</td>
<td>30.3</td>
<td>p = .47</td>
</tr>
<tr>
<td>Level of education</td>
<td>22.3</td>
<td>28.6</td>
<td>18.2</td>
<td>p = .82</td>
</tr>
<tr>
<td>Political party</td>
<td>43.3</td>
<td>42.9</td>
<td>48.5</td>
<td>p = .39</td>
</tr>
<tr>
<td>Martial status</td>
<td>8.3</td>
<td>11.4</td>
<td>3.0</td>
<td>p = .70</td>
</tr>
</tbody>
</table>

14. Results from studies 2.1, 2.2, and 3.1 show bias in decision making that many would classify as discrimination. Although not sufficient to reach the threshold of discrimination in the Allport framework of prejudice, the results, definitional debates aside, are an indication of strong partisan bias in the American public.

15. The study was fielded in 2013.
any amount up to $10,000. What amount, if any, do you think the city should fine the group of protesters?” Participants answered using a slider widget.

Only a small minority of participants agreed with the police action (i.e., gave responses above the midpoint), and this proportion is not significantly different when the protesters were in-party members or when the protesters were out-party members (15% vs. 20%, respectively). Participants tended to disagree with the police actions, with a small and insignificant difference ($b = -0.05$, 95% confidence interval $[-0.12, 0.01]$) in disagreement depending on whether participants were asked about the in-party or the out-party.16 The mean proposed fine toward the protesters is significantly lower in the in-party condition ($1,713.75$) than in the out-party condition ($3,000.21$), a difference of $1,286.46$ (95% confidence interval $[553.72, 2,019.19]$) or 57.12%.

The relationship between affective polarization and the suggested punishment is asymmetric in favor of co-partisans. The left panel of figure 3 shows that higher levels of affective polarization are strongly negatively related to agreeing with the police action when the target was a member of the in-party ($b = -0.31$, 95% confidence interval $[-0.52, -0.10]$). The right panel shows that affective polarization is not significantly related to agreeing with the police action when the target was a member of the out-party ($b = 0.06$, 95% confidence interval $[-0.17, -0.29]$). The differences between the slopes of affective polarization for both conditions is substantively larger than zero ($b = 0.37$, 95% confidence interval $[0.20, 0.54]$). Maximaly polarized participants, compared to minimally polarized participants, were more likely to oppose punishment of protesters from their party, but they also were no more likely to support punishing protesters from the opposing party.

Similarly, more affectively polarized participants recommended a lower fine when evaluating co-partisan protesters than less affectively polarized participants ($b = -6.06$, 95% confidence interval $[-8.82, -3.29]$; fig. 4, left panel).17 Affective polarization is not, however, related to the size of the fine when the protesters were members of the out-party ($b = .73$, 95% confidence interval $[-2.33, 3.82]$; fig. 4, right panel).18

While there was an asymmetry in outcomes—participants suggested less punishment when the protesters were co-partisans than when they were opposing partisans—the asymmetry was due to co-partisan favoritism, not a desire to harm the opposition (i.e., discriminate) against opposing partisans. We find no evidence that affective polarization predicts intent to discriminate against members of the opposition in this study.19

**Study 3.2: Tolerance for political corruption**

Scandals are fodder for both the media and political elites, but they represent a substantial challenge to democratic governance. Prior work shows that partisans respond differently to the scandalous or alleged criminal activities of co-partisans and opposition partisans (e.g., Dancey 2014; Peterson and Vonnahme 2014). Partisans, in general, are less likely today to believe in the guilt of co-partisan

---

16. The distribution of the fine is negatively skewed, but as results are not substantively different, we present difference in means and use OLS regressions instead of reporting the results of a quantile regression.

17. In fig. 4, we report the natural log of the fine amount. Both raw scores and other transformations give similar results.

18. We do see differences between Republicans and Democrats in this experiment (see appendix 4.6). However, because differences in response by party ID do not replicate in the next study and because the total N is small when the sample is split by party ID, we are not able to make strong conclusions about these results.

19. We repeated this study with 481 participants drawn from SSI with a treatment suppressing the identity of the group—the protesters were identified as students opposed to a tuition increase (rather than Young Democrats or Young Republicans). Nearly 33% of respondents strongly agreed with police actions when the group was not affiliated with partisanship, which is 13% higher than level of support for police intervention for co-partisans and 18% higher than the level of support for police intervention of opposing partisans.
of officials. But do such responses demonstrate a desire to prop up one’s own party or a desire to punitively attack the opposition? In this study, we replicate the results from study 3.1 but alter the domain to focus on a situation where actual wrongdoing has occurred. We use a scenario where a political elite has violated fund-raising laws to advance his political party. The structure of a newspaper article with randomly varied partisan actors is retained. (See table 5.)

Participants \((N = 475)\) recruited from the SSI panel read a faux newspaper article titled “Donations from Millionaire Businessman to [Republican/Democratic] Super PACs in Question.” The article reported that investigators were looking into possibly illegal donations that may have swayed an election. Participants indicated if they supported the investigation of the businessman on a seven-point bipolar scale ranging from “strongly oppose” to “strongly support,” which was recoded to range between 0 and 1.

Regardless of whether the target was a member of the out-party or the in-party, participants adopted the normatively “correct” position and favored an investigation of possibly illegal activities (see fig. 4). However, participants more fervently supported an investigation of an out-party member \((M = .74)\) than an investigation of an in-party member \((M = .61, 95\% \text{ confidence interval of the difference})\).
[\{-0.18, -0.08\}]. While there is in-party favoritism, it does not overcome the desire to investigate the wrong-doing of the middle-man, with the mean always above the midpoint on the disagree-agree scaler in the in-party condition (\$1,713.75) than in the out-party condition (\$3,000.21), a difference of \$1,286.46 (95% confidence interval [\$553.72, \$2,019.19]) or 57.12%.

High levels of affective polarization do move participants enough to be on the fence on the investigation of a co-partisan (b = -0.40, 95% confidence interval [-0.60, -0.21]; fig. 5, left panel). The probability of those at the highest levels of affective polarization supporting an investigation of a co-partisan is .49, while the probability for those at the lowest levels of affective polarization supporting the investigation is .89. Affective polarization is not at all related to supporting an investigation when the target is a member of the out-party (b = .05, 95% confidence interval [-0.18, 0.28]; fig. 5, right panel).

These results replicate study 3.1. While there is an asymmetry in attitudes toward the target of the investigation, this asymmetry is the result of co-partisan favoritism, not antagonism toward opposing partisans. Affective polarization is a reason to support co-partisans and not a source of intentional harm against opposing partisans.

Despite different contexts, these experiments show remarkably consistent results. Partisanship matters, but the relationship between partisan affect and decision making shows that partisans are more likely to help fellow partisans than they are to hurt those in the opposition. When placed in zero-sum decision tasks in other studies (picking a scholarship winner, a date, a spouse, etc.) partisans—especially those with higher levels of affective polarization—are more likely to select a co-partisan, but our results suggest that this is because of in-group favoritism and not out-group hatred.

**DISCUSSION**

We show that affective polarization, underpinned by the inherently competitive nature of partisanship, relates to some kinds of prejudicial behavior. However, affective polarization, like other in-group–out-group divides documented in psychology, is more about in-group love than out-group hate. While in line with years of psychology research, this is a surprising finding, as it is counter to recent work in political science on the nature of partisan bias. There is no doubt that partisan affective polarization is at historically high levels in America (Iyengar et al. 2012) and that hostility dominates politics (see Grimmer and King 2011; Sobieraj and Berry 2011). But the reality of partisan prejudice, we find, is likely more nuanced than most popular accounts. Partisan bias is significant, but partisan prejudice more consistently relates to behavior that benefits co-partisans than behavior that harms opposing partisans. Partisans willingly engage in some behaviors that might incidentally cause harm to the opposition (negative speech and avoidance), but they are not willing to impose harm (even when responding to a survey question where we would expect partisan bias to appear unfettered).

Throughout this paper, we use Allport’s framework to contextualize the effects of affective polarization on prejudicial behavior. But even for readers who are less convinced of the utility of Allport, our experiments demonstrate that partisan affect is related to behavior that benefits co-partisans but that is less conclusively related to a desire to directly harm (discriminate against) opposing partisans. Although the affectively polarized are more likely to deride and avoid the opposition, this behavior does not translate into bias against the opposition in situations where democratic rights are at stake.

While we do not find consistent evidence of out-group punishment, we do consistently find evidence of in-group
favoritism. The fact that we see such an asymmetry indicates that the emotions do not run as high as often portrayed. The two studies we used to test level 3 in Allport’s framework (discrimination) use a simple vignette design that, unlike work with trust games (Carlin and Love 2013; Iyengar and Westwood 2015) and work with actual cash payouts (Levendusky et al. 2015), include no incentive to temper partisan animosity. With no financial incentive to constrain partisan animus and no social norms to temper partisan hostility through social desirability bias, we should expect overwhelming evidence of partisan hate. The set of experiments presented here show that the effect of partisan animosity on behavior is perhaps less intense than prior work anticipates. This is consistent with decades of social psychological research that find a similar asymmetry in a variety of domains (e.g., Brewer 1999; Otten et al. 1996; Otten and Wentura 1999; Perdue et al. 1990). That is not to say that out-group punishment or derogation is not possible. Hostility (rather than favoritism) manifests when stronger emotions (e.g., fear) are attached to the out-group (Mackie, Devos, and Smith 2000). For instance, “an out-group that violates in-group norms may elicit disgust and avoidance; an out-group seen as benefiting unjustly (e.g., from government programs) may elicit resentment and actions aimed at reducing benefits; and an out-group seen as threatening may elicit fear and hostile actions” (Hewstone, Rubin, and Willis 2002, 580).

There is a link between affective polarization and promoting negative speech against the opposition, but the relationship is much stronger between negative attitudes and quashing negative speech against one’s own group than promoting hate speech against the opposition. We also find that affective polarization is related to seeking the company of co-partisans but that it is more strongly related to avoiding the out-party. In the Allport typology, these two forms of prejudice are significant but substantially less concerning than outright discrimination against opposing partisans, which causes material harm. That typology aside, the difference between more innocuous forms of prejudice—tolerance for negative speech about the opposition and avoidance of the opposition—is clearly distinct from discrimination as it is commonly understood and observed in other social divides such as race and gender. Partisan-based avoidance may create harm, which is troubling, but we show that avoidance and isolation are not an indication of more onerous discriminatory behavior.

Finally, we do not find consistent differences between the behavior of Democrats and Republicans, despite theoretical expectations to the contrary. Republicans were more likely to quash critical articles than Democrats, but they were not more likely to avoid the out-party. Additionally,
Republicans and Democrats appear to behave differently in the free speech experiment, but both groups behaved similarly in the corruption experiment. While some literature does indicate that Republicans should engage in higher levels of intergroup bias than Democrats, our results seem to indicate that these differences are not robust.

There are limitations to this work that should be considered. First, while our studies have high internal validity, concerns related to external and ecological validity must be addressed. The stimuli were contrived, and the participants were asked to make judgments about events they had just read about in a survey. Relatedly, our last set of studies do not use behavioral measures. On the one hand, the task was no less behavioral than those that asked respondents to make scholarship allocation decisions (Iyengar and Westwood 2015; Munro, Weih, and Tsai 2010) or studies that ask respondents to choose one candidate over another (e.g., Tomz and Van Houtweling 2008, 2009). On the other, we readily acknowledge the advantages of behavioral measures. We are comforted somewhat in the fact that the marginal effect of partisanship mirrors studies that use various economic games, such as those used in Carlin and Love (2013) or Iyengar and Westwood (2015).

Second, we look at how the affectively polarized respond to our treatments—an approach fully consistent with other work in political science that assesses the role of prejudice by assessing the interaction between attitudes and experimental conditions (see, e.g., Feldman and Huddy 2005; Mo 2014; Sniderman, Hagendoorn, and Prior 2004; Stenner 2005). Nonetheless, to truly identify the causal effects of prejudice, we would need to be able to randomly assign differing levels of affective polarization to respondents. This may be possible in the “minimal group” paradigm, but it is more difficult when identities are already formed.

Third, had real events similar to those described in the faux newspaper articles occurred, there would likely have been high-levels of media attention that could exacerbate possible partisan discrimination. There may be other factors that moderate the effects we observe. Future research should explore the role of respondent’s media diets, level of political interest, and past political engagement on partisan discrimination. Nonetheless, our results were assessed with a large number of participants, with realistic stimuli, and with nationally representative samples. There are counterexamples to the studies presented here, and many other manipulations that could and should be tested. Our work, nonetheless, produces results that are strikingly consistent, and when taken together, they present a clear pattern of behavior.

Finally, all our results focus on the American political system. The two-party structure of American politics and the growth of the permanent campaign creates a unique in-group and out-group dynamic that possibly magnifies partisan animus. Unlike other political systems with a multitude of parties aligned across the ideological spectrum that often form coalition governments, the United States is far more stable, which can lead to a stronger sense of group identity. However, recent work shows that in-group partisan identity is both common and relatively stable across various types of democratic systems with varying numbers of political parties (see Carlin and Love 2013, 2016; Westwood et al. 2015).

Our results have important implications for the modern American political arena. Political rancor permeates the political landscape, but while the rhetoric can be apocalyptic, most Americans, while willing to tolerate and support co-partisan preferential treatment, are not willing to punish or harm based on partisanship. Even the most affectively polarized—those with the strongest hatred for the opposition—are no more likely to intentionally inflict harm on the opposition than those with minimal affective political preferences. Wide tolerance for negativity is one explanation for why so many politicians and candidates are willing to engage in political attacks and utilize negative campaigning. But our results also show that there is a possible disconnect between elite behavior and citizen preferences. Citizens seem to approve of the negativity and bluster of representative rhetoric (Grimmer and King 2011), especially in uncompetitive districts. But citizens may not want their representatives to take it further, especially when their actions may violate the rights of members of the out-party.

Partisan redistricting represents one such disconnect. Ruling parties routinely gerrymander districts in a way that potentially violates the principle of one person, one vote (Raviv 2005). However, among citizens who are aware of redistricting (roughly half of citizens are not), large majorities believe it is unfair and would prefer that redistricting be left up to the courts or independent commissions (Fougere, Ansolabehere, and Persily 2010). It seems, then, that citizens would prefer that their representatives not discriminate against members of the out-party in the electoral arena. Similarly, support for voter identification laws drops when respondents are told that it keeps eligible citizens from voting (Wilson and Brewer 2013). Hostility is growing in the electorate, but for now it is largely avoidance, a tolerance for hostile speech, and support for preferential treatment of co-partisans.

ACKNOWLEDGMENTS

We thank Justin Grimmer, Shanto Iyengar, Neil Malhotra, Solomon Messing, Jonathan Mummolo, Paul Sniderman,
David Vannette and Rebeccca Weiss for helpful advice and encouragement. The authors contributed equally.

REFERENCES


