Polarization and Group Cooperation*

Andrea Robbett† Peter Hans Matthews‡

February 2021

Abstract

Does increased partisanship undermine the ability of politically heterogeneous groups to function and cooperate in apolitical settings? On the eve of the 2020 U.S. elections, we conducted an online experiment in which Democrats and Republicans played repeated public goods games, both with and without punishment. Absent punishment, mixed party groups are less cooperative and efficient than homogeneous groups. However, polarized groups fare no worse than those in which political affiliations are unknown. We find no differences in cooperation across groups that are able to punish free-riding behavior. Thus, knowing that one is in a group with likeminded individuals can serve as a substitute for an enforcement mechanism, but polarized groups can, at some efficiency cost, achieve similar contributions when sanctions are possible.

JEL Classification: C72, D9, H41

Keywords: Social dilemmas, political polarization, affective polarization, public goods, cooperation, experimental political science, experimental economics

*We are grateful to Corey Scheinfeld for stellar research assistance and to Sean Westwood and Eugen Dimant for useful discussions. Funding provided by Middlebury College.

†Department of Economics, Middlebury College, Middlebury, VT 05753. Email: arobbett@middlebury.edu.

‡Department of Economics, Middlebury College and Department of Economics, Aalto University School of Business. Email: pmatthew@middlebury.edu

1 Introduction

Have we reached the point where Democrats and Republicans can no longer work together to achieve common objectives in everyday apolitical interactions? The last four decades in U.S. politics have been characterized by steadily increasing partisan animosity and affective polarization within the general population (Iyengar, Sood, and Lelkes 2012, Abramowitz and Webster 2016), which is independent of ideological or issue-based differences (Mason 2015, Iyengar, Lelkes, Levendusky, Malhotra, and Westwood 2019). In 2020, the difference between reported like of one's own party relative to the opposing party reached its highest point since the "feeling thermometer" question measuring warmth toward each party was introduced to the American National Election Studies survey in 1978 (Finkel et al. 2020, Iyengar et al. 2019). The week before the 2020 election, a collective of prominent scholars of polarization proposed that our current divisions could be best defined as "political sectarianism," based not on ideological disagreements, but on dislike and distrust of opposing partisans, who are viewed as depraved and inherently "other" (Finkel et al. 2020). Even in decisions outside the political sphere, there is evidence that individuals treat co-partisans and opposing partisans differently, with implications for association and employment (Gift and Gift 2015, Iyengar and Westwood 2015, McConnell, Margalit, Malhotra, and Levendusky 2018, Dimant 2020).[1] This differential treatment sometimes manifests as antipathy for opposing partisans or, especially relevant for this paper, favoritism toward co-partisans.

At the same time, achieving sufficient cooperation to allow members to work together is essential for the functioning of most groups, communities, businesses, teams, neighborhoods, and organizations. Has our degree of partisan animosity undermined the ability of mixed-party groups to succeed? In this paper, we report the results of an online, pre-registered experiment,[2] conducted on the eve of the 2020 US elections, which was designed to test the

[1] Concern about political sectarianism has also proliferated in the popular press. For example, two months before Americans went to the polls, a New York Times headline asked "Can Love Survive This Election?" (Pajer 2020). Among those quoted in the story was Salvator Prano, the creator of *LiberalHearts.com*, who lamented the "whole new level of political polarization," as exemplified by the woman whose profile now announced "I divorced my husband. He voted for Trump!" The workplace isn't immune either: a recent survey revealed that "60% of employees felt that the election was impacting their ability to get work done" while "four out of 10 ... said that they have avoided coworkers due to their political views" (Jacobson 2021).

[2] The anonymized pre-registration plan can be viewed at: https://aspredicted.org/blind.php?x=67ye2k

extent to which politically polarized groups are able to work together to achieve common apolitical goals. To do so, we use the standard framework for studying group cooperation: a repeated linear voluntary contributions public goods game (Isaac, Walker, and Thomas 1984, Ledyard 1995). This game has been broadly applied to study a wide variety of group social dilemmas or collective action problems, in which what is best for the individual differs from what is best for the group as a whole, including volunteering for community projects, donating to charity, effort provision in work teams, or even compromising in marriage. Participants in our experiment are matched into groups of three and play a 10-period repeated game in the same fixed group, where partisan affiliation of each group member is either common knowledge or not revealed. To assess the extent to which groups are able to self-regulate and enforce cooperation, we additionally vary whether groups can engage in costly punishment (Fehr and Gächter 2000). Conducting this experiment in the days before the election provides a particularly strong test of whether cooperation can be sustained in polarized groups, since Michelitch and Utych (2018) show that citizens identify more closely with their political party as elections approach.

This paper builds on a relatively recent body of work assessing affective polarization and whether Democrats and Republicans treat each other differently in everyday interactions outside of the political sphere. Partisanship has emerged as one of the most salient social identities for many Americans (Huddy, Mason, and Aarøe 2015, Iyengar and Krupenkin 2018). There is evidence that partisans give more money to fellow citizens who are members of their own party rather than the opposing party in dictator and trust games games (Iyengar and Westwood 2015, Dimant 2020, Kranton, Pease, Sanders, and Huettel, 2020), have lower reservation wages when they believe their potential employer shares their political affiliation (McConnell et al. 2018), and, in hypothetical decisions, are more willing to award scholarships (Iyengar and Westwood), college admission (Munro, Lasane, and Leary 2010) or reduced price football tickets (Engelhardt and Utych 2020). A resume audit study has found that opposing partisans have a disadvantage in hiring decisions (Gift and Gift 2015). At the same time, Lelkes and Westwood (2017) found that partisan prejudice is primarily limited to bluster, avoidance of opposing partisans when forming groups, and favoritism toward co-partisans, rather than support for inflicting harm on opposing partisans.

The study that is perhaps closest to this paper is Dimant (2021), which examines whether players in a one-shot, two-person public goods game differ in their contributions, beliefs about contributions of their partner, and conditional contribution schedule depending on whether they are matched with someone whose opinion of Donald Trump or Joe Biden is aligned or misaligned with their own. The goal of this study was to assess the mechanisms underlying differential willingness to cooperate with people who share or do not share the same assessments of the presidential candidates, and not whether partisan identity affects groups' ability to sustain cooperation with and without punishment, and so our two papers are complementary. Dimant (2021) finds that participants' perception of closeness to their partner, amount given in dictator games, and amount contributed in public goods games are all significantly greater when matched with an in-group member compared to an out-group member; however, comparisons to a minimal group control condition and to a control with anonymous partners suggests that the perceptual closeness differences are attributable to in-group favoritism while the behavioral outcomes are due to out-group aversion. Additionally, players are similarly likely to be classified as unconditional cooperators, free-riders, or conditional cooperators regardless of who their partner is, but participants have more pessimistic beliefs about the contributions of those in the out-group, potentially suggesting that similar cooperation levels could be achieved if people could overcome their cynical assessments of opposing partisans.

In related work, Moore-Berg, Ankori-Karlinsky, Hameiri, and Bruneau (2020) provide an assessment of meta-perceptions between opposing partisans and find that, while Democrats and Republicans dislike each other, they also believe that the other side dislikes them even more. This underscores the potential importance of strategic interaction and, in particular, the possibility that cooperation could be fostered in repeated games in which participants receive feedback on the actual behavior of others.

It is therefore an open question as to whether partisan composition affects a group's ability to sustain cooperation. To this end, we report several main findings. First, without a mechanism for punishing free-riders, groups comprised of members of a single party achieve significantly higher cooperation than groups whose members belong to both parties. Second, this difference is due to increased cooperation in politically homogeneous groups, rather than

reduced cooperation in polarized groups. Mixed party groups and groups who do not know the affiliation of their group members generate similar contributions, which decline significantly over time, while homogeneous groups experience somewhat higher initial cooperation, which does not decline with repetition. When group members have a means of punishing each other, the advantage experienced by homogeneous groups disappears. We find that sanctions of free-riders are generally effective at increasing contributions, but that this is somewhat less true in mixed groups. At the same time, there are partisan differentials in punishment choices. As a result, heterogeneity in groups with punishment is something of a wash: some people are more likely to punish opposing partisans more for the same behavior, but the target of that punishment is also less likely to respond by increasing contributions. Last, we find that, relative to all other conditions, groups of co-partisans without punishment enjoy an efficiency premium of about 20 percentage points.

Our results thus suggest that affective polarization has not destroyed the ability of Democrats and Republicans to achieve cooperation in apolitical contexts: even during a week of heightened partisan animosity, groups comprised of Democrats and Republicans achieve similar cooperation and payoffs as those in which partisan affiliation is not revealed. Instead, we find a positive benefit of individuals knowing that they are in a group of likeminded partisans, and that politically homogeneity and punishment effectively are imperfect substitutes. This suggests that attempts to sort on the basis of partisanship (e.g., geographically as in Bishop 2009 or in work teams as in Lelkes and Westwood 2017) may actually lead to more efficient outcomes in situations when there is no other means of policing behavior and enforcing cooperation; however, when group members are able to monitor and sanction each other, cooperation can be achieved, at a cost, without matching co-partisans or avoiding opposing partisans.

2 Experimental Design

The experiment was conducted in the days leading up to the 2020 U.S. elections. Overall, 456 American political partisans participated, by playing a 10-period public goods games in fixed

groups of three.[3] We used a 2×2 design in which we independently varied whether participants knew the partisan affiliation of each group member and whether they were able to engage in costly punishment.

2.1 Recruitment of partisan participants

Participants were recruited and paid through Amazon's Mechanical Turk, which enables us to access a nationwide sample of Democrats and Republicans who participated synchronously and pay them differentially based on the outcome of the game. We took several measures to increase the likelihood that participants were engaged and attentive. First, we launched the study through the Cloud Research platform (Litman, Robinson, and Abberbock 2017) and screened out participants who failed Cloud Research's attention and engagement measures (along with limiting participation to those with a high past approval rating). Those who clicked on the task additionally had to pass through three separate screening stages to ensure that all participants in the game were attentive, American political partisans, who fully understood the public goods game and incentives. Before being able to participate, survey respondents first passed through a battery of four attention checks and anyone who failed any check was excluded. Those who passed the screener then filled out a one-minute survey of demographic questions, including party affiliation, for a base pay of \$0.25. Respondents who indicated that they were U.S. citizens and identified as Democrats or Republicans were then invited to participate in the experiment, and shown the instructions for the baseline public goods experiment (described in detail below). Finally, the potential participants were asked four comprehension check questions about the experiment (which involved calculating their own payoff under three different outcomes and a true/false question about the group matching) and only those answering all questions correctly ultimately entered the experiment. Notably, participants were sorted into treatments only after these screening stages were complete and the comprehension checks relied only on the baseline instructions to ensure participants were not differentially screened out across treatments.

[3] Approximately 62% of our participants were Democrats and 38% Republicans, a breakdown in line with most studies on mTurk.

2.2 The repeated public goods game

After completing the comprehension checks, participants were placed momentarily into a waiting room. As soon as two other participants were available, they were matched into a group of three and the group was randomly assigned to a treatment condition. Half of the groups were assigned to the *No Punishment* condition. In this case, they saw a summary page reminding them of the instructions and then began the game. The other half were assigned to the *Punishment* condition, in which case they received additional information about how to reduce the earnings of others, as described below. Independently, one fifth of the groups were assigned to the *Non-partisan* condition while four-fifths were assigned to the *Partisan* condition. We over-sampled the Partisan condition, since these groups further differ in the specific partisan breakdown. Those in the Partisan condition were shown the political affiliation of each group member prior to beginning the game.

In all four conditions, participants played a standard linear public goods game for ten periods (Isaac, Walker, and Thomas 1984). The number of periods and the fact that they were matched in a fixed group were common knowledge. The set-up and instructions for the *No Punishment* and *Punishment* conditions were directly adapted from Gächter, Renner, and Sefton (2008).

Each period consisted of two stages. First, each player received an endowment of 20 tokens and chose how many to contribute to a "group project." They were told that all contributions to the group project would be doubled and then divided equally among the three group members, while every token that they kept for themselves paid a return of 1. The payoff function for each individual i after the first stage is thus: $20 - e_i + \frac{2}{3}\sum_{j=1}^{3} e_j$, where e_i is player i's contribution and $\sum_{j=1}^{3} e_j$ is the sum of contributions in the three-person group.

In the second stage of the period, participants learned the total contributions in the group, their payoff from the first stage, and the individual contributions of each group member. In all conditions, the players were assigned a static ID for the duration of experiment (i.e., "participant A," "participant B," or "participant C"). The information about the player's randomly assigned ID was displayed at the very beginning of the experiment, at which time the participants in the *Partisan* condition also saw the affiliation of each person by their ID.

In the second stage of the period, each participant ID was listed on the screen, alongside the participant's contribution and earnings for the current period. In the *Non-partisan* condition, no additional information was given about the participants beyond their label. In the *Partisan* condition, each member's party affiliation was displayed below their label.

In the *No Punishment* condition, stage two was informative only: they viewed the contributions and earnings of each group member (and perhaps their party identification) and then clicked a button to progress to the next period. In the *Punishment* condition, each person made an additional decision in stage two: how many punishment points (described as "deduction points" in the instructions) to assign to each of the other two group members. They could assign each group member up to five punishment points. Each punishment point that the participant assigned reduced her own earnings by 1 and the earnings of the recipient by 3 (this punishment ratio was introduced in Gächter et al. 2008 and has subsequently become common place in the literature). In the event that the cost of punishment points received was greater than a participant's stage one earnings, their earnings were reduced to zero. However, participants still bore the full cost of punishment points they assigned to others. In this way, participants could not earn negative earnings purely as the result of punishments received, but could still earn negative earnings from their own punishment choices. Finally, each person learned the total punishments that they received (but not which group member assigned them) and their total earnings for the period after the punishments points in stage two were meted out. Specifically, participant's i final payoff for the period is given by $max\{0, 20 - e_i + \frac{2}{3}\sum_{j=1}^{3} e_j - 3\sum_{j\neq i} p_{ji}\} - \sum_{j\neq i} p_{ij}$, where p_{ij} is the punishment points assigned by player i to player j.

While nearly all of the protocols follow those established by Gächter et al., our design departs from their set-up in a few ways. First, we share the party identification of each group member in our *Partisan* condition. As such, we were unable to scramble the IDs of the group members in each period. Doing so would have ensured that participants' behavior could be tracked across all ten periods in some cases (whenever the two other members belonged to different parties) but not others (when the two other group members belonged to the same party). To avoid differences across partisan compositions and conditions, we thus fixed the IDs across all ten periods in all four conditions. This was known to the participants. Second, we

use a marginal per capita return of $\frac{2}{3}$ rather than 0.5, to simplify the description of the game to the participants (by explaining that contributions were multiplied by two and then equally distributed). Finally, we streamlined the instructions somewhat to appear on a single page and increase attention.

Participants received a flat payment of \$2 plus \$0.01 for each token earned in the game (which is comparable to Gächter et al. (2008)'s payment of £0.01 per token). The average payment was \$5.38 (including the base pay) and the game typically took 5-7 minutes once all group members had joined. The experiment was programmed in oTree (Chen, Schonger, and Wickens 2016).

3 Results

We begin by graphically presenting the overall group outcomes. We then formalize our results by analyzing individual contribution behavior and group outcomes in a regression framework. Finally, we consider punishment behavior and welfare results.

3.1 Group outcomes

We first assess the level of cooperation that groups of varying partisan compositions achieve with and without punishment. Figure 1 shows the average total group contribution over the course of the 10-period games in the *No Punishment* (left) and *Punishment* (right) conditions. In the absence of punishment, contributions are significantly higher in groups whose members all belong to the same party (denoted by square markers) than in groups comprised of both Democrats and Republicans (circle markers).[4] The difference arises in the very first period ($z = 2.418$, $p = 0.0156$) and, by the second half of the game, contributions in single-party groups are approximately 43% greater than those in mixed-party groups.

To ascertain whether these differences are due to individuals who know that they are interacting in a group of co-partisans contributing more than they otherwise would, individuals

[4]Treating each group of three interacting over the course of the 10 periods, we have $n = 25$ observations of single-party groups and $n = 35$ observations of mixed-party groups in the Partisan-No Punishment condition. A Mann-Whitney U test indicates that group contributions are significantly higher in single-party groups at the $p < 0.01$ level ($z = 2.74$, $p = 0.0061$).

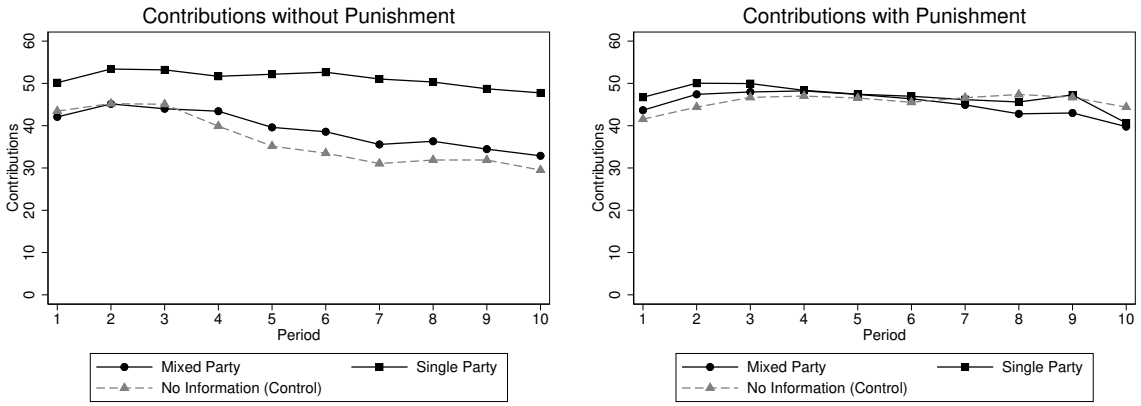

Figure 1: Group contributions over time without (left) and with (right) punishment

who know that they are in a group with members of the opposing party cooperating less than they otherwise would, or some combination, we consider the contributions in the *Non-Partisan* condition, where participants did not know the partisan identification of their groups members (shown by the gray line in figure 1).[5] Contributions in the non-partisan condition closely track those in mixed-party groups. There is no significant difference between contributions in mixed-party groups and those where members do not know the partisan affiliation ($z = 0.432$, $p = 0.667$), while contributions in single-party groups are significantly greater than the control ($z = 2.696$, $p = 0.007$). While we observe a significant decline in contributions over time in both the no information and mixed-party groups ($p < 0.01$), contributions do not decline significantly in single-party groups ($p > 0.20$).

The right panel of Figure 1 shows the outcomes among single-party, mixed-party, and no-information groups that can enforce cooperation by punishing free-riders. When groups have access to an enforcement mechanism, single-party groups no longer experience significantly greater contributions than those in mixed-party groups ($z = 0.147$, $p = 0.8828$) or groups without information on partisanship ($z = 0.34$, $p = 0.7341$), and there is likewise still no difference between mixed-party groups and the control ($z = 0.238$, $p = 0.8112$).

Our results thus far indicate that single-party groups achieve higher cooperation, but only

[5]Notably, there is no significant difference in contributions made by Democrats and Republicans either in this condition ($p = 0.72$) or overall ($p = 0.78$), indicating that this condition serves as an appropriate control comparison regardless of realized group composition.

in the absence of punishment opportunities, suggesting that partisan homogeneity is a substitute for an enforcement mechanism. In contrast, even polarized groups are able to achieve similarly high cooperation levels when they are able, at some cost, to punish free-riding behavior. These findings are confirmed by the regression results reported in Table 1.

Table 1: Group Contribution Outcomes

	(1) No Pun.	(2) Pun.	(3) No Pun.	(4) Pun.	(5) All	(6) All
Multi-Party	3.036	-0.953	4.517	-0.231	4.303	1.229
	(5.434)	(4.913)	(5.678)	(4.315)	(5.435)	(3.748)
Single-Party	14.49***	2.497	13.85**	1.613	15.00***	6.909**
	(4.989)	(4.782)	(5.293)	(4.673)	(5.045)	(3.219)
Number of Republicans	-1.897	1.445	2.755	-0.00665	-0.272	-0.272
	(2.003)	(2.306)	(4.200)	(3.432)	(2.533)	(2.537)
Period	-1.174***	-0.468**	-1.174***	-0.468**	-0.830***	-1.896***
	(0.268)	(0.199)	(0.269)	(0.200)	(0.170)	(0.590)
Punish					9.663*	-2.201
					(5.588)	(4.415)
Multi-Party x Punish					-4.563	3.369
					(7.112)	(5.633)
Single-Party x Punish					-14.59**	-1.638
					(6.911)	(5.432)
Period x Punish						2.157***
						(0.719)
Period x Single-Party						1.471**
						(0.671)
Period x Multi-Party						0.559
						(0.736)
Period x Single-Party x Punish						-2.355***
						(0.864)
Period x Multi-Party x Punish						-1.442
						(0.887)
Constant	45.12***	46.47***	38.41**	34.85*	31.92***	37.78***
	(4.142)	(4.234)	(15.25)	(18.62)	(12.10)	(11.81)
Controls	No	No	Yes	Yes	Yes	Yes
Observations	780	740	780	740	1520	1520

Standard errors in parentheses
Standard errors clustered at group level.
Columns (3) - (6) include controls for group demographics.
In all columns, the Non-partisan condition is the omitted condition.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In all the models reported in Table 1, there is one observation per group per period and standard errors are clustered at the group level. The first two columns regress the group contribution on indicators for whether both parties are represented in the group (*Multi-Party*) or

all members belong to the same party (*Single-Party*), with the *Non-Partisan* condition serving as the omitted condition. In the absence of punishment, we see that groups comprised of the same party achieve 14.5 tokens ($p < 0.01$) more in contributions (Col. 1), while there is no such effect with punishment (Col. 2). Columns (3) and (4) add demographic controls. Column (5) combines both punishment conditions into a single framework. The nearly equal and off-setting significant coefficients on *Single-Party* and *Single-Party* $\times$ *Punish* confirm that punishment wipes out the advantage of being in a single-party group. Finally, Column (6) interacts *Period* with our treatment variables and allows us to examine how the treatments affect contributions over time. In the omitted control condition (*Non-Partisan* and *No Punishment*), contributions decline significantly by almost two tokens ($-1.90, p < 0.01$) per period. This decline is significantly ameliorated when either all members belong to the same party (as indicated by the positive estimate of $1.47(p < 0.05)$ on *Period* $\times$ *Single-Party*) or when punishment is introduced (indicated by the positive coefficient $2.16(p < 0.01)$ on *Period* $\times$ *Punish*). Finally, the significantly negative coefficient on *Period* $\times$ *Single-Party* $\times$ *Punish* demonstrates that the two effects are substitutes, and that punishment is beneficial in moderating the decline in contributions only in the *Non-Partisan* and *Multi-Party* groups.

3.2 Individual contribution behavior

We next examine the determinants of an individual's decision to contribute and report the results in a series of regression models estimated in Table 2. The first two columns regress the individual's contribution on the number of co-partisans in the group (i.e., either 0, 1, or 2). The number of co-partisans has a significant positive effect $(2.36, p < 0.01)$ on contributions in the absence of punishment (Col. 1), but not $(0.527, p = 0.52)$ with punishment (Col. 2). However, this effect is not linear: being the lone member of one's party is no different from being in a group with one other co-partisan, and it is only when both other group members belong to one's party that we see a positive effect. Columns (3) and (4) regress the individual's contribution on indicators for whether there are zero, one, or two co-partisans in the group, with the *Non-Partisan* treatment serving as the omitted condition. Compared to knowing nothing about the party affiliation of their group members, individuals who know that they are with only

members of the other party or with one member of each party do not contribute significantly differently. However, individuals who know that they are in a group only with members of their own party contribute nearly five tokens $(4.82, p < 0.01)$more per period. Columns (5) and (6) include individual-level demographic controls and lead to similar conclusions. Column (7) includes observations from both punishment conditions and demonstrates that punishment has a positive overall effect, but offsets the beneficial effect of knowing that one is in a group of co-partisans.

Table 2: Individual Contribution Decision

	(1) No Pun.	(2) Pun.	(3) No Pun.	(4) Pun.	(5) No Pun.	(6) Pun.	(7) All
Number of Copartisans	2.359***	0.527					
	(0.861)	(0.825)					
Period	-0.319***	-0.208***	-0.391***	-0.156**	-0.391***	-0.156**	-0.277***
	(0.0981)	(0.0733)	(0.0892)	(0.0663)	(0.0893)	(0.0664)	(0.0567)
No Copartisans			0.645	-0.715	0.958	-0.316	0.834
			(1.854)	(1.643)	(1.907)	(1.523)	(1.872)
One Copartisan			0.950	0.0952	1.043	0.0542	1.216
			(1.792)	(1.552)	(1.779)	(1.443)	(1.774)
Two Copartisans			4.816***	0.413	4.784***	0.584	4.906***
			(1.677)	(1.568)	(1.686)	(1.476)	(1.669)
Republican					0.0450	-0.0852	-0.207
					(1.484)	(1.284)	(0.983)
Punish							3.051*
							(1.821)
No Copartisans x Punish							-1.365
							(2.426)
One Copartisans x Punish							-1.103
							(2.340)
Two Copartisans x Punish							-4.481**
							(2.253)
Constant	13.59***	15.79***	14.37***	16.08***	14.38***	13.00***	12.11***
	(1.330)	(1.209)	(1.273)	(1.217)	(2.442)	(2.333)	(2.006)
Controls	No	No	No	No	Yes	Yes	Yes
Observations	1800	1830	2340	2220	2340	2220	4560

Standard errors in parentheses
Standard errors clustered at group level.
Columns (1) and (2) include only Partisan condition.
In Columns (3) - (7) the No Partisan condition is the omitted condition.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3 Punishment: determinants and effectiveness

We turn next to punishment, and start with the observation that the sanctions required to sustain the cooperative contribution paths depicted in Figure 1 were quite modest. Across all punishment conditions, and counting both of the punishment opportunities available to each group member in each round in *Punishment* conditions, our participants acted on only 6.7% of these opportunities. Republicans (9.2%) did so more often, however, than Democrats (5.4%) but conditional on the decision to allocate punishment points, Republicans (2.9 points) were less harsh than Democrats (3.4 points). The first of these partisan differences dominates the latter: counting zeroes, the mean number of points meted out by Republicans (0.26) exceeds that by Democrats (0.18).

To understand their behavior in more detail, we estimate a two-limit Tobit specification, in which for each participant in each period there are two observations, corresponding to the two potential "targets" of individual group members, and report the results in Table 3.[6] Column (1) includes four regressors: *Target's contribution*, *Others' Average Contributions*, where "other" in this context refers to the punisher and non-target, whether or not the punisher is *Republican* and *Period*. We shall postpone, for a moment, a characterization of the separate effects on the intensive and extensive margins, but will note now that, consistent with previous research, the latent punishment variable decreases sharply in the target's contribution and increases in average contributions, with both effects significant at better than the 1% level. Furthermore, these effects are robust with respect to the addition of other controls. We also note that, congruent with our initial characterization of the data, the *Republican* indicator is positive, substantial and significant in all but one specification. Last, we discern in column (1) the first instance of another robust pattern, namely, the absence of any variation in punishment patterns over time.

We are especially interested to see whether the influence of partisanship extends to punishment, and Column (2) adds indicators for *Co-Partisan* and *Opposing Partisan* targets and their interactions with *Republican*, while Column (3) adds some additional demographic con-

[6]The substantial number of zeroes argues against the use of OLS, but we note that the results are qualitatively similar.

Table 3: Punishment Decisions

	(1)	(2)	(3)	(4)
PointsGiven				
Target's contribution	-0.926***	-0.931***	-0.936***	-0.630***
	(0.179)	(0.179)	(0.172)	(0.128)
Others' Avg Contributions	0.400***	0.401***	0.402***	0.552***
	(0.131)	(0.127)	(0.119)	(0.0913)
Period	-0.155	-0.149	-0.135	0.118
	(0.147)	(0.143)	(0.141)	(0.162)
Republican	2.943**	1.570	4.513*	5.279*
	(1.283)	(1.783)	(2.453)	(2.825)
Co-Partisan		-2.286	-2.375	-1.348
		(1.831)	(1.724)	(1.627)
Opposing Partisan		1.476	1.639	3.123**
		(1.524)	(1.393)	(1.543)
Republican x Opposing Partisan		-1.610	-2.536	-3.161
		(2.259)	(2.251)	(2.405)
Republican x Co-Partisan		4.181	2.971	0.674
		(2.908)	(2.845)	(2.876)
Constant	-7.428***	-6.435***	-10.58***	-14.74***
	(2.316)	(2.300)	(3.419)	(3.815)
Controls	No	No	Yes	Yes
Observations	4440	4440	4440	1650

Standard errors in parentheses
Tobit model of points given
Standard errors clustered at group level.
Column (4) includes targets who contributed less than their endowment.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

trols. The specification in Column (4) is identical to Column (3), except that we exclude cases in which the target contributed their entire endowment, punishment of which is "anti-social." All three share some common features. First and foremost, the coefficient on *Opposing Partisan* is large, positive and, in the last specification, significant at the the 5% level, while that on the interaction of *Opposing Partisan* and *Republican* is negative and of similar magnitude, if not quite significant at the 10% level. As our subsequent discussion of marginal effects will confirm, this suggests that relative to the default *No Partisan* condition, Democrats, in particular, allocate more punishment points to Republicans, but not vice versa. On the other hand, there is little evidence that relative to the default condition, Democrats or Republicans punish co-partisans any more or less.

To illustrate, and make more precise, the striking influence of opposing partisanship on punishment, the two panels in Figure 2 depict, respectively, the marginal effects on the likelihood of punishment and, conditional on punishment, the number of points chosen by Democrats and Republicans when the target is a member of the other party, by period. We see, first of all, that Democrats are more likely to punish Republicans than targets of unknown affiliation in the *No Partisan* condition, and that the effect is especially pronounced in the case of free-riders. In the case of a target who has contributed nothing, for example, Democrats are 4.9 *pp* ($p = 0.014$) more likely to punish a Republican. Reckoned against the 5.4% likelihood of Democratic punishment in the *No Partisan* condition, this is a very large effect. Even for a Republican target who contributes half their endowment, however, Democrats are still 3.9 *pp* ($p = 0.037$) more likely to punish. Republicans, on the other hand, are much more likely to inflict punishment (8.1% in the *No Partisan* condition) but are no more or less likely to allocate punishment points to Democrats in any period.

On the intensive margin, we observe that Democrats punish more (3.4 points) than Republicans (2.5 points) in the *No Partisan* condition, and the difference is significant at the 5% level. As Figure 2 reveals, Democrats punish significantly more when confronted with an opposing partisan and this declines only slightly with the target's contribution.

But is this punishment effective? That is, do those who are sanctioned increase their contributions? Consistent with previous literature, the answer is yes. In Table 4, we report

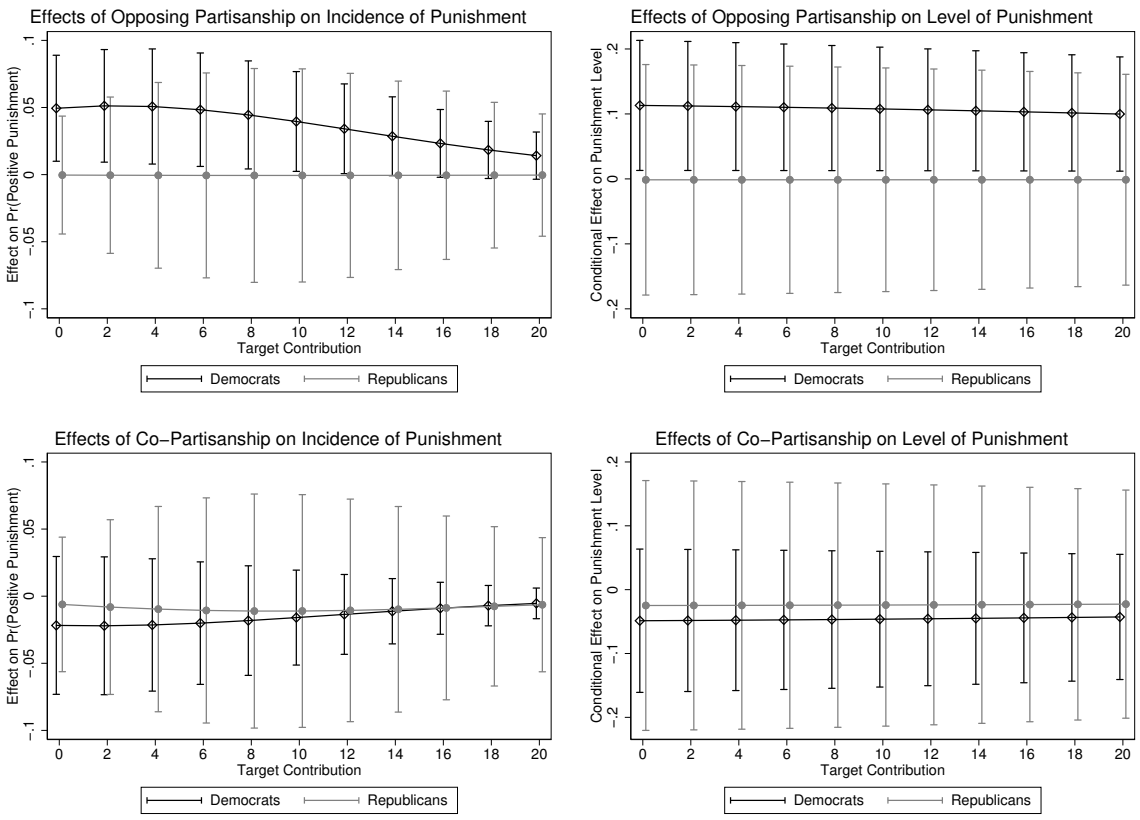

Figure 2: Marginal effects

the estimates of a series of regression models in which the outcome is the target's change in contribution from the previous period, denoted Δ*Contribution*, and the main explanatory variable is *Points Received*, the number of punishment points allocated to them in the previous period, with standard errors clustered at the group level.

Column (1), for example, is the simple bivariate regression of Δ *Contribution* on *Points Received*, and establishes what proves to be a robust conclusion, namely, that contributions do increase, significantly, after sanctions: across all conditions, each punishment point received is estimated to cause a 0.62 ($p < 0.001$) increase in contributions.

In Columns (2), (3), and (4), we decompose this effect with the addition of interactions between *Points Received* and *Single-Party* and *Multi-Party* and various controls, including *Period*. We see that people become less responsive to punishment over time. While the effect is marginal, we also observe that participants in multi-party groups are significantly less responsive to punishment. Column (5) replaces *Single-Party* and *Multi-Party* with indicators for the number of co-partisans and their interactions with *Points Received* and reveals a similar, if more nuanced, pattern. When individuals know that they are in a group with either one or two members of the opposing party, they are significantly less likely to correct their behavior after being punished (at the $p = 0.068$ and $p = 0.09$ levels, respectively). All of this together suggests that opposing partisans may punish one another more because sanctions are less effective.

3.4 Welfare effects

We have seen that for the punishment behavior just described, mixed partisan groups can sustain similar contribution paths to those that co-partisans achieve without punishment. (We leave aside for a moment the behavior of co-partisans when sanctions are available.) This does not mean, however, that welfare is also equalized across group types: because sanctions are not costless, the absence of partisan "fellow feeling" in mixed groups has welfare consequences. Our design allows us to measure the *welfare effects of partisanship*, at least in this stylized environment. To this end, consider first Figure 3, which plots group efficiency over time and across all conditions. Consistent with the experimental literature, we define efficiency as the group's realized payoffs, above and beyond the Nash equilibrium outcome, as a percentage

Table 4: Response to Punishment

	(1)	(2)	(3)	(4)	(5)
Previous Points Received	0.618***	1.273**	1.212***	1.220***	1.219***
	(0.141)	(0.482)	(0.444)	(0.439)	(0.440)
Multi-Party Group x Previous Points		-0.877*	-0.841*	-0.843*	
		(0.510)	(0.468)	(0.462)	
Single-Party Group x Previous Points		-0.480	-0.396	-0.390	
		(0.514)	(0.477)	(0.474)	
Multi-Party			-0.0649	-0.0510	
			(0.156)	(0.151)	
Single-Party			-0.262	-0.280	
			(0.219)	(0.239)	
Period			-0.166***	-0.166***	-0.166***
			(0.0393)	(0.0395)	(0.0397)
Republican				-0.106	-0.111
				(0.189)	(0.197)
No Copartisans					0.140
					(0.174)
One Copartisan					-0.133
					(0.167)
Two Copartisans					-0.279
					(0.240)
No Copartisans x Previous Points					-0.825*
					(0.480)
One Copartisan x Previous Points					-0.890*
					(0.481)
Two Copartisans x Previous Points					-0.390
					(0.474)
Constant	-0.376***	-0.380***	0.742***	0.320	0.306
	(0.0854)	(0.0838)	(0.242)	(0.389)	(0.386)
Controls	No	No	No	Yes	Yes
Observations	1998	1998	1998	1998	1998

Standard errors in parentheses
Dependent variable is the change in contribution (current contribution - previous period contribution)
Standard errors clustered at group level.
Includes all data from Periods 2-10 in Punishment condition
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of the difference between the socially efficient and equilibrium outcomes. That is, *Efficiency = (Group Payoff - 60)/(120 - 60)*, such that efficiency is 0 if the Nash equilibrium of zero contributions is realized and 1 if the efficient outcome of full contribution is realized. There is in all cases a small decline in efficiency over time, but as our initial observations suggest, what is most striking is the relative success of co-partisans in the absence of punishment who, averaged over all periods, achieve 85% efficiency.

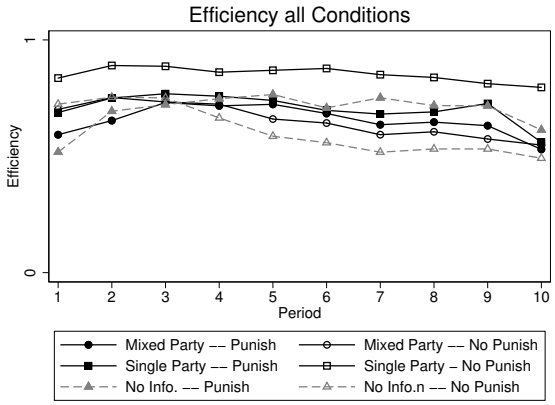

Figure 3: Group efficiency over time by condition

From this perspective, one plausible measure of these welfare effects would be the difference in efficiency between single party groups without punishment and mixed party groups with punishment. Averaged over all periods, the difference in efficiency was $20 = 85 - 65$ percentage points, a substantial effect. An alternative measure, the difference between single party and mixed groups in the absence of punishment, produces an almost identical estimate of 20 percentage points.

To explore these welfare costs in more detail, Table 5 reports the simple regression estimates for six models in which the outcome is group efficiency. Columns (1) and (3) are limited to cases in which punishment was not available, and include *Single Party*, *Multi-Party*, *Period* and *Number of Republicans* as main controls, where the difference is the inclusion of other controls in Column (3). The two sets of results are almost identical. Based on the estimates in Column (3), then, and relative to the default in which partisan affiliation was unknown, single party groups are estimated to cause a 23.1 *pp* ($p = 0.011$) efficiency gain. There appears to be a much smaller, and statistically insignificant, efficiency gain ($7.53pp, p = 0.429$) in mixed party

groups, and the difference between these effects is significant at better than the 10 percent level. It also seems that in the absence of punishment, efficiency declines steadily over time: the estimated effect is almost 2 percentage points per period ($p < 0.001$).

Table 5: Group Efficiency Outcomes

	(1) No Pun.	(2) Pun.	(3) No Pun.	(4) Pun.	(5) All	(6) All
Multi-Party	0.0506	-0.0379	0.0753	-0.0185	0.0732	0.0220
	(0.0906)	(0.104)	(0.0946)	(0.0906)	(0.0911)	(0.0635)
Single-Party	0.242***	0.00203	0.231**	0.0000599	0.251***	0.116**
	(0.0832)	(0.107)	(0.0882)	(0.109)	(0.0841)	(0.0547)
Number of Republicans	-0.0316	-0.0101	0.0459	-0.0518	-0.0280	-0.0280
	(0.0334)	(0.0602)	(0.0700)	(0.0804)	(0.0509)	(0.0510)
Period	-0.0196***	-0.00724	-0.0196***	-0.00724	-0.0136***	-0.0316***
	(0.00447)	(0.00564)	(0.00448)	(0.00567)	(0.00361)	(0.00984)
Punish					0.0940	-0.112
					(0.0989)	(0.0938)
Multi-Party x Punish					-0.101	0.0292
					(0.128)	(0.123)
Single-Party x Punish					-0.259**	-0.0264
					(0.127)	(0.118)
Period x Punish						0.0374**
						(0.0176)
Period x Single-Party						0.0245**
						(0.0112)
Period x Multi-Party						0.00931
						(0.0123)
Period x Single-Party x Punish						-0.0423**
						(0.0202)
Period x Multi-Party x Punish						-0.0237
						(0.0208)
Constant	0.752***	0.748***	0.640**	0.533	0.536**	0.635***
	(0.0690)	(0.0999)	(0.254)	(0.381)	(0.218)	(0.213)
Controls	No	No	Yes	Yes	Yes	Yes
Observations	780	740	780	740	1520	1520

Standard errors in parentheses
Standard errors clustered at group level.
Columns (4) and (6) include controls for group demographics.
In all columns, the No Partisan condition is the omitted condition.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (2) and (4) estimate the same models for those environments in which punishment was possible, and the differences are dramatic: there is now no advantage to being in either a single or mixed party group. Indeed, the efficiency gain associated with a politically homogeneous group is an almost precisely estimated zero. While it is perhaps not surprising

that mixed groups fare no better than those in the *No Partisan* condition, the question of why co-partisans do worse when punishment is introduced is a matter for future research. We conjecture, however, that this is another case in which the recourse to extrinsic incentives crowds out intrinsic, trust-based, motives.

To facilitate comparisons across punishment reasons, Column (5) uses the full sample, and adds the *Punish* indicator, as well as its interactions with the *Single-Party* and *Multi-Party* indicators, while Column (6) includes all of their respective interactions with *Period*. Consider the estimates in column (6). In the absence of punishment, the treatment effect of *Single-Party* is now $0.116 + 0.25 \times Period$ or $23.8pp (p = 0.003)$ in period 5, halfway through the game, relative to the *No Punishment/No Partisan* default. With punishment, the same effect is negligible: $0.006 = 0.116 - 0.026 + 0.25(5) - 0.042(5)$ or $0.6pp (p = 0.994)$ at the same point, where, predictably, the difference is significant at the 5% level.

And what about a mixed environment? In the absence of punishment, the effect of *Multi-Party* is now $0.022 + 0.009 \times Period$ or $6.8pp (p = 0.432)$, and with it, -2.1 *pp* $(p = 0.819)$, where the difference isn't statistically significant at any reasonable level. In short, diverse political groups seem to fare no better or worse than those in which affiliation is unknown. As observed in the previous section, some participants punish opposing partisans more harshly, but those sanctioned in mixed party groups are also less likely to raise their contributions in response.

4 Conclusion

In recent years, political pundits and academics alike have sounded the alarm on the increased partisan divide in the United States – a division driven less by ideological differences than by intense dislike and distrust for members of the opposing party. This animosity extends beyond simple expressions of dislike for opposing partisans on surveys and to interactions outside the political realm: People avoid associating with members of the opposing party – both socially and in the workplace, they are less generous toward out-group members, and they expect less generosity in return.

Despite the increased focus on affective polarization in recent years, the question remains

whether teams of mixed partisans who must work together in everyday interactions are able to overcome this animosity and sustain cooperation to accomplish a shared goal. In this paper, we contribute a controlled experimental test of this question. To do so, we employ the standard framework in experimental game theory for studying the determinants of cooperation in groups and teams: the repeated public goods games game. To further test whether partisan animosity undermines – or even enhances – groups' ability to self-regulate, we vary whether groups had the means of enforcing cooperation through a costly peer punishment mechanism.

We find that groups whose members share a common partisan affiliation sustain higher contributions to a public good – and experience higher welfare – than members of polarized groups. There are at least two reasons this result is less dire than first seems, however. First, contributions in polarized groups are similar in magnitude and trend to contributions in groups in which people do not know the political affiliation of the other members. This suggests that our core finding is due to increased cooperation among co-partisans rather than decreased cooperation with opposing partisans. It is possible, of course, that partisans who know nothing about group members conclude that at least one is likely to be an opposing partisan.

Second, when groups can enforce cooperation through costly punishment, we find that single-party groups lose their advantage, suggesting that co-partisanship and punishment are imperfect substitutes. The imperfection is that groups who enforce cooperation through punishment have lower welfare – because of the costs of norm enforcement – than homogeneous groups without punishment.

This in turn suggests that partisan sorting into more homogeneous groups – of the type observed in the United States (e.g., Lang and Pearson-Merkowitz 2015) – might be associated with benefits other than the direct benefits of shared affiliation inasmuch as it enables groups to sustain cooperation without reliance on costly sanctions.

5 References

Abramowitz, Alan I. and Steven Webster (2016). "The Rise of Negative Partisanship and Nationalization of U.S. Elections in the 21st Century." *Electoral Studies*, 41: 12-22.

Bishop, Bill (2009). *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart.* New York: Houghton Mifflin Harcourt.

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree: An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance*, 9: 88-97.

Dimant, Eugen (2021). "Hate Trumps Love: The Impact of Political Polarization on Social Preferences."

Engelhardt, Andrew M. and Stephen M. Utych (2020). "Grand Old (Tailgate) Party? Partisan Discrimination in Apolitical Settings." *Political Behavior*, 42: 768-789.

Fehr, Ernst and Simon Gächter (2000). "Cooperation and Punishment in Public Goods Experiments." *The American Economic Review*, Vol 90(4): 980-994.

Finkel, Eli J., Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, Mary C. McGrath, Brendan Nyhan, David G. Rand, Linda J. Skitka, Joshua A. Tucker, Jay J. Van Bavel, Cynthia S. Wang, and James N. Druckman (2020). "Political Sectarianism in America." *Science*, 370(6516): 533-536.

Gächter, Simon, Elke Renner, and Martin Sefton (2008). "The Long-Run Benefits of Punishment." *Science.* 322: 1510.

Gift K, Gift T. (2015). "Does Politics Influence Hiring? Evidence from a Randomized Experiment." *Political Behavior*, 37(3): 653-675.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe (2015) "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review*, 109(1): 1-17.

Isaac, R. Mark, James M. Walker, and Susan H. Thomas (1984). "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice*, 43(2): 113-149.

Iyengar, Shanto and Masha Krupenkin (2018). "Partisanship as Social Identity; Implications for the Study of Party Polarization." *The Forum*, 16(1): 23-45.

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: 129-146.

Iyengar, Shanto, Gaurav Sood and Yphtach Lelkes (2012). "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly*, 76(3):405-431.

Iyengar Shanto and Sean J. Westwood (2015). "Fear and Loathing across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science*, 59(3): 690-707.

Jacobson, Adam (2021). "Defusing Political Conflict in the Workplace" *Risk Management*, 1 February. http://www.rmmagazine.com/2021/02/01/defusing-political-conflict-in-the-workplace

Kranton, Rachel Matthew Pease, Seth Sanders, and Scott Huetteld. " Deconstructing Bias in Social Preferences Reveals Groupy and Not-Groupy Behavior." *Proceedings of the National Academy of Sciences*, 35: 21185-21193.

Lang, Corey and Shanna Pearson-Merkowitz (2015). "Partisan sorting in the United States, 1972?2012: New evidence from a dynamic analysis." *Political Geography*, 48: 119-129.

Lelkes, Yphtach and Sean J. Westwood (2017). "The Limits of Partisan Prejudice." *Journal of Politics*, 79(2): 485-501.

Litman, Lieb, Jonathan Robinson, and Tzvi Abberbock (2017). "TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences." *Behavior Research Methods*, 49(2), 433-442.

Mason, Lilliana. (2015). "I Disrespectfully Agree: The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science*, 59(1): 128-145.

McConnell, Christopher, Yotam Margalit, Neil Malhotra, and Matthew Levendusky (2018). "The Economic Consequences of Partisanship in a Polarized Era." *American Journal of Political Science*, 62(1): 5-18.

Michelitch, Kristin and Stephen Utych. (2018). "Electoral Cycle Fluctuations in Partisanship: Global Evidence from Eighty-Six Countries." *The Journal of Politics*, 80(2), 412-427.

Moore-Berg, Samantha L., Lee-Or Ankori-Karlinsky, Boaz Hameiri, and Emile Bruneau (2020). "Exaggerated Meta-Perceptions Predict Intergroup Hostility between American Political Partisans." *Proceedings of the National Academy of Sciences*, 117(26): 14864-14872.

Munro, Geoffrey D., Terell P. Lasane, and Scott P. Leary (2010). "Political Partisan Prejudice: Selective Distortion and Weighting of Evaluative Categories in Collge Admissions Applications." *Journal of Applied Social Psychology*, 40(9): 2434-2462.

Pajer, Nicole (2020). "Can Love Survive This Election?" *New York Times*, 25 August. https://www.nytimes.com/2020/08/25/fashion/weddings/can-love-relationships-survive-this-election.html

Online Appendix: Instructions

Screen 1. Thank you for your responses. We will now ask you to participate in the experiment. You will receive an additional flat bonus payment of $2 if you complete the game and could earn also earn bonuses up to an extra $4. You will be able to participate only if you correctly answer the comprehension questions in a couple of pages so please read the instructions carefully.

This is an experiment on decision making. Your earnings will be based on the decisions made by you and others. It is therefore very important that you read these instructions carefully. Both how you will be paid and the participation of other people are exactly as we describe in the instructions.

During the experiment, you will earn tokens. At the end of the experiment, you will be paid $0.01 US dollars per token you have earned (as described on the next page), plus the $0.25 for filling out the initial survey and the flat payment of $2 for completing the game.

This experiment is divided into different periods. There will be 10 periods in total.

Participants are divided into groups of three. You are with the same people in a group for all ten periods.

Screen 2. In each period you will be given tokens and have to decide how many tokens you would like to contribute to a project. Your earnings in the period will depend on the decisions made by you and the other two group members.

Each period consists of two stages:

In the first stage, you will receive tokens and have to decide how many tokens you contribute to a group project.

In the second stage, you will learn how much each member of your group contributed to the project.

The following pages describe the course of the experiment in detail.

Screen 3.

The first stage

At the beginning of each period, each participant in your group receives 20 tokens. In the first stage, you have to choose how many tokens you want to contribute to a group project and how many of them to keep for yourself. You can contribute any number of the 20 tokens to the group project. Each person will make this same choice.

Your earnings consists of two parts:

Your earnings from tokens kept: For each token that you keep for yourself you earn 1 token.

Your earnings from the group project: The tokens contributed to the group project are doubled and equally divided among the three group members. Thus your earnings from the group project are: (Total contributions to the project by all members) x 2/3

Your total earnings for the period is therefore:

(20 - your contribution to the project) + 0.67 (total contributions to the project)

The earnings of each group member from the project is calculated in the same way. Assume, for example, that the sum of the contributions of all group members is 9 tokens. This amount is doubled and then divided three ways, so each group member receives 9 x 2 / 3 = 6 tokens from the project, and your earnings for the period are equal to 6 plus however many tokens you personally kept.

Notice that, for each token which you keep for yourself, you earn 1 token. If, instead, you contribute this token to the group project, then the total contribution to the project will rise by one token. Your income from the group project will rise by 0.67 tokens. Moreover, each of the other group members? income from the project will rise by 0.67 tokens. Your contribution to the group project therefore also raises the income of the other group members. For each token contributed to the project the total earnings of the group will rise by 2 tokens. Similarly, you also earn tokens for each token contributed to the group project by the other group members. For each token contributed by any member you earn 0.67 tokens.

In summary, your earnings in tokens at the first stage of a period are equal to: 20 - your contribution + (2/3) (sum of contributions)

The second stage

Each member in the group will be assigned an ID: either Participant A, Participant B, or Participant C. This ID will be the same for all ten periods.

After you have viewed your earnings for the period, you will see an information screen showing how much each group member contributed to the project. Specifically, the screen will report the contributions as well as the first stage earnings for Participant A, Participant B, and Participant C.

After the second stage, the period is over and the next period commences.

You will be matched with the same group of people for all 10 periods. (True/False)

If each person keeps all 20 tokens (that is, contributes 0 tokens), what will your payment be for that period? (0/14/20/40)

If each person contributes all 20 tokens, what will your payment be for that period? (0/14/20/40)

If the other two people each contribute 20 tokens and you contribute 0 tokens, what will you payment be for that period? (0/14/20/40/47)

(Those who pass the comprehension check are then bounced to oTree and see the following screens)

Screen 4.

In your group, the three participants will be labeled: Participant A, Participant B, and Participant C.

These roles will the stay the same for all 10 periods.

You are Participant A.

[This part appears for those in the Partisan conditions with Democrat and Republican color coded as blue or red respectively]

Your group is as follows:
Participant A (you) – Democrat
Participant B – Republican

Participant C – Democrat

Screen 5 (No Punishment condition)

You will play the game for 10 periods. Each period will have two stages.

Stage 1: Each person will decide how many of their 20 tokens to contribute the group account.

Stage 2: After you have viewed your earnings for the period, you will see an information screen showing how much Participant A, Participant B, and Participant C contributed to the group account. You will then decide whether to reduce their earnings in the first stage by distributing points to them.

Screen 5 (Punishment condition)

You will play the game for 10 periods. Each period will have two stages.

Stage 1: Each person will decide how many of their 20 tokens to contribute the group account.

Stage 2: After you have viewed your earnings for the period, you will see an information screen showing how much Participant A, Participant B, and Participant C contributed to the group account and their earnings.

In Stage 2 you can decide whether to decrease the earnings of each other group member by assigning deduction points. The other group members can also decrease your earnings if they wish to. You will have to decide how many deduction points to assign to each of the other two group members. You can assign up to 5 points to each group member, or assign 0 points if you don't wish to change a group member's earnings. For each deduction point that you allocate to another group member his or her earnings are reduced by 3 tokens and your own earnings are reduced by 1 token.

Note that you will not be allowed to reduce the earnings of a group member to less than zero. After everyone has made a decision, you will be informed how many deduction points you received from the other group members and also what your total earnings in tokens for

that period are. Note that you do not get to know how individual group members spend their deduction points.

Your total earnings for the period will equal: Stage 1 earnings - 3 x deduction points received - deduction points allocated It is, in principle, possible that you make negative earnings in a period. However, you can always avoid negative earnings with certainty through your own choices.