Discussion Paper Series – CRC TR 224

Discussion Paper No. 238
Project A 01

The Cost of a Divided America: An Experimental Study Into Destructive Behavior

Wladislaw Mill [1]
John Morgan [2]

November 2020

[1] University of Mannheim, Department of Economics, L7 3-5, 68131 Mannheim, Germany, Email: mill@uni-mannheim.de, Telefon +49 621 181-1897
[2] morgan@haas.berkeley.edu; Haas School of Business, University of California, Berkeley, CA 94720, United States

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

The cost of a divided America: an experimental study into destructive behavior*

Wladislaw Mill¶ John Morgan‡

November 20, 2020

Does the polarization in the US lead to dysfunctional behavior? To study this question, we investigate the attitudes of supporters of Donald Trump and of Hillary Clinton towards each other and how these attitudes affect spiteful behavior. We find that both Trump and Clinton supporters have less positive attitudes towards the opposing supporters compared to coinciding supporters. More importantly, we show that significantly more wealth is destroyed if the opponent is an opposing voter. Surprisingly, this effect is mainly found for Clinton voters. This provides the first experimental evidence that the divide in the nation leads to destructive behavior.

Keywords: Spite; Voting; Experiment; Hillary Clinton; Donald Trump.
JEL: C91, D01, D62, D72, D74

*We thank Stefano DellaVigna, Noam Yuchtman, Ned Augenblick, Ori Weisel, Oliver Kirchkamp, Christoph Engel, Thomas Kessler, Rainer Rilke, Roi Zultan, Benedikt Werner, and participants at numerous conferences and seminars for helpful comments. We especially thank Michael Matthieu for logistic support. We gratefully acknowledge funding from the Max-Planck-Society through the IMPRS-Uncertainty and the German Research Foundation (DFG) through CRC TR 224 (Project A1).

¶University of Mannheim, Department of Economics, L7 3-5, 68131 Mannheim, Germany. Email: mill@uni-mannheim.de, Telefon +49 621 181-1897.

‡morgan@haas.berkeley.edu; Haas School of Business, University of California, Berkeley, CA 94720, United States.

1. Introduction

Fundamentally, people are tribal. While an individual person has many tribal identities, at particular times—certainly around the time of presidential elections—a person's main identity is political. This was notably true during the 58th presidential election.

In addition to anecdotal evidence, recent papers show that the polarization of the Democratic and Republican Parties is increasing and is higher than at any other time since the Civil War (Hare and Poole, 2014). A current survey by the Pew Research center also shows that partisanship division increased in recent years and not only on the party level (Pew Research Center, 2017a). The Pew Research Center (2017a) show that in 1994 more than one-third of self-identified Republicans were more liberal than the median Democrat—compared to just 5 % in 2017.[1] The same trend can be seen for Republicans. The study also shows that "Republicans and Democrats both say their friend networks are predominantly made up of people who are like-minded politically"(Pew Research Center, 2017a) and most of Democrats and Republicans state that they have few or no friends in the opposing party.[2] Further, due to social media, information bubbles are rising, which in turn might lead to an increasing polarization all around the world (Bakshy et al., 2015; Flaxman et al., 2016; Vicario et al., 2016).

This American divide manifests itself not only in opposing opinions and intellectual debates but also seem to result in aggressive behavior. Repeated reports of violence at rallies (Kite, 2016), increasing incidents of hate crime[3], and mounting fights between protesters – especially after the election of Donald Trump (Kite, 2016; Reilly, 2016; Shallwani et al., 2016) – are just a few examples of the division signified.[4]

Taken together, one might have the impression that the supporters of opposing parties were genuinely aggressive towards each other. However, the question is whether these reports are representative of the extremists of the parties or whether

[1] Media coverage of this topic can be seen, for example, in http://edition.cnn.com/2017/10/05/politics/pew-poll-ideology.

[2] Media coverage of this topic can be seen, for example, in http://edition.cnn.com/2017/10/05/politics/friends-political-party.

[3] A recent FBI report shows for example that the number of hate crimes rose in 2016, especially after the election of Donald Trump (see http://edition.cnn.com/2017/11/13/politics/hate-crimes-fbi-2016-rise).

[4] The Southern Poverty Law Center has kept track of hateful harassment and violence since the election of Donald Trump: https://www.splcenter.org/hatewatch. An interactive map with reported violence and hate speech after the election can be found here: (Manning, 2016) Further reports of attacks on both sides can be found here: Yan et al. (2016), here: Shallwani et al. (2016), here: Reilly (2016) and, here: Kite (2016).

negative attitudes also lead to aggressive actions for a broader part of society. We know that politically active people might behave aggressive and spiteful – such as parties ending up in gridlock – but what about the population at large? Have they also become two warring tribes willing to harm each other? And if so, do these two tribes differ in their aggressiveness?

Our research examines the degree to which tribalism reflects extremist views or occurs more generally in the population at large. In particular, we examine whether partisan preferences "spill over" into non-political, especially destructive behavior. To examine this, we recruit from the population of American online workers without revealing anything about our interest in partisan spillovers. We then compare individual choices, outside of the voting context but based on the (revealed) partisan preferences of the co-player. Our main question is whether partisanship produces spiteful decisions among the population at large outside of political contexts.

We, therefore, measured attitudes of supporters of Donald Trump and of supporters of Hillary Clinton towards coinciding voters and opposing voters. More importantly, we measured whether the participants were willing to harm their counterparts (by reducing their payoff). We do so by adopting a design where there is no trade-off between the own and the others' payoffs; that is, a setting of pure spite. To the best of our knowledge, we are the first to study spite in a political context, and more importantly, we are the first to study the "spill over" of partisan preferences into destructive behavior.

Previous research has already focused on whether partisanship affects non-political behavior. For example, Fowler and Kam (2007) demonstrates that partisanship influences decisions in dictator games. They show ingroup favoritism among co-partisans, in the sense that dictators share more money with copartisan recipients than non-partisans. In a similar vein Carlin and Love (2013) investigate the effect of partisanship on trust behavior. They show that similar to the literature on in-group favoritism, partisanship biases trust behavior in favor of co-partisans. While there is a trade-off between the own payoff and the other's payoff in dictator and trust games, we use a game where there is no benefit for the decider in destroying the payoff of the other. Thus, we look at a systematically different situation. While e.g., Fowler and Kam (2007) can show that partisanship influences decisions in dictator games, it seems like this behavior is to benefit the ingroup without harming the out-group. Yet, the dictator game does not allow for a distinction between outgroup-hate and ingroup-love. We use a design that identifies the dysfunctional aspect of behavior. Hence, the difference between a dictator game and our game is comparable to the difference between omitting help and directly harming others. Thus, unlike Fowler and Kam (2007) and Carlin and Love (2013) the main goal of

our paper is to study dysfunctional behavior directly.

Investigating the effect of partisanship on non-political behavior is also closely linked to the literature on ingroup-outgroup bias.[5] This literature investigates whether group affiliation (either induced or existing) leads to discriminative behavior towards ingroup and/or outgroup members. One of the main findings in this strand of literature is that people discriminate positively towards the ingroup (ingroup favoritism). However, hostile discrimination against the outgroup (outgroup hate) is rarely found.[6] One exeption using natural groups is Weisel and Böhm (2015), who showed that findings of outgroup hate are actually more a form of help avoidance and less direct harm to the outgroup.

The ingroup-outgroup bias was also recently introduced into an economic setting by the seminal work of Chen and Li (2009).[7] Chen and Li (2009) show that ingroup-bias transcends to economic decision making. More specifically they show that a match with an ingroup-member result in greater charity concerns and lower envy.[8] The most current advancements are made by Kranton and Sanders (2017) and Kranton et al. (2018). Kranton and Sanders (2017) show that people exhibit groupy and non-groupy preferences. More specifically, Kranton and Sanders (2017) use a minimal-group paradigm and political affiliation to investigate whether a person's social preferences change dependant on the type of matched partner. They show that 40% of participants exhibit no bias, i.e., the participants do not change their social preferences, while 60% of participants switch from one social-preference classification to another (e.g., selfish to inequality-averse). Interestingly, the results in Kranton and Sanders (2017) indicate that ingroup-bias might result in a higher likelihood of being classified as dominance-seeking. Thus our work can be seen as complementary to Kranton and Sanders (2017) and provides even stronger support for the importance of partisanship of changes in social preferences. Thus, while the literature so far has mainly provided compelling evidence of the ingroup-bias, there has been no sole-focus on the destructive side of this bias. Different from the previous literature, we focus particularly on outgroup hate in the form of direct harm.

[5] For some literature reviews see e.g. Hewstone et al. (2002); Greenwald and Pettigrew (2014) and see Riek et al. (2006) for a meta-analytic review of threat and outgroup negativity.

[6] e.g. Abbink et al. (2010); Chen and Li (2009); Charness et al. (2007); Bernhard et al. (2006); Tajfel (1970); Mummendey and Schreiber (1984); Brewer (1999); Riek et al. (2006); Vignoles and Moncaster (2007); Halevy et al. (2008); Greenwald and Pettigrew (2014); Weisel and Böhm (2015); De Dreu et al. (2016); Weisel (2015); Filippin and Guala (2013); Margolis and Sances (2016); Fowler and Kam (2007); Loewen (2010); Carlin and Love (2013); Iyengar and Westwood (2015); Alabastro et al. (2013); Weisel and Zultan (2016).

[7] Chen and Li (2009) can be considered part of the emerging field of identity economics (see the seminal paper by Akerlof and Kranton, 2000)

[8] See also the work of Chen et al. (2014), Chen et al. (2010) and Chen and Chen (2011).

While the question at hand has not been answered so far, there has been some research on antisocial behavior. A prevalent theme in economic decision making is the study of prosocial behavior. It has been shown very compellingly that people have prosocial preferences (see De Dreu, 2010; Fischbacher and Gächter, 2010; Murphy et al., 2011, for example), behave as conditional cooperatiors(e.g. Fischbacher et al., 2001; Herrmann and Thöni, 2009; Kocher et al., 2008) and are willing to sacrifices the own payoff for the good of another person (altruism)(e.g. Bernhard et al., 2006; Fehr and Gächter, 2002; Levine, 1998). However, the dark side of economic decision making is emerging recently. This rather new literature shows comprehensively that a significant amount of participants in laboratory and field experiments are not behaving payoff-maximizing (egotistically) or prosocially but are punishing antisocially (Herrmann et al., 2008), are burning money of others (Prediger et al., 2014; Abbink and Sadrieh, 2009; Sadrieh and Schröder, 2017; Abbink and Herrmann, 2011), are behaving spitefully (Fehr et al., 2008; Saijo and Nakamura, 1995; Kimbrough and Reiss, 2012; Bartling and Netzer, 2016) and are even willing to pay for the antisocial behavior (Kirchkamp and Mill, 2019; Zizzo and Oswald, 2001; Abbink and Dogan, 2019). We relate particularly to these papers as we specifically focus on spiteful behavior as a form of dysfunctional behavior.

In line with previous literature on destructive behavior, we find that some participants are willing to destroy the resources of coinciding voters. More importantly, however, we show that participants behave significantly more spitefully towards their voting counterpart (i.e., opposing voters) —they increase the probability of destroying an opposing voter's payoff by almost 17% relative to the probability of destroying a fellow voters' payoff — which constitutes our main finding. We also see that participants significantly dislike their voting counterparts. Concerning heterogeneity effects, we find that Trump voters do not behave more spitefully towards opposing voters—there is no statistically significant effect of their opponents' partisanship on their choices. By contrast, Clinton voters do behave increasingly more spitefully towards opposing voters—they increase the probability of destroying a Trump voters' payoff by almost 30% relative to the probability of destroying a fellow Clinton voter' payoff. The difference in behavior seems to arise from an asymmetry in the intensity of ingroup-outgroup bias. Clinton voters express strong antipathy toward Trump supporters, whereas Trump voters have a much weaker aversion towards Clinton voters. Thus, we offer three main results:

1. Participants report having more negative attitudes towards opposing voters.
2. Participants behave significantly more spitefully towards their voting counterparts.

3. Clinton and Trump voters differ substantially in their attitudes and behavior towards opposing voters.

Altogether, we provide the first evidence of spiteful behavior due to opposing political affiliation, and we find first empirical support for the destructive effects of political division.

The rest of this paper is structured as follows: In Section 2, we will explain the design of the experiment. Section 3 shows the results of the experiment. In Section 4, we conclude.

2. Design

In the following, we will present the recruitment, the procedure of the experiment, and the measures of interest.[9] The experiment was implemented using Qualtrics. We conducted the experiment in four waves: before the 58th US presidential election in late November 2016, after the inauguration of the president-elect in late January 2017, before the midterms in late October 2018 and after the midterms in early November 2018. Participants did not take part in multiple waves, which is why we do not have a panel data structure.[10]

2.1. Recruitment

Participants were recruited via Amazon's Mechanical Turk (MTurk) which is an online labor market and frequently used by social scientists for conducting experiments.[11] Workers in MTurk can choose from human intelligence tasks (HITs), and

[9] This paper is part of a bigger research project in the field of partisanship and economics. The focus of the companion paper (Mill and Morgan, 2018) lies on auction behavior and builds on the same approach – using partisanship – as this paper and uses the same participants, however in Mill and Morgan (2018) we focus solely on auction-theory testing. Participants made the auction decisions in Mill and Morgan (2018) before the spite measure was introduced, and no feedback was given in between. The reason we do not combine both papers is threefold: 1) both papers are aimed at a different audience, 2) combining both papers would make the paper too long and most importantly 3) the paper would lose its focus as both papers are aimed at very different questions.

[10] As there are no relevant differences between the waves for neither measure (neither the behavioral nor the attitudinal measure), we postpone the discussion of wave differences and all questions related to the wave effects into appendix C. In the main part of the paper, we pool the data over all waves and present the aggregate results.

[11] For example : Jordan et al. (2016, 2017); Peysakhovich et al. (2014); Rand et al. (2014); Suri and Watts (2011); Mao et al. (2017); Horton et al. (2011).

they will be paid by the requester after performing the task. These tasks are relatively simple and quick. Common tasks are answering surveys, transcribing data, classifying images, transcribing audio clips, translation rating and so on. (Horton et al., 2011; Berinsky et al., 2012; Paolacci et al., 2010; Mason and Suri, 2012).

More importantly, MTurk samples tend to be more representative of the US population than typical student samples as MTurk samples are usually more diverse in age, ethnicity, education, and geographical location than student samples (Buhrmester et al., 2011; Berinsky et al., 2012; Paolacci et al., 2010). Most importantly, the data obtained in MTurk is at least as reliable as data obtained via traditional methods (Buhrmester et al., 2011; Horton et al., 2011; Berinsky et al., 2012; Paolacci et al., 2010). In particular, in a recent paper Arechar et al. (2018) show that even interactive experiments can be conducted very reliably online and that behavioral patterns observed in the lab can be replicated using an online experiment with an MTurk sample.

In addition to the more diverse sample, there are several advantages of using an online design for our experiment. First, participants' anonymity can be sufficiently ensured, as we have only participants' MTurk-ID, which might result in more reliable results concerning antisocial behavior. Secondly, reciprocity concerns can be minimized as participants have no way of meeting the other participants nor figuring out who was assigned as their partner (which might be possible in a laboratory setting, which might bias behavior in a more pro-social direction). Third, peer effects can be excluded for the reason mentioned above.

To ensure a qualitative sample (i.e., participants understanding the task and paying attention), we restrict eligibility criteria. We restrict recruitment – as common in practice – to US-based individuals with an approval rate of 97% or higher.[12] Moreover, we restrict recruiting to individuals with approved HITs of more than 500. Further, individuals have to pass an attention check to take part.

2.2. Procedure

Participants willing to take part in the study were directed to the online survey tool Qualtrics, where they were asked about their vote in the 58th US presidential election. Only participants who indicated to vote for either Donald Trump or Hillary Clinton were directed towards the consent form. Undecided or independent voters were excluded from participation in the survey (as we do not have clear

[12]Participants' location will be verified through their IP addresses. Requesters can review the work done by MTurkers and decide to approve or reject the work. Approved work is paid as indicated in the contract, and rejected work is not paid. Hence, higher approval rates of workers indicate a higher quality of work.

predictions for them), and the survey ended for them. All of the remaining participants were directed, after reading the consent form, to answer socio-demographic questions (gender, age, income, education).[13]

Thereafter, the participants were presented with the experimental manipulation. The manipulation of the experiment was to let the participants either interact with a coinciding voter or an opposing voter.

The participants were told that at the beginning of the experiment, Trump voters were assigned the group color red, while Clinton voters were assigned the color blue. The manipulation across participants was to tell them which color their matched opponent will have (either red or blue).

To test comprehension and attentiveness, we asked whether participants understand the elements that appear on her screen, as some recent studies indicate the use of bots on MTurk. These simple questionnaire-elements include choices, payoffs, as well as information of her coplayer, i.e., whether their matched competitor was assigned the color blue, red, or green (which was a filler). Inattentive participants, those not comprehending the task, as well as potential bots were filtered out, as we are only interested in participants who have a basic comprehension of the task. Thus, failing to answer these questions correctly led to the exclusion of the experiment and the payment.

One potential concern of asking – as a basic comprehension test – which color the competitor was assigned to, might make the manipulation salient. There are several responses to this concern. First, it is crucial for this study that participants have a basic comprehension of the task and the situation. Second, there have been several comprehension questions which would diffuse the focus on this particular manipulation question.[14] Third, even if the question would result in a higher salience for the political position of the opponent, this would just result in a more realistic setting. Political attitudes and views are often presented and highlighted very saliently by real-world actors, as setting up yard signs, having political bumper stickers, wearing MAGA-hats, etc.[15]

[13] After answering the demographic questions, the participants first took part in an auction experiment without feedback, as reported in Mill and Morgan (2018).

[14] Note: even if experimenter demand is indeed induced, it is unlikely that the effect is very strong. de Quidt et al. (2018) estimate the demand effect for dictator games in case the experimenter explicitly tells participants what to do at less than one standard error. In case the experimenter just indicates his expectation, the demand effect is estimated at about .2 standard errors. In our experiment, participants *might* expect the demand to be a change in behavior. However, the design allows for several ways of doing so, and thus a demand effect would be even more restricted.

[15] The motivation of people engaging in such behavior in real life is manifold, and some motives might raise concerns for the research question if participants could endogenously choose how

Overall, we have a 2 (Own Vote ∈ {Clinton; Trump}) x 2 (Opponent's Vote ∈ {Clinton; Trump}) design.

After answering the control question, we measured the spite behavior towards their interacting counterpart (either coinciding voter or opposing voter), which is explained in greater detail in section 2.3. After completing the incentivized tasks, the participants had to answer a set of post-experimental questions.

To measure the participant's attitudes towards the opposing voters, we used an adjusted version of the social distance questionnaire (Crandall, 1991) and the feeling thermometer (Weisberg, 1980). We also elicited the general spite tendency by using a questionnaire (Marcus et al., 2014).[16]

At the end of the experiment, one of the items from the spite behavior was randomly selected to become payoff-relevant to ensure incentive compatibility.[17] The participants were informed that they would be paid within one week after determining the payoff, depending on their and their opponent's decision.[18] A graphical representation of the procedure can be seen in Figure 14.

2.3. Own Spite Measure

Our main measurement was aimed to mimic basic market interactions where spite has been observed. More specifically, our measure reflects a simplified and condensed version of a second-price auction where one player can reduce the payoff of the opponent by increasing the own bid (see Kimbrough and Reiss, 2012, for exactly such a situation). This measure consists of three distribution-decisions upon money. These distributions are shown in Table 1. We call this our own spite measure. We asked the participants to decide three times among nine possible allocations, similar to the SVO-Slider measure by Murphy et al. (2011). The participants were told that either their decision or their opponent's decision would be implemented, depending on a computerized random draw.

The first row resembles a situation in a second-price auction where the decision-maker will lose the auction but can increase the winning price, and thereby reduce the payoff of the opponent (a situation explicitly induced in Kimbrough and Reiss,

strongly to signal their attitude. However, the strength of the signal was kept constant across treatment as participants were merely informed about the voting decision of their opponent.

[16] The spite questionnaire is explained and analyzed in more detail in Appendix E.

[17] Hence, only one random problem was selected to become payoff relevant, which is arguably the only incentive-compatible mechanism (for a detailed argument see Azrieli et al., 2015).

[18] After finishing the collection, we matched the participants according to the instructions and paid them their bonus. The bonus payment was automated and implemented via AMS and R.

2012). The second row resembles a situation in a second-price auction where the decision-maker might win the auction. In this situation, the decision-maker can increase the winning price, and thereby reduce the payoff of the opponent in case the opponent wins. However, at the same time, the decision-maker also increases the chance of overbidding and, therefore, reduces the own payoff. The third row resembles a situation as in the first row but with a different endowment. At the same time, the third row reflects a simple situation of spite, where the difference in payoffs is increasing in spite. The aggregated rows are essentially aimed at representing spite in an auction-like market setting.

Note, that the spite behavior in the first row coincides with inequality aversion (see Kimbrough and Reiss, 2012). The second row could potentially also be driven by inequality aversion; however, the corresponding parameters would exceed the typically observed inequality-aversion parameters ($\alpha > 1$), and thus, the second row is better characterized by the preference of spite. The third row can only be explained by spite.

In all sets, the allocation with the highest payoff for the other player also maximizes the own payoff. However, any deviation from this allocation reduces the payoff of the other player and never increases the own payoff. In contrast to a standard dictator game – where there is a trade-off between the own payoff and the payoff of the opponent – in this game, the participants who do not choose the Pareto-efficient outcome do this in order to harm the other player. Therefore, any deviation from the Pareto-efficient outcome resembles spiteful behavior in a market setting and can be interpreted as spite or joy-of-destruction.

Table 1: Spite measure.

You receive	70	70	70	70	70	70	70	70	70
	○	○	○	○	○	○	○	○	○
Other receives	100	98	96	94	92	91	89	87	85
You receive	70	68	65	62	60	58	55	52	50
	○	○	○	○	○	○	○	○	○
Other receives	100	96	92	89	85	81	78	74	70
You receive	100	100	100	100	100	100	100	100	100
	○	○	○	○	○	○	○	○	○
Other receives	100	98	96	94	92	91	89	87	85

The table depicts the nine allocation choices in each of the three decisions of participants in our own spite measure. For each choice, the upper row denotes the payoff in experimental currency units for the deciding participants, while the bottom rows each denote the payoff for the other player.

In our measure, the *spite score* is the amount taken away relative to the maximally

possible amount. The amount taken away can range between 0 and 60 points (reducing the payoff of the opponent in all three distributions) and, therefore, the spite score ranges between 0 and 1.[19] As spite is arguably rather rare, we designed the task such that the incentive to behave spitefully is rather high. In particular, we decided to have low stakes (i.e., each point in the distribution represents 0.2 cents)[20] so that behaving spitefully is easily detected. This way, first, we potentially increase the chance of finding spiteful behavior, and second, any null-findings would have a greater bearing.

2.4. Measures of Attitudes

To measure the participant's attitudes towards their opponent, we utilized the social distance questionnaire and the feeling thermometer.

2.4.1. Social distance questionnaire

The social distance questionnaire is designed "to measure social rejection and willingness to interact with an individual member of a social group"(Robinson et al., 1999, p. 341 ff). In our experiment, the respective social groups were Trump voters and Clinton voters. The questionnaire elicits the agreeableness upon seven items on a scale between one and seven. The participants were asked to rate how strongly they agree with statements about a person. For example, participants were asked to indicate how strongly they agree with the following statements made about a Trump voter: "This appears to be a likable person" or "I would like this person to move into my neighborhood." The social distance score is the mean of seven item answers.[21] Higher scores indicate feeling closer to the individual member of the

[19] Note that using only the third decision as the main measure of spite leads to the same (and even stronger) results as presented below.

[20] Hence, the payoff for this task ranged between 3 and 20 cents. Thus, this experiment can be seen as a low stake experiment. Note that Kirchkamp and Mill (2019) use the same measure (in a very different context and with a student sample in the lab) while paying 6 euro-cents per point (thus, the payoff for this task ranged between 90 and 600 euro-cents). The distribution of choices in Kirchkamp and Mill (2019) is very similar to the patterns observed in this paper, indicating that the stake-size does not affect behavior substantially. Interestingly, Forsythe et al. (1994) and Carpenter et al. (2005) provide further compelling evidence that mean allocations in dictator games with low stakes do not differ from allocations in dictator games with high stakes. Additionally, Camerer and Hogarth (1999) survey the experimental economics literature and shows that behavior is impacted mainly if tasks are incentivized. Thus, by making the experiment having low stakes, we do not distort the results but rather nudge behavior into the direction of spite to have sufficient variance in the data.

[21] Note: we use the mean as this is common practice for the social distance measure (Parrillo and Donoghue, 2005; Robinson et al., 1999). However, using the median as the main statistic

respective social group.

2.4.2. Feeling Thermometer

The feeling thermometer is commonly used in polling (f.e. American national election studies), political sciences (Greene, 1999; Kaid et al., 1992; Miller and Wlezien, 1993) and also in medicine (Patrick et al., 1994; Jacobson et al., 1992; Schünemann et al., 2003). The feeling thermometer asks participants to imply how warm they feel towards a specific group or person. We asked participants to indicate their feeling towards Clinton voters, Trump voters, Republicans in general, and Democrats in general, on a scale between 0 and 10. Participants were told that if they had a positive feeling towards a group or feel favorably towards it, they should give it a score somewhere between 5 and 10, depending on their feeling. If they felt negatively, they should give a score between 0 and 5, and in case of no feeling, they should give a score of 5.

3. Results

3.1. Participants and Demographics

A total of 2616 participants (1428 females, 1188 males, 1621 Clinton voters, 995 Trump voters) finished the survey. The experiment was conducted with the online survey tool Qualtrics.[22]

As with most experimental studies, our sample does not perfectly represent the American population.[23]
The ages of our participants ranged from 18 to 88 years, with most participants in the age between 30 and 44 (48 %) and 24 %, 24 % of participants in the age between 18 and 29, 45 and 64, respectively (Median = 36). Hence, our sample is younger than the average American with a median age of 37.9 and with 15% of the population older than 65 years (compared to 4 % in our sample).
In addition, 55 % of our participants were female compared to 50.8% females in the US population.

produces the same results.

[22] To see whether participants paid sufficient attention during our study and whether participants' indicated vote for either Donald Trump or Hillary Clinton was sincere we controlled for coherence in Appendix A.1.

[23] For comparison estimates see the census aggregates: https://www.census.gov/quickfacts/fact/table/US/PST045216 and https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf.

Concerning the ethnicity in our sample: 83 % of participants are White; 6 % of participants are African American; 4 % of participants are Hispanic, and 6 % of participants are Asian compared to 61.3% Whites; 13.3% African Americans; 17.8% Hispanics and 5.7% Asians in the US population.

Moreover, our participants indicated to have higher education than the typical American. 69 % of participants implied to have at least a Bachelor's degree as the highest qualification compared to roughly 33% in the United States as a whole.

Hence, our sample is younger, more female, more white, and better educated than the average American.

In addition, looking at the location of the participants (see Figure 1), we find that the participants mainly come from populated and urban areas. This can also explain the discrepancy in the distribution of Trump and Clinton voters in our study (38 % vs. 62 %) compared to the distribution in the general election (46 % vs. 48 %).

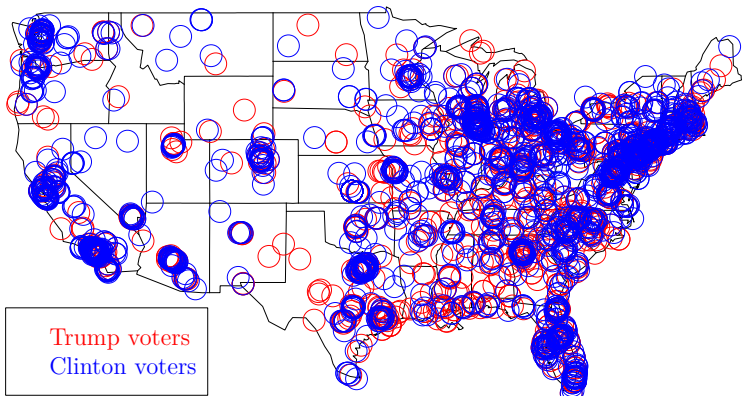


Figure 1: Participants' location by vote.
The figure depicts the location of our participants. Blue circles represent Clinton voters while red circles depict Trump voters.

Nevertheless, the participants in our study exhibit similar demographic voting patterns as reported in exit-pollings:[24] Trump voters in our sample are significantly less educated and are on average older than Clinton voters. Significantly more man voted for Donald Trump and significantly more white people voted for Trump than for Clinton. Table 2 shows the demographic differences between Clinton and Trump voters in our sample.

[24]See Alcantara et al. (2016) or Kirk and Patrick (2016).

Table 2: Demographics of participants.

Test	Clinton voters	Trump voters	T	Df	p	95% CI	Sign.
Female	0.56	0.53	1.39	2614.00	0.17	[-0.01 , 0.07]	
Age	37.72	40.18	-5.14	2614.00	0.00	[-3.39 , -1.52]	$***$
Race=White	0.80	0.89	-6.41	2614.00	0.00	[-0.12 , -0.07]	$***$
College-Ed or Higher	0.76	0.66	5.75	2614.00	0.00	[0.07 , 0.14]	$***$
Income > $70k	0.43	0.47	-1.80	2614.00	0.07	[-0.08 , 0]	.
Notes:						$^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;	

The table depicts summary statistics on Clinton and Trump voters in our sample, and compare whether demographic characteristics differ between the two voter types. Female is a dummy with value one if the voting participant is a female and zero otherwise. Age is a continuous variable denoting the age of the voting participant. Race=White is a dummy with value one if the voting participant indicated to be white and zero otherwise. Similarly, College-Ed or Higher, Income > $70k is a dummy with value one if the voting participant has at least a college degree, annual income of more than $70k, respectively, and zero otherwise. We use two-sample t-tests to compare characteristics.

Even though our sample does not fully represent the typical American, we are able to show the same tendencies as in the American population. Accounting for our selected samplewe see that Trump voters in our sample reflect Trump voters in the general election quite well, as do the Clinton voters in our sample compared to Clinton voters in the general election.

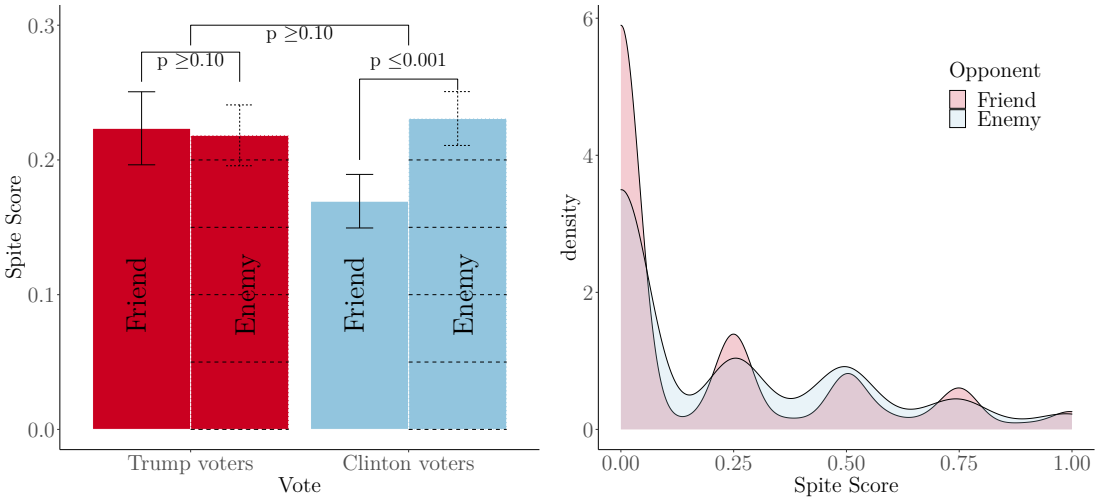
More specifically, an analysis of the voter data just after the election revealed that "in the 2016 election, a wide gap in presidential preferences emerged between those with and without a college degree. College graduates backed Clinton by a 9-point margin (52%-43%), while those without a college degree backed Trump 52%-44%"(Pew Research Center, 2016). In our data, college graduates voted for Clinton 51 % of the time while they voted Trump 42 %. People without a college degree in our data voted for Clinton 41 % of the time compared to Trump with 57 %.
Moreover, the analysis shows that "older voters (ages 65 and older) preferred Trump over Clinton 53%-45%."(Pew Research Center, 2016). In our data, the numbers are 56%-42%.
In addition, women supported Clinton over Trump by 54% to 42% (Pew Research Center, 2016). In our data, the margin is 49%-45%.

Further, young adults (18-25) preferred Clinton over Trump by a wide 55%-37% margin (Pew Research Center, 2016). In our data, the margin is 52%-40%.
The analysis by the Pew Research Center (2016) also shows that "Trump won whites with a college degree 49% to 45%" and he won whites without a college

degree 67% to 28%. In our data, Trump won whites with a college degree 45% to 49%, and he won whites without a college degree 60% to 39%.

Thus, our selected sample shows a striking similarity to the general populations' patterns and reflects the attitudes of general Clinton and Trump voters rather reliably.[25]

3.2. Spite



(a) Spite towards opponents

(b) Distribution of the Spite-Score by opponent.

Figure 2: Results of the spite-score.
The figure on the left depicts how spitefully participants behave towards their friends (i.e., coinciding voter) and enemies (i.e., opposing voter). The left two red bars show the spite behavior of Trump voters while the right two blue bars show the spite behavior of Clinton voters. Bars with the label "Friend" denote the spite behavior towards friends (i.e., coinciding voters) while bars with the label "Enemy" denote the spite behavior towards enemies (i.e., opposing voters). Tie fighters depict 95% confidence intervals. P-values are calculated using t-tests.
The figure to the right depicts the distribution of the spite score by opponent. The red distribution denotes the spite behavior toward friends while the blue distribution denotes the spite behavior towards enemies.

In the following, we will refer to all those opponents who were assigned the same color as the deciding participant as "friends" (Trump voter vs. Trump voter; Clinton voter vs. Clinton voter), opponents who were assigned a non-matching color as "enemies" (Trump voter vs. Clinton voter and vice versa).

[25] To deal with potential selection effects, we reestimate all regressions by reweighting our sample to make it more representative in Appendix B.2. All results prevail.

To see how participants are behaviorally influenced by their opponent, we examine the spite score in this section. A distribution of the spite score is shown in Figure 2b.[26] Figure 2a depicts the aggregated spite behavior by opponent and vote.

3.2.1. Estimation approach: zero-inflated beta regression

The spite score is a proportion between 0 and 1, and it represents how many points were taken away from the opponent relative to the maximally possible amount. A common approach for this kind of data is a beta regression.[27] A beta regression assumes that the data is distributed by a beta distribution—in contrast to the standard approach for non-proportion data which assumes a normal distribution. In particular, the problem with estimating proportions with a Gaussian distribution is that the support of the Gaussian distribution is the whole set of real numbers, while a proportion is a number between 0 and 1. Hence, the beta distribution (with support between 0 and 1) is the better estimation choice on theoretical grounds. In addition, we also discuss further estimation possibilities in A.5 and show that the beta distribution performs best in estimating the spite-score. Hence, there are theoretical and practical reasons to explain why the beta regression is more appropriate than the normal distribution.

Another issue is that the spite score has an excessive amount of zeros, as can be seen in Figure 2b. To deal with this, we will use a zero-inflated beta model regression. The zero-inflated model estimates the decision to be spiteful in two parts: first, it estimates whether a participant decided to be spiteful or not (using a logistic regression); and second, it estimates conditionally on deciding to be spiteful, how spiteful participants decided to behave (using a beta regression). This allows us to use the data more efficiently and, more importantly, is less biased. The following example should clarify why a zero-inflated model gives a better picture than a model ignoring zero inflation. Imagine two very different groups, A and B. In group A, 80 % of participants never behave spitefully, but 20 % do, and these 20 % behave 100 % spitefully. In group B, 100 % of participants behave 20 % spitefully. If we would just use a simplistic approach and ignore the zero inflation, then we would find no significant difference between these two groups because both groups would show a spite score of .2 on average. However, this approach ignores

[26] We also examine the properties of the spite score in Appendix A.3. It can be seen that participants with better attitudes towards their opponent behave less spitefully, and that participants with a higher score on the spite-personality measure also behaved more spitefully.

[27] See Jönsson and Thor (2012); Seow et al. (2012); Laliberté et al. (2012); Rogers et al. (2012) for applications of the beta regression. For theoretical papers, see: Smithson and Verkuilen (2006); Elgers et al. (1979); Cribari-Neto and Zeileis (2010); Grün et al. (2012); Schmid et al. (2013); Ospina and Ferrari (2012).

essential information and does not reflect the situation adequately. In contrast, a zero-inflated model would show that group A is significantly more likely not to behave spitefully compared to group B and that conditional on the spiteful behavior participants of group A are significantly more likely to be fully spiteful. Hence, taking zero inflation into account gives us a clearer picture of the spite behavior.

Therefore, we use a zero-inflated beta model regression to estimate our spite score. Nevertheless, we additionally include more ordinary estimation methods in Appendix B, where the main results can also be found together with the more common estimators, such as OLS and Tobit regressions. All results prevail if we use a standard linear regression or a Tobit regressions on the spite score. Thus, our results do not rely on the zero-inflated beta regression being used.

3.2.2. Estimation

We stipulate that our conditional spite score (y) follows a beta distribution with mean μ and precision σ, i.e. $y \sim \mathcal{B}(\mu, \sigma)$. Here, $\mathbb{E}(y) = \mu$ and $\mathbb{V}ar(y) = \mu(1+\mu)/(1+\phi)$. The beta density function is described by:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}$$

with $0 < \mu < 1$, $0 < y < 1$[28], and $\phi > 0$.

Let $y_1, ..., y_n$ be our observations of the conditional spite score, where every $y_i \sim \mathcal{B}(\mu, \sigma), 1 = 1, ..., n$.

Hence, the econometric model we use is described by:

$$\begin{aligned} g(\mu_i) &= \sum_{j=1}^{k} x_{ij}\beta_j \\ g(v_i) &= \sum_{j=1}^{k} x_{ij}\alpha_j \\ v_i^* &= 1 \text{ if } v_i > 0 \\ \mu_i^* &= \mu_i \cdot v_i^* \end{aligned} \tag{1}$$

[28]Note: Because y is stipulated to be strictly smaller than 1, we change the spite score of 1 to $1 - 10^{-5}$, as suggested in Smithson and Verkuilen (2006). Furthermore, because the conditional spite score is per definition bigger than 0, we do not need to adjust y at 0.

where $\beta = (\beta_1, ..., \beta_k)^T$ and $\alpha = (\alpha_1, ..., \alpha_k)^T$ are vectors with the unknown regression parameters ($\beta, \alpha \in \mathbb{R}^k$) and $x_{i1}, ..., x_{ik}$ are the observations. $g(\cdot)\colon (0,1) \to \mathbb{R}$ is the link function. We will use the logit link: $g(\mu) = log(\mu/(1-\mu))$, as this is easier to interpret. We will assume $x_{i1} = 1$ to be the intercept.

3.2.3. Estimation Results

Our main model, which we will denote by Ψ, estimates the Equation 1 with $x_{i1} = 1$ and $x_{i2} =$ Enemy, where Enemy is a dummy variable with value one if the opponent is an enemy and zero if the opponent is a friend.[29] This model aims to directly answer the question of whether or not partisanship (more specifically, a divided nation) leads to dysfunctional behavior. The resulting estimation is reported in Table 3 in Columns 1 and 2.

We can see that spite behavior towards enemies is significantly more likely than spite behavior towards friends. Hence, the probability of a participant to behave spitefully towards an opposing voter ($\frac{e^{-0.4 \cdot 0.3}}{(1+e^{-0.4 \cdot 0.3})} = 0.47$) is significantly higher than the probability of a participant to behave spitefully towards a fellow voter ($\frac{e^{-0.4}}{(1+e^{-0.4})} = 0.4$). Thus, the probability of spiteful behavior increases by 17 % if a participant is to interact with an opposing voter.

However, the intensity of spite, conditional on being spiteful, does not increase significantly if the deciding participant is interacting with an enemy compared to a friend. The probability of behaving fully spitefully conditional on being spiteful towards friends is ($\frac{e^{0.22}}{(1+e^{0.22})}$) 0.56 and it is ($\frac{e^{0.22 \cdot 0.02}}{(1+e^{0.22 \cdot 0.02})}$) 0.56 towards enemies.

Hence, the main result of this paper is:[30]

> **Result B1 (Main Result)** Participants are significantly more likely to behave spitefully towards an opposing voter compared to a coinciding voter.

Thus, the key insight is that the participants are more likely to behave spitefully towards people who voted differently in a non-political situation.[31] Hence, partisanship spills over into the non-political realm.

[29] Several additional robustness checks, such as including sociodemographic controls or taking information about crime, poverty, and religion of the county of the participants into account, can be found in Appendix A.4. The results prevail in all models.

[30] Our results are split into two categories: behavioral and attitudinal. Thus, the first result is labeled B1 as we focus first on the behavioral results and discuss the attitudinal results thereafter.

[31] Note: the results also allow for an alternative interpretation, namely that partisanship reduces spitefulness. While this interpretation is also valid, the key insight of this paper, i.e., parti-

To investigate whether there are differences between Trump and Clinton voters in their behavior towards enemies we estimate Equation 1 with $x_{i1} = 1$, $x_{i2} =$ Enemy, $x_{i3} =$ Trump voter and $x_{i4} =$ Enemy$\cdot$Trump voter. Here Trump voter is a dummy variable with one if the participant voted for Donald Trump and zero if she voted for Hillary Clinton. Similarly, Enemy is a dummy variable with value one if the opponent is an enemy and zero if the opponent is a friend. The results can be found in Table 3 in Column 3 and 4. Columns 5 & 6 report the main model (Ψ) for Trump voters and Columns 7 & 8 report the main model for Clinton voters separately.[32]

It can be seen from Table 3 Column 3 that the probability of a Clinton voter to behave spitefully towards a Trump voter is significantly higher (0.47) than the probability of a Clinton voter to behave spitefully towards a Clinton voter (0.36). Once decided to behave spitefully (Column 4), the probability of a Clinton voter to behave fully spitefully (taking away all points) towards a coinciding voter (0.57) did not differ significantly from the probability of a Clinton voter to behave fully spitefully towards a Trump voter (0.59). Concerning Trump voters, we see that they are significantly more likely to be spiteful towards a coinciding voter (0.48) compared to the Likelihood of a Clinton voter behaving spitefully to a coinciding voter (0.36). We can also see a significant interaction of vote and enemy on the decision to be spiteful. Thus, while Clinton voters increase the probability of spiteful behavior towards enemies significantly from 36 % to 47 %, i.e. a relative increase of 31 %, Trump voters have no significant increase in their probability of spiteful behavior towards Clinton voters (0 %). Conditional on behaving spitefully, the interaction of vote and enemy was not significant. Hence, the probability of a Trump voter to behave fully spitefully to a non-partisan (0.51) was not significantly smaller than the probability of Clinton voter to behave fully spitefully to a non-partisan (0.59).

Due to the significant interaction effect, we split the data with regard to the vote (as can be seen in Columns 5-8 in Table 3). It is evident that Clinton voters are more likely to behave spitefully towards enemies compared to friends. This pattern cannot be found for Trump voters. Trump voters do not significantly differentiate between Trump voters and Clinton voters neither in their decision to

sanship affects spite behavior, does not change. We prefer our interpretation as we typically expect no or little spite behavior to be observed. For example, Lange et al. (1997), and Murphy et al. (2011) show that only 7%, and 4% of participants have dominance seeking preferences, respectively. Furthermore, using the same measure as here Kirchkamp and Mill (2019) find only a spite score of 0.05 on average in a standard student sample. Thus, higher values on our spite score are more the exception than the norm.

[32] To deal with potential selection effects between Clinton and Trump voters, we reestimate all regressions using propensity score matching in Appendix B.1. All results prevail.

behave spitefully nor in their conditional behavior.

Overall, Clinton voters were more likely to behave spitefully towards enemies relative to friends than were Trump voters. Trump voters differ substantially in their behavior from Clinton voters as they are do not differentiate between friends and enemies. This means that Clinton voters engaged in more dysfunctional behavior than Trump voters if paired with an enemy.

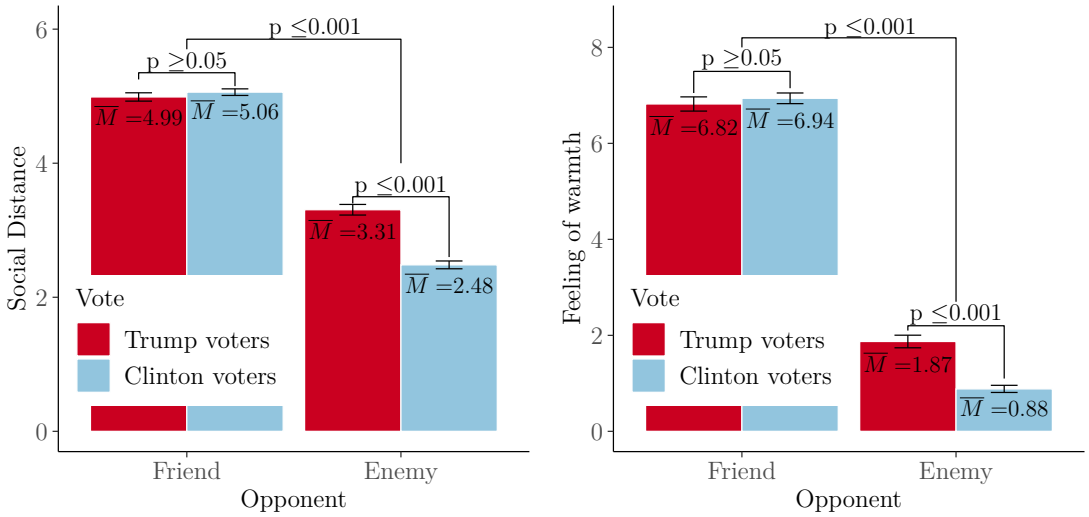
Result B2 Clinton voters are more likely to behave spitefully towards Trump voters compared to fellow Clinton voters. Trump voters, on the other hand, do not differentiate in their behavior between Clinton and Trump voters.

Table 3: Estimation of the spite behavior

	Spite?	Spite Score	Spite?	Spite Score	Spite?	Spite Score	Spite?	Spite Score
	Full sample		Full sample		Trump voters		Clinton voters	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-0.40^{***}	0.22^{***}	-0.59^{***}	0.29^{***}	-0.07	0.15^{*}	-0.59^{***}	0.27^{***}
	(0.06)	(0.05)	(0.08)	(0.07)	(0.10)	(0.08)	(0.08)	(0.07)
Enemy	0.30^{***}	0.02	0.47^{***}	0.09	0.01	-0.11	0.47^{***}	0.08
	(0.08)	(0.07)	(0.10)	(0.09)	(0.13)	(0.10)	(0.10)	(0.09)
Trump voter			0.52^{***}	-0.15				
			(0.12)	(0.11)				
Enemy x Trump voter			-0.46^{**}	-0.19				
			(0.16)	(0.14)				
Model	Logistic	Beta	Logistic	Beta	Logistic	Beta	Logistic	Beta
Observations	2,616	1,156	2,616	1,156	995	480	1,621	676
Log Likelihood	$-1{,}788.64$	64.64	$-1{,}779.69$	72.56	-689.06	2.33	$-1{,}090.62$	82.56

Notes: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;
This table depicts a zero-inflated beta regression model of the spite behavior. Enemy denotes a dummy with value one if the opponent is an enemy (i.e., opposing voter) and zero otherwise. Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Models (1), (3), (5), and (7) denoted by "Spite?" estimate the decision to behave spitefully or not with a logistic regression. Models (2), (4), (6), and (8) denoted by "Spite Score" estimate the decision on how spitefully to behave conditionally on behaving spitefully using a beta regression. Model (1)-(4) display the full sample while Models (5)-(6) and (7)-(8) display the spite behavior of Trump and Clinton voters, respectively. Standard errors are in parenthesis.

3.3. Attitudes



(a) Social distance and opponent.

(b) Feeling of warmth and opponent.

Figure 3: Attitudes and opponent.
The figure shows the attitudes towards opponents. The figure on the left shows how *close* participants report to feel towards their friends and enemies. The figure on the right shows how *warm* participants report to feel towards their friends and enemies. The left two columns show the attitudes towards friends (i.e., coinciding voters) while the two right columns indicate the attitudes towards enemies (i.e., opposing voters). Red bars denote the attitudes of Trump voters, while blue bars denote the attitudes of Clinton voters. Tie fighters denote 95% confidence intervals. P-values are obtained using Wilcoxon signed-rank tests for within comparisons and Wilcoxon-Mann-Whitney-Test for between comparisons.

To have a better understanding of the underlying reasoning for the observed behavior, we will now look at the reported attitudes. Thus, we will analyze the attitudes of Trump and Clinton voters towards coinciding and opposing voters. For that purpose, we first examine the results of the social distance questionnaire (as can be seen in Figure 3a). As the second measure of attitude, we compare how warm participants felt towards their enemies compared to their friends (as can be seen in Figure 3b) by using the feeling thermometer.[33] There are no substantial or systematic changes over time in attitudes or behavior,[34] which is why we pool the data and report on the results for each individual wave in the appendix C.2.

It can be seen that on average friends were considered much closer ($M = 5.04$, $SD = 0.97$) compared to enemies ($M = 2.8$, $SD = 1.24$), $t(2615) = 67.5$, $p \leq 0.001$.

[33] We discuss attitudes to democrats and republicans in general and compare it to other studies in Appendix A.2.

[34] See section C for a detailed discussion of wave effects.

Hence, we find that there is indeed a divide in the nation, as there is a significant and substantial gap (2.24 point difference on a 7 point scale, or a relative decrease in closeness of 45%) in the participants' attitudes towards fellow partisans and opposing voters. Concerning the feeling of warmth we see that on average friends are felt much warmer towards ($M = 6.7$, $SD = 2.28$) compared to enemies ($M = 1.52$, $SD = 1.83$), $t(2615)= 96.6$, $p \leq 0.001$. This again represents a significant and substantial difference (a 5.18 point difference on a 10 point scale, or a relative decrease in feeling of warmth of 77%)), and provides again evidence for a divided nation.[35]

Result A1 (Divided nation) Attitudes towards enemies are substantially worse compared to the attitudes towards friends.

Surprisingly, it can also be seen that the difference in social distance between friends and enemies was bigger for Clinton voters (M_{Diff}=2.59, SD_{Diff} =1.61) compared to Trump voters (M_{Diff}=1.81, SD_{Diff} =1.55), $t(2614)= -12.2$, $p \leq 0.001$.

In particular, this was due to Clinton voters' attitudes towards enemies. Clinton voters felt less close towards their enemies than did Trump voters.[36] At the same time Clinton and Trump voters did not differ significantly in their closeness towards friends.[37]

Similar results are obtained concerning the second measure of attitudes. Considering the feeling of warmth towards friends there is again no difference between Clinton voters and Trump voters.[38] However, we find again that Clinton voters show a stronger reaction towards their enemies than Trump voters in their second measure of attitudes.[39]

Taken together: Trump voters' closeness towards Clinton voters was only 66% of the closeness expressed towards fellow Trump voters (3.29 / 5.01). Similarly, Trump voters feeling of warmth towards Clinton voters was only 31% of the warmth expressed towards fellow Trump voters (2.07 / 6.67).

Clinton voters closeness towards Trump voters was, however, only 49% of the

[35] Similar results are obtained using linear regressions on the differences in attitudes and mixed-effects regressions on the repeated measures of attitudes. The results for the social distance measure are reported in Table 14. The result for the feeling of warmth measure is reported in Table 13.

[36] $M_{Clinton} = 2.49$, $SD_{Clinton} = 1.16$; $M_{Trump} = 3.29$, $SD_{Trump} = 1.22$, $t(2614)= 16.9$, $p \leq 0.001$.

[37] $M_{Clinton} = 5.06$, $SD_{Clinton} = 0.98$; $M_{Trump} = 5.01$, $SD_{Trump} = 0.95$, $t(2614)= -1.4$, $p \geq 0.05$.

[38] $M_{Clinton} = 6.72$, $SD_{Clinton} = 2.24$; $M_{Trump} = 6.67$, $SD_{Trump} = 2.34$, $t(2614)= -0.7$, $p \geq 0.05$.

[39] $M_{Clinton} = 1.18$, $SD_{Clinton} = 1.6$; $M_{Trump} = 2.07$, $SD_{Trump} = 2.03$, $t(2614)= 14.2$, $p \leq 0.001$.

closeness expressed towards fellow Clinton voters (2.49 / 5.06). A similar result is observed for the feeling of warmth: Clinton voters feeling of warmth towards Trump voters was only 18% of the warmth expressed towards fellow Clinton voters (1.18 / 6.72).

Thus, Clinton voters express almost 17%, 13% more antipathy toward Trump voters than the reverse in the measures of social distance and feeling of warmth, respectively.

Result A2 Clinton voters show more negative attitudes towards their enemies than Trump voters do.

Thus, in conclusion, we find the same pattern in the attitudes as we have observed in the spite behavior. Enemies are considered as more distant/less warm, and the participants behaved more spitefully towards their enemies. Interestingly: Clinton voters behave more spitefully towards their enemies compared to their friends while Trump voters are indifferent. The same can be found in attitudes: Clinton and Trump voters have the same attitudes towards friends but Clinton voters have significantly worse attitudes towards their enemies than Trump voters do.[40]

4. Discussion

This paper investigates whether tribalism—understood as the self-identified party affiliation—leads to dysfunctional behavior. In particular, we studied which attitudes and, more importantly, which behavior voters showed towards voters casting the same or the opposite vote. To the best of our knowledge, we are the first to study spite in a political context. For this purpose, we collected decisions of self-reported Clinton and Trump voters online shortly before the 58th US presidential election, after the inauguration of the president-elect, before the midterms 2018 and after the midterms.

Most importantly, we were able to show that dysfunctional behavior—understood as the destruction of wealth—was significantly more likely if an opposing voter was impacted compared to a coinciding voter. Peculiarly, we found that Clinton voters were significantly more likely to behave spitefully towards Trump voters compared towards fellow Clinton voters. This effect was not found for Trump voters—they do not differentiate between Trump voters and Clinton voters in their behavior.

[40] In Appendix D we further investigate how the attitudes mediate the decision to behave spitefully.

In addition, we were able to show that Trump voters were significantly less likely to behave spitefully towards enemies than Clinton voters were.

These effects are supported by the attitudes of the voters: attitudes towards opposing voters were substantially and significantly more negative than attitudes towards coinciding voters. Surprisingly, this effect was significantly stronger for Clinton voters. Further, the timing of the experiment did not substantially change attitudes, and it had no significant effect on dysfunctional behavior (discussed in detail in Appendix C).

While we believe the results to be robust, two possible limitations should be noted. First, our experiment might be prone to experimenter demand effects as the opponent's political orientation is made salient. While this saliency is essential for the treatment to work, it might reveal the experiment's purpose and, thus, lead participants to shift their behavior. However, we find in the results that Trump voters do not differentiate between Clinton voters and fellow Trump voters, which indicates that if a possible demand effect is at work, it is at least not straightforward. It might be the case that the demand effect affects Clinton voters and Trump voters differently. However, this seems rather unlikely as, to the best of our knowledge, no heterogeneous experimenter demand effects have been reported in the literature. And even if experimenter demand shifts Clinton and Trump voters' behavior differently, these effects are expected to be relatively small. For example, de Quidt et al. (2018) show that if the experimenter indicates his expectation, the demand effect is estimated at .2 standard errors and more strikingly, Mummolo and Peterson (2018) show that even financial incentives to respond inline with researcher expectations fail to consistently induce demand effect. Still, a demand effect cannot be excluded, and future research might use methods suggested by, e.g., de Quidt et al. (2018) to account for this limitation. The second limitation of our experiment is the non-representativeness of our sample. While our sample is much more representative of the US population than typical student samples, it is still not representative of the US population, as discussed in detail in Section 3.1. Thus, such a selection might bias our results and reduce the generalizability of our findings. In Appendix B.2, we try to deal with this issue by adjusting the weights of our estimations to make our sample artificially representative. While our results remain robust, we cannot exclude the possibility that our findings would differ using a representative sample.

Several aspects of the results are worth elaborating on.

First, it is worth pointing out that we are not the first to show that there is a

divided nation (see Pew Research Center, 2017a), but we are the first to show that this division leads to significantly increased destructive behavior. This is the main point of the paper, and this presents a significant and important contribution. We are able to show that even in a low key situation, like an online experiment, people are more likely to behave spitefully if matched with opposing voters. Hence, it seems plausible that in more salient situations where partisanship is even easier to detect and of more importance (e.g., collaborative work), the effect would be even stronger. More importantly, we know now that an increasing division of the United States leads to increased social and economic costs. This in itself presents the importance of a unifying leadership and the potential harm of a divisive presidency.

Second, it is interesting that the timing of the experiment hardly influenced attitudes and behavior. In particular, because Donald Trump won unexpectedly, we would have hypothesized that the attitudes and behavior would change substantially, but this seems not to be the case. A possible explanation is that the division in the United States has been increasing for years (Pew Research Center, 2017a) and that the presidency of Donald Trump was not as crucial to the division as largely portrayed in the media.

Third, the differences between Clinton voters and Trump voters are worth elaborating on. On the one hand, it seems not too surprising that Trump voters did not differentiate between friends and enemies because this would be perfectly in line with most papers on outgroup-bias who show that outgroup-bias lead rarely to purely hostile behavior (Tajfel, 1970; Brewer, 1999; Riek et al., 2006; Halevy et al., 2008; Weisel and Böhm, 2015; De Dreu et al., 2016). It is also not too surprising that attitudes towards opposing voters are negative because this has also been shown in other papers (Weisel and Böhm, 2015; Margolis and Sances, 2016; Fowler and Kam, 2007; Loewen, 2010; Carlin and Love, 2013; Iyengar and Westwood, 2015; Michelitch, 2015).

However, it is surprising that Clinton voters have significantly less positive attitudes towards their enemies compared to Trump voters. More importantly, Clinton voters behave more spitefully towards enemies than Trump voters do.

One possible explanation for the asymmetry in behavior is that Trump voters are considered inferior or even morally wrong in supporting Donald Trump. In that case, Mummendey and Wenzel (1999) argue theoretically that inferior groups are more likely to experience discrimination and hostility. Similarly, Brewer (1999) argues that outgroup hate might be present if participants are fighting for political power. Further support is provided by Parker and Janoff-Bulman (2013), who

show that morality based groups lead to less positive emotions. More importantly, Weisel and Böhm (2015) show a significant increase in help avoidance if the group difference is morality-based: "When given the chance to benefit a strong-enmity outgroup, and even more so a morality-based outgroup, many group members decline to do so" (Weisel and Böhm, 2015, p. 118).

In case Clinton supporters really see a morality based difference between them and Trump supporters, while Trump supporters do not see the difference as morality based, we might explain the significant difference in attitudes and behavior towards enemies.

The question now is whether Clinton voters really see themselves as morally superior to Trump supporters, while Trump supporters do not see a morality based difference, and if so, then what is driving this perception.

It might be that constant reports of Donald Trump's mishaps and especially presenting Trump supporters as racist, sexist, and/or homophobic might have manifested a selective picture of Trump supporters. The following quote from Hillary Clinton might represent the picture a Clinton supporter might have of a Trump supporter:

> "To just be grossly generalistic, you could put half of Trump's supporters into what I call the "basket of deplorables,"[...] The racist, sexist, homophobic, xenophobic, Islamaphobic, you name it. And unfortunately, there are people like that, and he [Donald Trump] has lifted them up."(Chozick, 2016)

Similarly, recent polls indicate that a majority of Democrats indicate to feel angry going into the midterm elections of 2018 while only 30 percent of Republicans say the same.[41] In Appendix F we also discuss whether and how morality differs between Clinton and Trump voters in our experiment.

However, these explanations are only stipulations, and future research has to provide empirical evidence and clarifications. It might be valuable for future research to take a closer look at the justifications and motivations of Clinton and Trump supporters for spite behavior.

Given the recent indications of a growing division in the US, it is essential to understand which further-reaching consequences it has. This question is particularly relevant with regard to destructive behavior as it might reduce our societal progress, threaten our democratic values, and potentially curtail economic growth.

[41] See https://www.politico.com/story/2018/11/05/poll-generic-ballot-narrows-on-eve-of-midterms-960757.

Overall, the central message of this paper is: A divided America means a bruised America. The division of the United States not only manifests itself in differences in attitudes but also results in increased destructive behavior. The goal of future research has to be to figure out how to combat such destructive behavior and how to mitigate the political division.

References

Abbink, K., Brandts, J., Herrmann, B., and Orzen, H. (2010). Intergroup conflict and intra-group punishment in an experimental contest game. *American Economic Review*, 100(1):420–47.

Abbink, K. and Dogan, G. (2019). How to choose your victim. *Games and Economic Behavior*, 113:482 – 496.

Abbink, K. and Herrmann, B. (2011). The moral costs of nastiness. *Economic Inquiry*, 49(2):631–633.

Abbink, K. and Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters*, 105(3):306 – 308.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Alabastro, A., Rast, D. E., Lac, A., Hogg, M. A., and Crano, W. D. (2013). Intergroup bias and perceived similarity: Effects of successes and failures on support for in- and outgroup political leaders. *Group Processes & Intergroup Relations*, 16(1):58–67.

Alcantara, C., Uhrmacher, K., and Guskin, E. (Oct. 16, 2016). Clinton and trump's demographic tug of war. *The Washington Post*. Available at https://www.washingtonpost.com/graphics/politics/2016-election/the-demographic-groups-fueling-the-election/.

Arechar, A. A., Gächter, S., and Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, 21(1):99–131.

Azrieli, Y., Chambers, C. P., and Healy., P. J. (2015). Incentives in experiments: A theoretical analysis. Mimeo.

Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

Bartling, B. and Netzer, N. (2016). An externality-robust auction: Theory and experimental evidence. *Games and Economic Behavior*, 97:186 – 204.

Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20(3):351.

Bernhard, H., Fehr, E., and Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *American Economic Review*, 96(2):217–221.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3):429–444.

Buhrmester, M., Kwang, T., and Gosling, S. D. (2011). Amazon's mechanical turk. *Perspectives on Psychological Science*, 6(1):3–5.

Camerer, C. F. and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1):7–42.

Carlin, R. E. and Love, G. J. (2013). The politics of interpersonal trust and reciprocity: An experimental approach. *Political Behavior*, 35(1):43–63.

Carpenter, J., Verhoogen, E., and Burks, S. (2005). The effect of stakes in distribution experiments. *Economics Letters*, 86(3):393 – 398.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352.

Chen, R. and Chen, Y. (2011). The potential of social identity for equilibrium selection. *The American Economic Review*, 101(6):2562–2589.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–98.

Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1):431–57.

Chen, Y., Li, S. X., Liu, T. X., and Shih, M. (2014). Which hat to wear? impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, 84:58 – 86.

Chozick, A. (Sept. 10, 2016). Hillary clinton calls many trump backers "deplorables," and g.o.p. pounces. *The New York Times*. Available at https://www.nytimes.com/2016/09/11/us/politics/hillary-clinton-basket-of-deplorables.html?_r=0.

Crandall, C. S. (1991). Multiple stigma and aids: Illness stigma and attitudes toward homosexuals and iv drug users in aids-related stigmatization. *Journal of Community & Applied Social Psychology*, 1:165–172.

Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, 34(2):1–24.

De Dreu, C. K. W. (2010). Social value orientation moderates ingroup love but not outgroup hate in competitive intergroup conflict. *Group Processes & Intergroup Relations*, 13(6):701–713.

De Dreu, C. K. W., Gross, J., Méder, Z., Giffin, M., Prochazkova, E., Krikeb, J., and Columbus, S. (2016). In-group defense, out-group aggression, and coordination failures in intergroup conflict. *Proceedings of the National Academy of Sciences*, 113(38):10524–10529.

de Quidt, J., Haushofer, J., and Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, 108(11):3266–3302.

Elgers, P. T., Haltiner, J. R., and Hawthorne, W. H. (1979). Beta regression tendencies: Statistical and real causes. *The Journal of Finance*, 34(1):261–263.

Federal Bureau of Investigation, United States Department of Justice (2016). Uniform crime reporting program data: County-level detailed arrest and offense data, 2014. Available at http://doi.org/10.3886/ICPSR36399.v2.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Fehr, E., Hoff, K., and Kshetramade, M. (2008). Spite and development. *American Economic Review*, 98(2):494–99.

Filippin, A. and Guala, F. (2013). Costless discrimination and unequal achievements in an experimental tournament. *Experimental Economics*, 16(3):285–305.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–56.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Flaxman, S., Goel, S., and Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1):298–320.

Forsythe, R., Horowitz, J. L., Savin, N., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3):347 – 369.

Fowler, J. H. and Kam, C. D. (2007). Beyond the self: Social identity, altruism, and political participation. *Journal of Politics*, 69(3):813–827.

Glassman, B. and United States Census Bureau (2016). Selected economic characteristics by state: 2014 and 2015. Available at https://www.census.gov/data/tables/2016/demo/income-poverty/glassman-acs.html.

Greene, S. (1999). Understanding party identification: A social identity approach. *Political Psychology*, 20(2):393–403.

Greenwald, A. G. and Pettigrew, T. F. (2014). With malice toward none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, 69(7):669–684.

Grün, B., Kosmidis, I., and Zeileis, A. (2012). Extended beta regression in R: Shaken, stirred, mixed, and partitioned. *Journal of Statistical Software*, 48(11):1–25.

Halevy, N., Bornstein, G., and Sagiv, L. (2008). "in-group love" and "out-group hate" as motives for individual participation in intergroup conflict. *Psychological Science*, 19(4):405–411.

Hare, C. and Poole, K. T. (2014). The polarization of contemporary american politics. *Polity*, 46(3):411–429.

Herrmann, B. and Thöni, C. (2009). Measuring conditional cooperation: a replication study in Russia. *Experimental Economics*, 12(1):87–92.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Hewstone, M., Rubin, M., and Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53(1):575–604.

Horton, J. J., Rand, D. G., and Zeckhauser, R. J. (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, 14(3):399 – 425.

Iyengar, S. and Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3):690–707.

Jacobson, J. J., Maxson, B. B., Mays, K., and Kowalski, C. J. (1992). A utility analysis of dental implants. *The International journal of oral & maxillofacial implants*, 7(3):381 – 388.

Jönsson, M. T. and Thor, G. (2012). Estimating coextinction risks from epidemic tree death: Affiliate lichen communities among diseased host tree populations of fraxinus excelsior. *PLOS ONE*, 7(9):1–10.

Jordan, J. J., Hoffman, M., Bloom, P., and Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, 530(7591):473–476.

Jordan, J. J., Sommers, R., Bloom, P., and Rand, D. G. (2017). Why Do We Hate Hypocrites? Evidence for a Theory of False Signaling. *Psycholical Science*.

Kaid, L. L., Leland, C. M., and Whitney, S. (1992). The impact of televised political ads: Evoking viewer responses in the 1988 presidential campaign. *Southern Communication Journal*, 57(4):285–295.

Kimbrough, E. O. and Reiss, J. P. (2012). Measuring the distribution of spitefulness. *PLOS ONE*, 7(8):1–8.

Kirchkamp, O. and Mill, W. (2019). Spite vs. risk: Explaining overbidding. a theoretical and experimental investigation. CESifo Working Paper No. 7631.

Kirk, A. and Patrick, S. (Nov. 7, 2016). Us election: How age, race and education are deciding factors in the race for president. *The Telegraph*. Available at http://www.telegraph.co.uk/news/0/us-election-how-age-race-and-education-are-deciding-factors-in-t/.

Kite, A. (November 7, 2016). Election 2016: A violent season marked with arrests. *The Washington Post*. Available at https://www.wsj.com/articles/election-2016-a-violent-season-marked-with-arrests-1478552509.

Kocher, M., Cherry, T., Kroll, S., Netzer, R. J., and Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters*, 101(3):175–178.

Kranton, R., Pease, M., Sanders, S., and Huette, S. (2018). Groupy and not groupy behavior: Deconstructing bias in social preferences. Technical report. Mimeo.

Kranton, R. E. and Sanders, S. G. (2017). Groupy versus non-groupy social preferences: Personality, region, and political party. *American Economic Review*, 107(5):65–69.

Laliberté, E., Adair, E. C., and Hobbie, S. E. (2012). Estimating litter decomposition rate in single-pool models using nonlinear beta regression. *PLOS ONE*, 7(9):1–16.

Lange, P. A. M. V., Bruin, E. M. N. D., Otten, W., and Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73(4):733–746.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1(3):593 – 622.

Loewen, P. J. (2010). Affinity, antipathy and political participation: How our concern for others makes us vote. *Canadian Journal of Political Science*, 43(3):661–687.

Manning, N. (November 15, 2016). We stand for transparency, justice and human rights: Monitoring post election violence and protest in the usa. *Ushahidi*. Available at https://www.ushahidi.com/index.php?p=blog/2016/11/15/championing-justice-transparency-and-human-rights-monitoring-post-election-violence-and-protest-in-usa.

Mao, A., Dworkin, L., Suri, S., and Watts, D. J. (2017). Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoner's dilemma. *Nature Communications*, 8.

Marcus, D. K., Zeigler-Hill, V., Mercer, S. H., and Norris, A. L. (2014). The psychology of spite and the measurement of spitefulness. *Psychological Assessment*, 26(2):563–574.

Margolis, M. F. and Sances, M. W. (2016). Partisan differences in nonpartisan activity: The case of charitable giving. *Political Behavior*, pages 1–26.

Mason, W. and Suri, S. (2012). Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods*, 44(1):1–23.

Michelitch, K. (2015). Does electoral competition exacerbate interethnic or interpartisan economic discrimination? evidence from a field experiment in market price bargaining. *American Political Science Review*, 109(1):43–61.

Mill, W. and Morgan, J. (2018). Competition between friends and foes. Mimeo.

Miller, A. H. and Wlezien, C. (1993). The social group dynamics of partisan evaluations. *Electoral studies*, 12(1):5–22.

Mummendey, A. and Schreiber, H.-J. (1984). 'different' just means 'better': Some obvious and some hidden pathways to in-group favouritism. *British Journal of Social Psychology*, 23(4):363–367.

Mummendey, A. and Wenzel, M. (1999). Social discrimination and tolerance in intergroup relations: Reactions to intergroup difference. *Personality and Social Psychology Review*, 3(2):158–174.

Mummolo, J. and Peterson, E. (2018). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*, 113(2):517–529.

Murphy, R. O., Ackerman, K. A., and Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8):771–781.

Ospina, R. and Ferrari, S. L. (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis*, 56(6):1609 – 1623.

Paolacci, G., Chandler, J., and Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419.

Parker, M. T. and Janoff-Bulman, R. (2013). Lessons from morality-based social identity: The power of outgroup "hate", not just ingroup "love". *Social Justice Research*, 26(1):81–96.

Parrillo, V. N. and Donoghue, C. (2005). Updating the bogardus social distance studies: a new national survey. *The Social Science Journal*, 42(2):257 – 271.

Patrick, D. L., Starks, H. E., Cain, K. C., Uhlmann, R. F., and Pearlman, R. A. (1994). Measuring preferences for health states worse than death. *Medical Decision Making*, 14(1):9–18.

Pew Research Center (2014). Pew research center 2014 u.s. religious landscape study. Available at http://www.pewforum.org/datasets/pew-research-center-2014-u-s-religious-landscape-study/.

Pew Research Center (November, 2016). Behind trump's victory: Divisions by race, gender, education. *Mimeo*. Available at https://www.pewresearch.org/fact-tank/2016/11/09/behind-trumps-victory-divisions-by-race-gender-education/.

Pew Research Center (Oktober, 2017a). The partisan divide on political values grows even wider. *Mimeo*. Available at http://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/.

Pew Research Center (September, 2017b). Partisans differ widely in views of police officers, college professors. *Mimeo*. Available at http://www.people-press.org/2017/09/13/partisans-differ-widely-in-views-of-police-officers-college-professors/.

Peysakhovich, A., Nowak, M. A., and Rand, D. G. (2014). Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, 5:4939.

Prediger, S., Vollan, B., and Herrmann, B. (2014). Resource scarcity and antisocial behavior. *Journal of Public Economics*, 119(C):1–9.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., and Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5:3677.

Reilly, K. (November 13, 2016). Racist incidents are up since donald trump's election. these are just a few of them. *Time*. Available at http://time.com/4569129/racist-anti-semitic-incidents-donald-trump/.

Riek, B. M., Mania, E. W., and Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, 10(4):336–353.

Robinson, J. P., Shaver, P. R., and Wrightsman, L. S., editors (1999). *Measures of Political Attitudes. Measures of Social Psychological Attitudes*, volume 2. Academic Press, San Diego, CA, US, 1 edition.

Rogers, J. A., Polhamus, D., Gillespie, W. R., Ito, K., Romero, K., Qiu, R., Stephenson, D., Gastonguay, M. R., and Corrigan, B. (2012). Combining patient-level and summary-level data for alzheimer's disease modeling and simulation: a beta regression meta-analysis. *Journal of Pharmacokinetics and Pharmacodynamics*, 39(5):479–498.

Sadrieh, A. and Schröder, M. (2017). Acts of helping and harming. *Economics Letters*, 153(Supplement C):77 – 79.

Saijo, T. and Nakamura, H. (1995). The "spite" dilemma in voluntary contribution mechanism experiments. *Journal of Conflict Resolution*, 39(3):535–560.

Schmid, M., Wickler, F., Maloney, K. O., Mitchell, R., Fenske, N., and Mayr, A. (2013). Boosted beta regression. *PLOS ONE*, 8(4):1–15.

Schünemann, H. J., Griffith, L., Jaeschke, R., Goldstein, R., Stubbing, D., and Guyatt, G. H. (2003). Evaluation of the minimal important difference for the feeling thermometer and the st. george's respiratory questionnaire in patients

with chronic airflow obstruction. *Journal of Clinical Epidemiology*, 56(12):1170 – 1176.

Seow, W. J., Pesatori, A. C., Dimont, E., Farmer, P. B., Albetti, B., Ettinger, A. S., Bollati, V., Bolognesi, C., Roggieri, P., Panev, T. I., Georgieva, T., Merlo, D. F., Bertazzi, P. A., and Baccarelli, A. A. (2012). Urinary benzene biomarkers and dna methylation in bulgarian petrochemical workers: Study findings and comparison of linear and beta regression models. *PLOS ONE*, 7(12):1–8.

Shallwani, P., King, K., and Thadani, T. (November 13, 2016). Protests against president-elect donald trump continue across the u.s. *The Wall Street Journal*. Available at https://www.wsj.com/articles/donald-trump-protests-continue-in-u-s-person-injured-in-portland-1478955985.

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol Methods*, 11(1):54–71.

Suri, S. and Watts, D. J. (2011). Cooperation and contagion in web-based, networked public goods experiments. *PLOS ONE*, 6(3):1–18.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5):96–102.

United States Department of Agriculture (2017). Poverty estimates for the u.s., states, and counties, 2015. Available at https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data.aspx.

Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559.

Vignoles, V. L. and Moncaster, N. J. (2007). Identity motives and in-group favouritism: A new approach to individual differences in intergroup discrimination. *British Journal of Social Psychology*, 46(1):91–113.

Weisberg, H. F. (1980). A multidimensional conceptualization of party identification. *Political Behavior*, 2(1):33–60.

Weisel, O. (2015). Negative and positive externalities in intergroup conflict: exposure to the opportunity to help the outgroup reduces the inclination to harm it. *Frontiers in Psychology*, 6:1594.

Weisel, O. and Böhm, R. (2015). "ingroup love" and "outgroup hate" in intergroup conflict between natural groups. *Journal of Experimental Social Psychology*, 60:110 – 120.

Weisel, O. and Zultan, R. (2016). Social motives in intergroup conflict: Group identity and perceived target of threat. *European Economic Review*, 90:122 – 133.

Yan, H., Sgueglia, K., and Walker, K. (December 22, 2016). "make america white again": Hate speech and crimes post-election. *CNN*. Available at http://edition.cnn.com/2016/11/10/us/post-election-hate-crimes-and-fears-trnd/.

Zizzo and Oswald (2001). Are people willing to pay to reduce others' incomes? *Annales d'Économie et de Statistique*, (63/64):39.

A. Robustness checks

A.1. Coherence

To ensure that participants are not randomly choosing a candidate and are really paying attention, we asked participants also for their preferred political party, and we included several attention checks.

A.1.1. Consistency

More specifically, we asked participants in the general demographics part “With which party do you normally identify yourself most with?” and later in the study, we asked “Which political party do you usually feel closest to”.

99.5 % of those participants who indicated to identify most with Democrats also felt closest to Democrats, and similarly, 99.1 % of those participants who indicated to identify most with Republicans also felt closest to Republicans.

Furthermore, 91.5 % of those participants who indicated to have voted for Hillary Clinton also felt closest to Democrats, and similarly, 85.1 % of those participants who indicated to have voted for Donald Trump also felt closest to the Republicans.

Additionally, 77.5 % of those participants who indicated to have voted for Hillary Clinton usually are identifying themselves with Democrats (6.1 % usually identify themselves with Independents), and similarly 69.4 % of those participants who indicated to have voted for Donald Trump usually identify themselves with Republicans (24.4 % usually identify themselves with Independents).

A.1.2. Attention Checks

In some of the questionnaires, we included additional attention checks by asking questions, for example, "Click on agree" or "This is another control. We ask you to select the second option." We included four of those attention checks (without having any impact on the participants). Only 2 % of all participants failed one or more of those attention checks (some of the participants, however, reported to have misunderstood the meaning of "second option" as this might have been ambiguous in regard to the reference point).

Overall, participants seem to be attentive and consistent in their political attitudes in our study.

A.2. Feeling of warmth

In this section, we examine the properties of the feeling of warmth measure compared to results from the literature.

Notably, we asked participants for their feeling of warmth towards Trump and Clinton voters and additionally towards Republicans and Democrats. A recent study by the Pew Research Center also asked Republicans and Democrats for their feeling of warmth towards, among others, Republicans and Democrats. They also found a very similar gap. What is, however, striking is how close the estimations from our study are to the estimates of the study by the Pew Research Center (2017b).

Table 4 shows the estimates in our online study compared to the representative study by the Pew Research Center (2017b).

Even though our results are very similar to the results by the study of the Pew Research Center (2017b), our participants indicated to have, on average, lower feelings of warmth than participants in Pew Research Center (2017b). This might be partly explained by the fact that our experiment was conducted shortly before or after the election, and at that point, people were somewhat fed up with politics. Another explanation might be that we used a 10-point scale, and Pew Research Center (2017b) used a 100-point scale. Of course, it is also possible that our participants are particularly negative. However, the key observation from this comparison is that reported feelings of warmth in our online study are very similar to the representative study by the Pew Research Center (2017b).

Table 4: Comparing feelings of warmth towards Democrats and Republicans in our study to the study by Pew Research Center (2017b).

	Attidues towards	Our study	Pew study
Among Democrats	Democracts	6.52 ± 0.11	7.10
	Democracts	1.49 ± 0.08	2.40
Among Republicans	Republicans	6.5 ± 0.15	7.00
	Republicans	2.29 ± 0.13	2.30

A.3. The spite score

In this section, we examine the properties of the spite score.

For that purpose, we investigate how the spite score, and the spite questionnaire, feeling of closeness and feeling of warmth relate and whether the Vote interacts

with those measures. For that purpose we estimate Equation 1 with $x_{i1} = 1$, $x_{i2} =$ Trump voter, $x_{i3} = Ind.Var$ and $x_{i4} = Ind.Var \cdot$ Trump voter for each of the measures ($Ind.Var \in$ {Spite Questionnaire, Feeling of closeness, Feeling of Warmth}). Here Trump voter is a dummy variable with 1 if the participant voted for Donald Trump and 0 if she voted for Hillary Clinton. The result of these estimations can be found in Table 5. It can be seen that log odds of being spiteful at all and also the log odds of behaving maximally spitefully increased with increasing scores in the spite questionnaire. Comparably, it can be seen that the log odds of behaving spitefully and also the log odds of behaving maximally spitefully decreased with the distance and the warmth participants felt towards their opponent. Thus, the nicer the attitude towards participants are the less spitefully is the behavior.

It is also interesting to see that none of the measures interact significantly at the 5% level with the indicated vote. Hence, an increase in e.g. feeling of warmth has the same effect for Trump voters as for Clinton voters. Thus, Clinton and Trump voters show the same reaction towards their opponents for a given level of attitudes.

Taken together, the results suggest that the spite score exhibits valuable properties such that spite increases with antisocial answers in a questionnaire and decreases with increasing sympathy for the opponent.

Table 5: Attitudes as predictors of spite behavior.

	Spite? Spite Questionnaire	Spite Score Questionnaire	Spite? Social Distance	Spite Score Social Distance	Spite? Feeling Thermometer	Spite Score Feeling Thermometer
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.22^{***}	-0.05	0.13	0.57^{***}	-0.11	0.41^{***}
	(0.14)	(0.13)	(0.12)	(0.10)	(0.07)	(0.06)
Ind.Var	0.38^{***}	0.16^{***}	-0.13^{***}	-0.06^{*}	-0.06^{***}	-0.02
	(0.06)	(0.05)	(0.03)	(0.03)	(0.01)	(0.01)
Trump voter	0.04	0.01	0.10	-0.54^{**}	0.13	-0.35^{***}
	(0.23)	(0.19)	(0.23)	(0.19)	(0.12)	(0.10)
Ind.Var x Trump voter	0.09	-0.11	0.05	$0.08^{\cdot}$	0.04	0.02
	(0.09)	(0.07)	(0.06)	(0.05)	(0.02)	(0.02)
Model	Logistic	Beta	Logistic	Beta	Logistic	Beta
Observations	2,616	1,156	2,616	1,156	2,616	1,156
Log Likelihood	$-1,742.56$	77.51	$-1,780.25$	74.89	$-1,780.66$	72.86

Note: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;

Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Ind.Var. indicates the independent variable with is either the spite questionnaire, the feeling of closeness or the feeling of warmth. Models (1), (3), and (5) estimate the decision to behave spitefully or not with a logistic regression. Models (2), (4), and (6) estimate the decision on how spitefully to behave conditionally on behaving spitefully using a beta regression. Models (1) and (2) estimate the spite behavior using the spite questionnaire as the independent variable. Models (3) and (4) estimate the spite behavior using the social distance questionnaire as the independent variable. Models (5) and (6) estimate the spite behavior using the feeling thermometer as the independent variable. Standard errors are in parenthesis.

A.4. Estimations with Controls

To ensure robustness of our results, we extend in this section the estimation of section 3.2 to further controls.

At first we also estimate the Equation 1 with $x_{i1} = 1$, $x_{i2} =$ Enemy (the main model); and $x_{i1} = 1$, $x_{i2} =$ Enemy, $x_{i3} =$ Trump voter and $x_{i4} = \textit{Opponent} \cdot$ Trump voter, where Trump voter is a dummy variable with 1 if the participant voted for Donald Trump and 0 if she voted for Hillary Clinton. Similarly Enemy is a dummy variable with 1 if the opponent is an enemy and 0 if the opponent is a friend.

Additionally, we control for socio demographics and relevant information of the participants county/state. In the first three sets of models we include a control

of socio demographics of: gender $\in$ { Male; Female }, Education $\in$ {College-ed. or higher; below College }, Age $\in \{18, \ldots, 88\}$, and reported Income $\in \{< 70k, > 70k\}$ with the control vector X_m, with $m = 5, ..., 8$. The result can be found in Table 6.

We also control for information about poverty, crime, and religion of the participants' county. In the same table, we additionally include socio-demographics.

The crime data was obtained from the "Uniform Crime Reporting Program Data: County-Level Detailed Arrest and Offense Data, 2014" reported by the United States Department of Justice, Federal Bureau of Investigation[42]. To control for crime in the county of the participating participant, we include the relative crime (amount of reported crimes relative to the population of the county) and the relative violent crime (amount of reported violent crimes relative to the population of the county).

The data on poverty on the county level was obtained from the United States Department of Agriculture[43] and we also obtained data on the state level from the United States Census Bureau[44]. To control for poverty on the state level, we included the poverty as reported in (United States Department of Agriculture, 2017). We also controlled for the percentage of people living below the poverty level on the state level from the Census Data.

The data on how religious the county was, was obtained from the Pew Research Center 2014 U.S. Religious Landscape Study.[45] We controlled in two ways for how religious the county of the participant was. We included the percentage of participants answering yes on the question, whether they believe in god, and we secondly included the percentage of participants answering "very important" on the question: "How important is religion in your life".

Table 6 replicates Table 3 from section 3.2.3 and additionally controls for socio demographics and other controls.

The resulting estimates show that the results are essentially robust to controls.

[42]The data can be found here: http://doi.org/10.3886/ICPSR36399.v2 (Federal Bureau of Investigation, United States Department of Justice, 2016)

[43]The data can be found here: https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/ (United States Department of Agriculture, 2017). The poverty estimates reported in this data are model estimates from the U.S. Census Bureau's Small Area Income and Poverty Estimate.

[44]The data can be found here: https://www.census.gov/data/tables/2016/demo/income-poverty/glassman-acs.html Glassman and United States Census Bureau (2016).

[45]The data can be found here: http://www.pewforum.org/datasets/pew-research-center-2014-u-s-religious-landscape-study/ (Pew Research Center, 2014).

	Spite? Full sample		Score		Spite? Trump voters		Score		Spite? Clinton voters		Score	
Constant	−0.59***	(0.08)	0.29***	(0.07)	−0.07	(0.10)	0.15*	(0.08)	−0.59***	(0.08)	0.27***	(0.07)
Enemy	0.47***	(0.10)	0.09	(0.09)	0.01	(0.13)	−0.11	(0.10)	0.47***	(0.10)	0.08	(0.09)
Trump voter	0.52***	(0.12)	−0.15	(0.11)								
Enemy x Trump voter	−0.46**	(0.16)	−0.19	(0.14)								
Age												
EduLow												
IncomeLow												
Male												
PovertyCounty												
PovertyState												
Viol.Crimes												
Crimes												
Religion												
BeliefInGod												
Model	Logistic		Beta		Logistic		Beta		Logistic		Beta	
LogLik	-1779.69		72.56		-689.06		2.33		-1090.62		82.56	
Observations	2,616		1,156		995		480		1,621		676	

	Spite? Full sample		Score		Spite? Trump voters		Score		Spite? Clinton voters		Score	
Constant	−1.05***	(0.16)	0.09	(0.14)	−0.31	(0.25)	−0.12	(0.19)	−1.19***	(0.20)	0.11	(0.18)
Enemy	0.47***	(0.10)	0.09	(0.09)	−0.003	(0.13)	−0.13	(0.10)	0.48***	(0.10)	0.09	(0.09)
Trump voter	0.54***	(0.12)	−0.10	(0.11)								
Enemy x Trump voter	−0.48**	(0.16)	−0.22	(0.14)								
Age	0.01·	(0.003)	0.001	(0.003)	0.002	(0.01)	0.01·	(0.004)	0.01·	(0.004)	−0.002	(0.004)
EduLow	−0.12	(0.09)	−0.16*	(0.08)	−0.01	(0.14)	−0.19·	(0.11)	−0.18	(0.12)	−0.13	(0.11)
IncomeLow	0.14·	(0.08)	0.11	(0.07)	0.15	(0.13)	−0.01	(0.10)	0.13	(0.10)	0.19*	(0.09)
Male	0.34***	(0.08)	0.21**	(0.07)	0.15	(0.13)	0.13	(0.10)	0.45***	(0.10)	0.25**	(0.10)
PovertyCounty												
PovertyState												
Viol.Crimes												
Crimes												
Religion												
BeliefInGod												
Model	Logistic		Beta		Logistic		Beta		Logistic		Beta	
LogLik	-1767.5		79.17		-687.51		5.72		-1077.66		88.36	
Observations	2,616		1,156		995		480		1,621		676	

	Spite? Full sample		Score		Spite? Trump voters		Score		Spite? Clinton voters		Score	
Constant	−2.01	(1.73)	1.10	(1.45)	−2.96	(2.97)	2.83	(2.29)	−0.90	(2.16)	0.01	(1.86)
Enemy	0.50***	(0.11)	0.07	(0.09)	0.02	(0.13)	−0.07	(0.11)	0.50***	(0.11)	0.07	(0.10)
Trump voter	0.50***	(0.13)	−0.19·	(0.11)								
Enemy x Trump voter	−0.47**	(0.17)	−0.14	(0.15)								
Age												
EduLow												
IncomeLow												
Male												
PovertyCounty	0.03**	(0.01)	−0.01	(0.01)	0.06***	(0.02)	0.003	(0.01)	0.01	(0.01)	−0.02*	(0.01)
PovertyState	0.02	(0.02)	−0.05*	(0.02)	−0.04	(0.04)	−0.03	(0.03)	0.06*	(0.03)	−0.06*	(0.03)
Viol.Crimes	−72.85	(60.03)	−66.24	(52.94)	−230.34*	(103.31)	100.37	(84.66)	11.42	(75.60)	−122.63·	(67.75)
Crimes	−7.04	(18.22)	18.29	(15.74)	20.53	(28.96)	3.41	(23.25)	−21.61	(23.78)	17.86	(21.17)
Religion	−1.49	(1.26)	1.97·	(1.07)	−3.13	(2.05)	2.21	(1.61)	−0.11	(1.62)	1.62	(1.43)
BeliefInGod	1.85	(2.55)	−1.11	(2.13)	5.09	(4.30)	−4.20	(3.29)	−0.64	(3.21)	0.72	(2.78)
Model	Logistic		Beta		Logistic		Beta		Logistic		Beta	
LogLik	-1667.66		76.73		-640.52		5.6		-1019.47		88.84	
Observations	2,462		1,088		940		449		1,522		639	

	Spite? Full sample		Score		Spite? Trump voters		Score		Spite? Clinton voters		Score	
Constant	−2.62	(1.75)	0.55	(1.45)	−3.17	(2.99)	2.34	(2.29)	−1.80	(2.19)	−0.60	(1.87)
Enemy	0.51***	(0.11)	0.07	(0.09)	0.02	(0.13)	−0.08	(0.11)	0.51***	(0.11)	0.07	(0.10)
Trump voter	0.52***	(0.13)	−0.12	(0.11)								
Enemy x Trump voter	−0.49**	(0.17)	−0.18	(0.15)								
Age	0.01·	(0.004)	0.001	(0.003)	0.003	(0.01)	0.01*	(0.004)	0.01*	(0.005)	−0.004	(0.004)
EduLow	−0.10	(0.10)	−0.22**	(0.08)	0.05	(0.14)	−0.24*	(0.12)	−0.17	(0.13)	−0.20·	(0.11)
IncomeLow	0.11	(0.08)	0.10	(0.07)	0.11	(0.14)	−0.05	(0.11)	0.11	(0.11)	0.17·	(0.10)
Male	0.35***	(0.08)	0.24**	(0.07)	0.15	(0.14)	0.19·	(0.11)	0.49***	(0.11)	0.28**	(0.10)
PovertyCounty	0.03**	(0.01)	−0.01	(0.01)	0.06***	(0.02)	0.01	(0.01)	0.01	(0.01)	−0.02*	(0.01)
PovertyState	0.01	(0.02)	−0.05*	(0.02)	−0.05	(0.04)	−0.03	(0.03)	0.06·	(0.03)	−0.06·	(0.03)
Viol.Crimes	−66.87	(60.15)	−64.77	(52.78)	−225.88*	(102.90)	117.88	(84.49)	17.04	(76.45)	−128.11·	(67.56)
Crimes	−9.88	(18.35)	16.95	(15.70)	19.97	(29.03)	0.68	(23.19)	−25.54	(24.22)	18.69	(21.14)
Religion	−1.68	(1.27)	1.71	(1.07)	−3.11	(2.06)	2.08	(1.60)	−0.42	(1.64)	1.28	(1.43)
BeliefInGod	2.20	(2.56)	−0.51	(2.13)	5.15	(4.31)	−3.88	(3.28)	−0.18	(3.25)	1.52	(2.78)
Model	Logistic		Beta		Logistic		Beta		Logistic		Beta	
LogLik	-1655.52		85.4		-639.14		11.16		-1005.47		96	
Observations	2,462		1,088		940		449		1,522		639	

Notes: ·p<0.10;*p<0.05;**p<0.01;***p<0.001;

Table 6: Zero-inflated beta regression of spite using socio-demographics and further controls.

Enemy denotes a dummy with value one if the opponent is an enemy (i.e., opposing voter). Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. EduLow denotes participants who have at most a high school degree. IncomeLow denotes participants who have an income below $70k$ a year. PovertyCounty denotes the poverty level reported in (United States Department of Agriculture, 2017) on the county level. PovertyState denotes the poverty level reported in (Glassman and United States Census Bureau, 2016) on the state level. Viol.Crimes(County)/Crimes(County) denotes the percentage of (violent) crimes relative to the population of a county reported in (Federal Bureau of Investigation, United States Department of Justice, 2016). Religion denotes the percentage of people in a county indicating that religion is very important in their lives. BeliefInGod indicates the percentage of people in a county stating that they believe in God. Models denoted by "Spite?" estimate the decision to behave spitefully or not with a logistic regression. Models denoted by "Score" estimate the decision on how spitefully to behave conditionally on behaving spitefully using a beta regression. Standard errors are in parenthesis.

A.5. Model comparisons conditional spite

In this section, we compare further potential models to explain the spite score against the beta regression. As the spite score has an excessive amount of zeros, it seems plausible to assume a two-step process while the first step is the decision to be spiteful or not. This decision is a binary decision, and hence, a logistic regression was used for this step. However, for the conditional spite score, we argued for a beta regression as this is the common method for this kind of data. To further support our claim, we compare modeling the spite score conditional on behaving spitefully with a beta regression compared to other regressions. It can be seen from Table 7 that the beta regression outperforms the other regressions (standard Gaussian-, Gamma-, Weibul-, log-normal-, exponential and exponential Gaussian regression) on the conditional spite score.[46]

Table 7: Model fit of alternative models.

Comparison	DF	χ^2_{diff}	P-value
Beta (LogLik=72.559) vs. Exp. (LogLik=-291.375)	1.00	727.87	$\leq$0.001***
Beta (LogLik=72.559) vs. Log-Normal (LogLik=-31.065)	0.00	207.25	$\leq$0.001***
Beta (LogLik=72.559) vs. Gaus (LogLik=-19.549)	0.00	184.22	$\leq$0.001***
Beta (LogLik=72.559) vs. Exp.Gaus (LogLik=20.53)	-1.00	104.06	$\leq$0.001***
Beta (LogLik=72.559) vs. Gamma (LogLik=37.988)	0.00	69.14	$\leq$0.001***
Beta (LogLik=72.559) vs. Weibul (LogLik=49.93)	0.00	45.26	$\leq$0.001***

This table compares the Log-likelihoods of different models estimating the conditional spite score. Beta denotes a beta regression. Exp. denotes an exponential regression. Log-Normal denotes a log-normal regression. Gaus denotes a standard Gaussian regression. Exp.Gaus denotes an exponential gaussian regression. Gamma denotes a gamma regression and Weibul denotes a weibul regression. χ^2_{diff} denotes the Chi-squared test statistic of the differences.

B. Alternative estimations

In this section, we will estimate the same models as in section 3.2.3. However, we will use rather ordinary tools to estimate the spite score. Hence, in this section, we will not use the zero-inflated beta regression, and therefore, we will ignore that the spite score is between 0 and 1.

If we do not account neither for the bounded outcome of the spite score nor for the zero inflation we could simply use a linear regression. The outcome is reported

[46] It is also noteworthy that the main results of this paper do not change, if one would use a different kind of model.

in Table 8 in Models (1)-(4). It shows the same estimation as Table 3 of section 3.2.3.

However, as it seems very reasonable to account for the extensive amount of zeros in the spite score, we use also a standard Tobit regression. Table 8 shows in models (5)-(8) the same estimation as Table 3 of section 3.2.3.

It can be seen from the results, that all main results still prevail using the ordinary estimation approach. In particular, we see, on average, more spiteful behavior towards enemies than towards friends. Also, as before, we find that most of the spite behavior is driven by Clinton voters while Trump voters do not behave significantly more spitefully towards enemies.

Table 8: Estimating spite with ordinary methods.

	Linear regressions				**Tobit regressions**			
	Full sample		Trump voters	Clinton voters	Full sample		Trump voters	Clinton voters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.19^{***}	0.17^{***}	0.22^{***}	0.17^{***}	-0.10^{***}	-0.15^{***}	0.02	-0.18^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)
Enemy	0.04^{**}	0.06^{***}	-0.01	0.06^{***}	0.09^{***}	0.15^{***}	-0.01	0.15^{***}
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.04)	(0.03)
Trump voter		0.05^{**}				0.15^{***}		
		(0.02)				(0.04)		
Trump voter x Enemy		-0.07^{**}				-0.15^{**}		
		(0.02)				(0.05)		
Model	OLS	OLS	OLS	OLS	Tobit	Tobit	Tobit	Tobit
Observations	2616	2616	995	1621	2616	2616	995	1621
LogLik	-438.29	-433.07	-139.99	-292.12	-1914.36	-1906.84	-715.99	-1186.4

Notes: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;

Enemy denotes a dummy with value one if the opponent is an enemy (i.e. opposing voter). Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Models (1), (2), and (5),(6) report on the full sample while Models (3),(7), and (4),(8) report on Trump and Clinton voters, respectively. Models (1)-(4) show the estimations obtained using a linear regression. Estimations in Models (5)-(8) are performed using a Tobit regression. Standard errors are in parenthesis.

B.1. Propensity score matching

Even though participants were assigned an opponent randomly, participants did select whether to be a Trump or a Clinton voter. Throughout most of the paper, we only compare the behavior towards an opponent within a group (either Trump or Clinton voters), but in Table 3, we also compare whether Clinton and Trump voters differ in their spite behavior.

To control for self-selection we use nearest neighbor propensity score matching to estimate Tables 3, and 8. Moreover, we reestimate Tables 14, and 13. More specifically, we match participants on the following demographic characteristics: Age, gender, education, ethnicity, and income. Table 9 shows the estimation using propensity score matching. It is evident from Table 9 that all results are qualitatively not influenced by participants self-selecting into Clinton and Trump voters.

Table 9: Reestimating results using propensity score matching.

	Spite								Thermo		Distance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	−0.10***	−0.16***	0.19***	0.16***	−0.40***	−0.60***	0.20***	0.25***	6.90***	6.91***	5.07***	5.10***
	(0.02)	(0.03)	(0.01)	(0.01)	(0.06)	(0.08)	(0.05)	(0.07)	(0.04)	(0.05)	(0.02)	(0.03)
Enemy	0.09***	0.16***	0.04***	0.07***	0.32***	0.51***	0.005	0.06	−5.64***	−6.03***	−2.24***	−2.57***
	(0.02)	(0.03)	(0.01)	(0.01)	(0.08)	(0.10)	(0.07)	(0.09)	(0.05)	(0.07)	(0.03)	(0.04)
Trump voter		0.15***		0.06***		0.53***		−0.11		−0.06		−0.10*
		(0.04)		(0.02)		(0.12)		(0.11)		(0.08)		(0.04)
Enemy x Trump voter		−0.16***		−0.07**		−0.50**		−0.16		1.02***		0.85***
		(0.05)		(0.02)		(0.16)		(0.14)		(0.11)		(0.06)
Model	Tobit	Tobit	OLS	OLS	Logistic	Logistic	Beta	Beta	Mixed effects	Mixed effects	Mixed effects	Mixed effects
LogLik	-1904.95	-1896.55	-830.41	-824.32	-1894.17	-1884.35	50.39	55.6	-11831.62	-11762.08	-8749.93	-8602.76
AIC	3815.89	3803.09	1666.83	1658.63	3792.33	3776.7	-94.79	-101.21	23671.24	23536.15	17507.86	17217.51
Observations			2,616	2,616	2,616	2,616	1,156	1,156	5,232	5,232	5,232	5,232

Notes: *p<0.05;**p<0.01;***p<0.001;
Enemy denotes a dummy with value one if the opponent is an enemy (i.e., opponent voted not the same candidate on election day) and zero otherwise. Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. The first two columns reestimate the spite behavior reported in Table 8 using a Tobit regression. Models (3) and (4) reestimate the spite behavior reported in Table 8 using a linear regression. Models (5) and (6) reestimate the spite behavior reported in Table 3 using a logistic regression on the decision to be spiteful or not. Models (7) and (8) reestimate the spite behavior reported in Table 3 using a beta regression on the conditional spite behavior. Models (9), (10) and (11), (12) reestimate the feeling of warmth reported in Table 14 and the social distance reported in Table 13 using mixed-effects models. Standard errors are in parenthesis. All reestimations are performed using new weights obtained from nearest neighbor propensity score matching on the following demographics: Age, gender, education, ethnicity, and income.

B.2. Weighted sample

As pointed out in section 3.1, our sample is not representative. Even though the assignment of the opponent is random, and all results can be interpreted causally, the result might not be representative of the US population. Thus, we reweight our sample to make the results more representative. Specifically, we change the

weights of our sample by using the R `anesrake` method, which performs raking to obtain the weights such that our demographic variables coincide with the target demographic of the US population. The target demographic variables are: age, gender, ethnicity, and educational achievement.[47]

Using the obtained weights we estimate spite behavior reported in Tables 3, and 8. Moreover, we reestimate the attitudes reported in Tables 14, and 13. Table 10 shows the estimation. It is evident from Table 10 that all results are qualitatively not influenced by our sample not being representative.

Table 10: Reestimating results using population weights.

	Spite		Spite		Spite		Spite		Thermo		Distance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	−0.13***	−0.22***	0.17***	0.13***	−0.43***	−0.76***	0.13*	−0.01	6.74***	6.56***	4.92***	4.84***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.05)	(0.07)	(0.05)	(0.06)	(0.05)	(0.06)	(0.02)	(0.03)
Enemy	0.08***	0.15***	0.04***	0.07***	0.21**	0.37***	0.03	0.25**	−5.22***	−5.24***	−1.94***	−2.07***
	(0.02)	(0.03)	(0.01)	(0.01)	(0.08)	(0.10)	(0.07)	(0.09)	(0.05)	(0.06)	(0.02)	(0.03)
Trump voter		0.35***		0.15***		1.18***		0.33**		0.46***		0.15**
		(0.04)		(0.02)		(0.13)		(0.10)		(0.09)		(0.05)
Enemy x Trump voter		−0.24***		−0.11***		−0.67***		−0.56***		0.08		0.44***
		(0.05)		(0.02)		(0.18)		(0.15)		(0.10)		(0.05)
Model	Tobit	Tobit	OLS	OLS	Logistic	Logistic	Beta	Beta	Mixed effects	Mixed effects	Mixed effects	Mixed effects
LogLik	-1853.4	-1802.99	-1769.7	-1728.75	-1571.04	-1512.61	42.05	49.66	-13375.26	-13358.1	-10046.6	-9977.32
AIC	3712.81	3615.97	3545.4	3467.5	3146.09	3033.23	-78.11	-89.32	26758.53	26728.2	20101.2	19966.64
Observations			2,616	2,616	2,616	2,616	1,156	1,156	5,232	5,232	5,232	5,232

Notes: *p<0.05;**p<0.01;***p<0.001;

Enemy denotes a dummy with value one if the opponent is an enemy (i.e., opponent voted not the same candidate on election day) and zero otherwise. Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. The first two columns reestimate the spite behavior reported in Table 8 using a Tobit regression. Models (3) and (4) reestimate the spite behavior reported in Table 8 using a linear regression. Models (5) and (6) reestimate the spite behavior reported in Table 3 using a logistic regression on the decision to be spiteful or not. Models (7) and (8) reestimate the spite behavior reported in Table 3 using a beta regression on the conditional spite behavior. Models (9), (10) and (11), (12) reestimate the feeling of warmth reported in Table 14 and the social distance reported in Table 13 using mixed-effects models. Standard errors are in parenthesis. All reestimations are performed using weights to make our sample more representative. The weights are obtained by using the R `anesrake` method, which performs raking to obtain the weights such that our demographic variables coincide with the target demographic of the US population (i.e., age, gender, ethnicity, and educational achievement).

C. Wave effects

As mentioned in the design section, we conducted the experiment in four waves: in late November 2016 (before the 58th US presidential election), late January 2017 (after the inauguration), late October 2018 (before the midterms) and early

[47] Age and gender demographics can be found here: https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf. Information on educational achievement can be found here: https://www.census.gov/data/tables/2018/demo/education-attainment/cps-detailed-tables.html. The distribution of ethnicity can be found here: https://www.census.gov/quickfacts/fact/table/US/PST045218.

November 2018 (after the midterms). As there are no relevant changes in our data over time, we pooled the data in the main paper. In this section, however, we direct our attention on the changes and differences in spite behavior and attitudes between the waves. Figure 4 shows the changes in spite behavior and attitudes between the waves. From Figure 4, we can already see that there are some changes over time, which, however, do not follow a clear pattern, nor are the changes substantial and really meaningful. Figure 5 shows the differences between two subsequent waves in spite behavior and attitudes. From Figure 5, we also see that for spite, no change is significantly greater than zero. A similar result is obtained for the social distance with the exception that Clinton voters changed to slightly more positive attitudes towards Trump voters from before the 58th US presidential election to after the inauguration – which however, represents a relative change of just 8.86%. For the feeling of warmth, we do indeed see some significant changes over time: from before the 58th US presidential election to after the inauguration Trump voters consider Clinton voters more positive (which reverses again from the inauguration to before the midterms) and similarly Clinton voters consider fellow Clinton voters more positive – which also represents a relative change of just 4.86%. It should also be taken into account that we compare multiple values against zero, with only three being significant. Taking any p-value adjustment method into account would render all significant differences void. Thus, we can see that for each subgroup, there are no substantial or relevant changes happening.

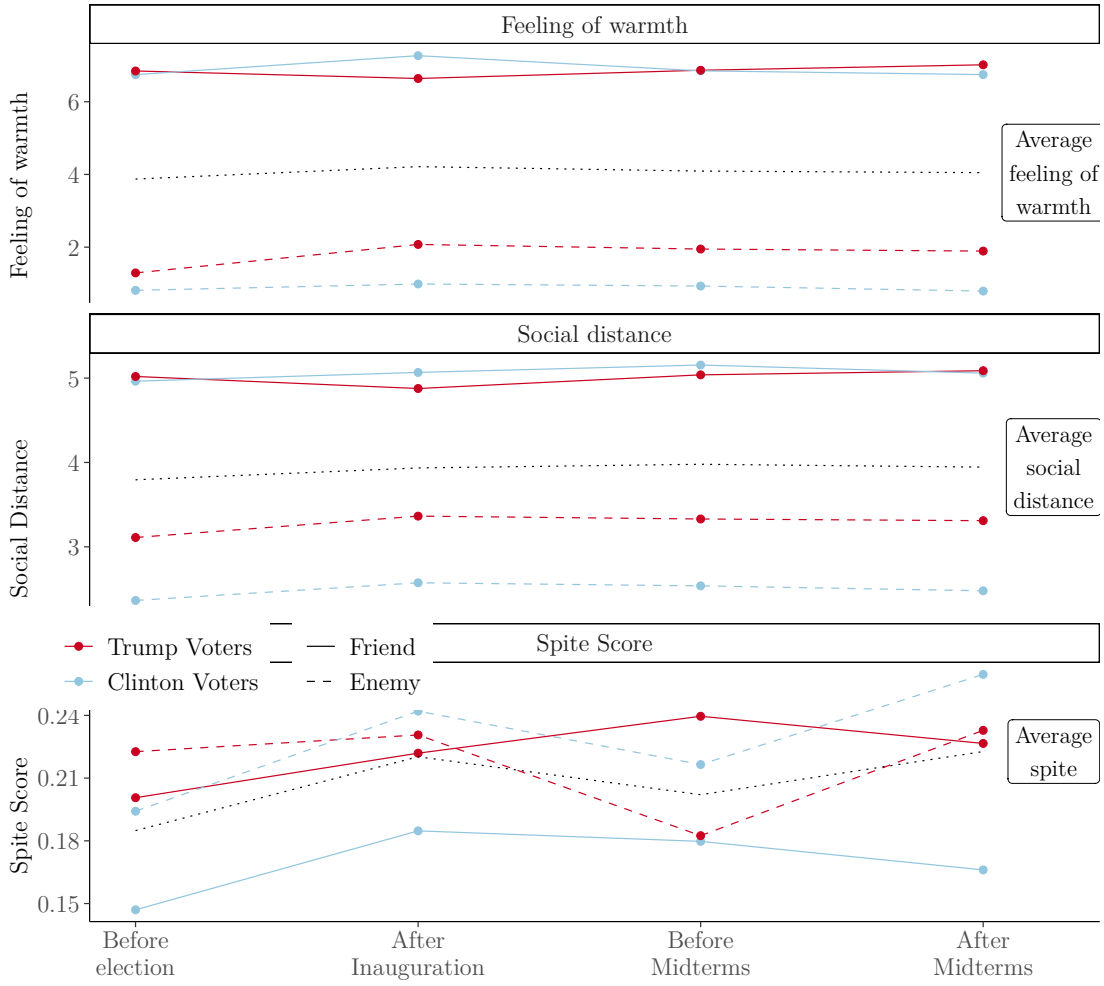


Figure 4: Attitudes and spite over time

The figure depicts the spite behavior and attitudes over time (i.e., for each of the four waves). The panel on top represents the reported feeling of warmth. The panel in the middle represents the reported social distance. The panel on the bottom depicts the spite behavior. Black dotted lines show the average behavior/attitudes for each measure. Blue lines denote the behavior and attitudes of Clinton voters, while red lines represent the behavior and attitudes of Trump voters. Solid lines depict the behavior and attitudes towards friends (i.e., coinciding voter) while dashed lines depict the behavior and attitudes towards enemies (i.e., opposing voter). The four waves were conducted: in late November 2016 (before the 58th US presidential election), late January 2017 (after the inauguration), late October 2018 (before the midterms), and early November 2018 (after the midterms).

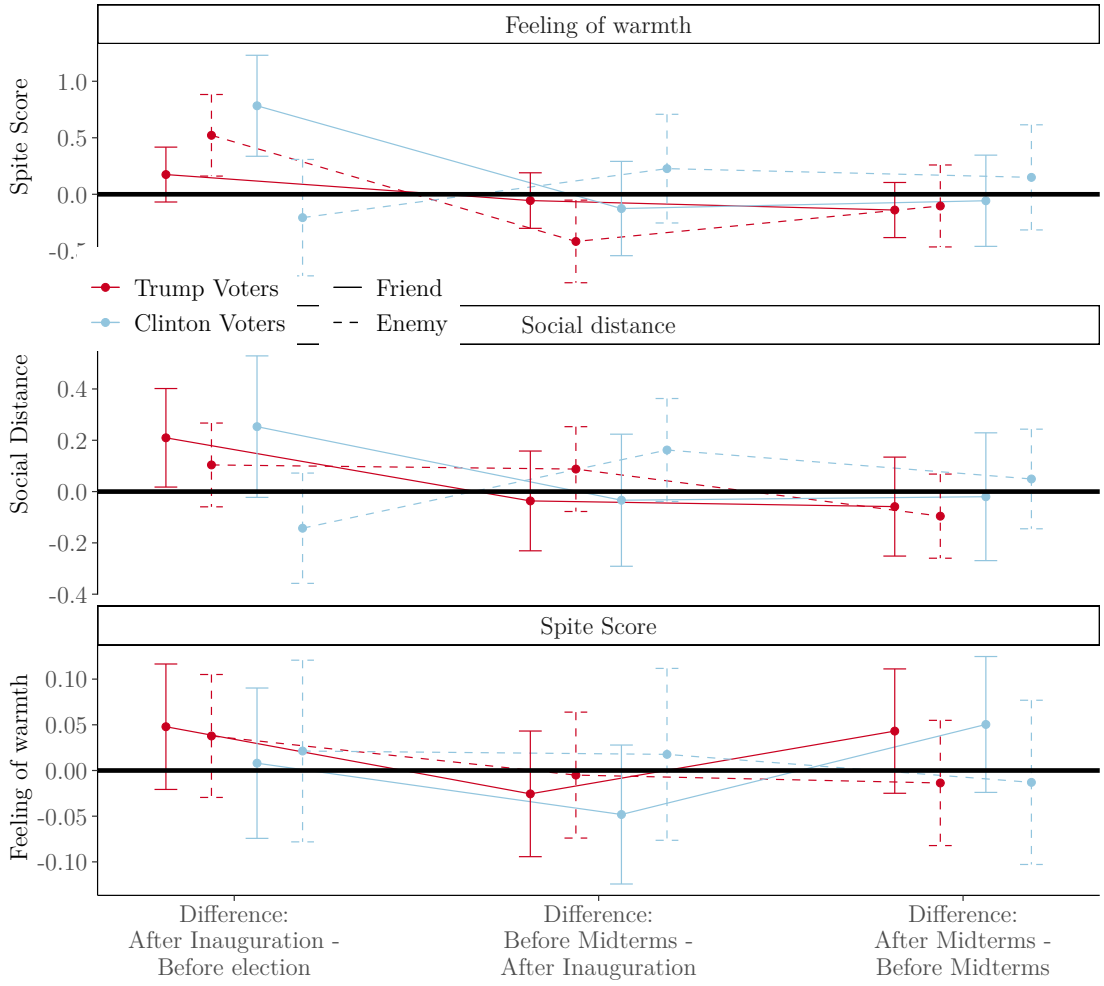


Figure 5: Differences in attitudes and spite over time
The figure depicts the difference in spite behavior and attitudes between two subsequent waves over time. The panel on the bottom represents the reported feeling of warmth. The panel in the middle represents the reported social distance. The panel on top depicts the spite behavior. Black dotted lines show the average behavior/attitudes for each measure. Blue lines denote the behavior and attitudes of Clinton voters, while red lines represent the behavior and attitudes of Trump voters. Solid lines depict the behavior and attitudes towards friends (i.e., coinciding voter) while dashed lines depict the behavior and attitudes towards enemies (i.e., opposing voter). The differences between waves are: comparing the change from after the inauguration (late January 2017) to before the 58th US presidential election (late November 2016); comparing the change from before the midterms (late October 2018) to after the inauguration (late January 2017) and comparing the change from after the midterms (early November 2018) to before the midterms (late October 2018).

Even though neither spite nor attitudes change systematically over time, we ac-

count for the timing effects in the following subsections were we focus on each measure separately.

C.1. Time and Spite

In this subsection, we discuss how our results change if we take the timing into account and if we focus on the results for each wave individually.

Table 11 reports the estimation of a zero-inflated beta regression following Equation 1 while controlling for the wave specific effects. In addition, Table 12 reports the estimation of a linear model and a Tobit model of spite while controlling for the wave specific effects. We can see that all results prevail if accounting for wave specific effects.

Table 11: Zero-inflated beta regression model of spite behavior.

	Spite?	Spite Score	Spite?	Spite Score	Spite?	Spite Score	Spite?	Spite Score
	Full sample		Full sample		Trump voters		Clinton voters	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-0.39^{***}	0.31^{***}	-0.58^{***}	0.38^{***}	-0.06	$0.22^{\cdot}$	-0.59^{***}	0.37^{***}
	(0.09)	(0.08)	(0.10)	(0.09)	(0.14)	(0.11)	(0.11)	(0.11)
Enemy	0.29^{***}	0.01	0.46^{***}	0.08	0.01	-0.12	0.46^{***}	0.07
	(0.08)	(0.07)	(0.10)	(0.09)	(0.13)	(0.10)	(0.10)	(0.09)
Trump voter			0.50^{***}	-0.16				
			(0.12)	(0.11)				
Enemy x Trump voter			-0.45^{**}	-0.19				
			(0.16)	(0.14)				
Model	Logistic	Beta	Logistic	Beta	Logistic	Beta	Logistic	Beta
Wave fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,616	1,156	2,616	1,156	995	480	1,621	676
Log Likelihood	−1,784.84	66.50	−1,776.50	74.51	−688.61	3.36	−1,086.75	84.09

Notes: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;
Enemy denotes a dummy with value one if the opponent is an enemy (i.e. opposing voter). After election denotes a dummy with value one if the decision was made after the inauguration of Donald Trump and zero if the decision was made before the election of Donald Trump. Models (1), (3), and (5) estimate the decision to behave spitefully or not with a logistic regression. Models (2), (4), and (6) estimate the decision on how spitefully to behave conditionally on behaving spitefully using a beta regression. All estimations account for wave specific effects. The omitted category is after the inauguration. Standard errors are in parenthesis.

Table 12: Ordinary regressions of the spite score

	Linear regressions				**Tobit regressions**			
	Full sample		Trump voters	Clinton voters	Full sample		Trump voters	Clinton voters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.20^{***}	0.18^{***}	0.23^{***}	0.18^{***}	-0.08^{**}	-0.13^{***}	0.03	-0.16^{***}
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)
Enemy	0.04^{**}	0.06^{***}	-0.005	0.06^{***}	0.09^{***}	0.15^{***}	-0.005	0.15^{***}
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.04)	(0.03)
Trump voter		0.05^{**}				0.14^{***}		
		(0.02)				(0.04)		
Trump voter x Enemy		-0.07^{**}				-0.15^{**}		
		(0.02)				(0.05)		
Model	OLS	OLS	OLS	OLS	Tobit	Tobit	Tobit	Tobit
Wave fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2616	2616	995	1621	2616	2616	995	1621
LogLik	-434.84	-429.93	-139.37	-289.11	-1910.6	-1903.6	-715.39	-1182.95

Notes: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;

Enemy denotes a dummy with value one if the opponent is an enemy (i.e. opposing voter). Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Models (1), (2), and (5),(6) report on the full sample while Models (3),(7), and (4),(8) report on Trump and Clinton voters, respectively. Models (1)-(4) show the estimations obtained using a linear regression. Estimations in Models (5)-(8) are performed using a Tobit regression. All estimations account for wave specific effects. The omitted category is after the inauguration. Standard errors are in parenthesis.

Moreover, Figure 6 depicts the spite behavior within each wave. It is evident that the patterns reported in section 3.2 can also be seen within each wave. Trump voters are seemingly indifferent between Clinton and Trump voters (with a small tendency towards more spite behavior towards Clinton voters, with the exception of the third wave); Clinton voters are less spiteful towards friends than are Trump voters and we see that Clinton voters significantly differentiate between friends and enemies.

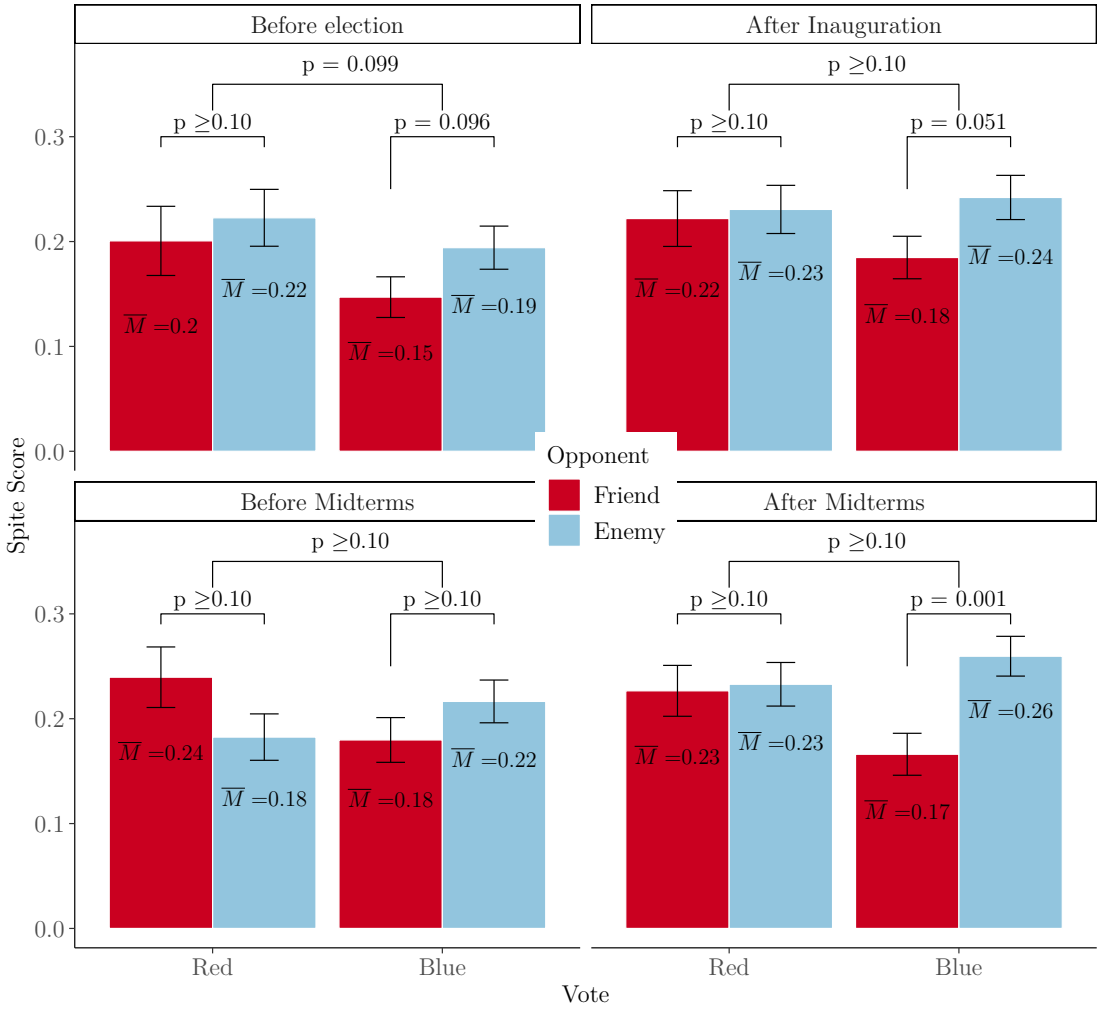


Figure 6: Spite towards opponents
The figure depicts how spiteful participants behave towards their friends (i.e., coinciding voter) and enemies (i.e., opposing voter). The left two bars show the spite behavior of Trump voters while the right two bars show the spite behavior of Clinton voters. Red bars denote the spite behavior towards friends (i.e., coinciding voters) while blue bars denote the spite behavior towards enemies (i.e., opposing voters). Tie fighters depict standard errors. P-values are calculated using t-tests. The upper left panel shows attitudes before the election. The upper right panel shows attitudes after the inauguration. The lower left and right panel depict the attitudes before and after the midterms, respectively.

C.2. Time and Attitdues

In this subsection, we focus on the attitudes participants reported to have in each of the four waves.

C.2.1. Time and feeling of warmth

Concerning the feeling of warmth, Table 13 reports the attitudes indicated by participants accounting for wave specific effects. We can see that accounting for wave specific effects does not change the results substantially and that all main results prevail. In addition, Figure 7 depicts the feeling of warmth in each wave. We can also see that, as reported before, typically Clinton and Trump voters do not differ in their attitudes towards friends; that Clinton and Trump voters have substantially and significantly less positive attitudes towards their opposing voters and that Clinton voters are also substantially and significantly more negative towards their enemies than Trump voters are. We can see that the attitudes do vary slightly between waves, which, however, does not change the overall pattern nor any results.

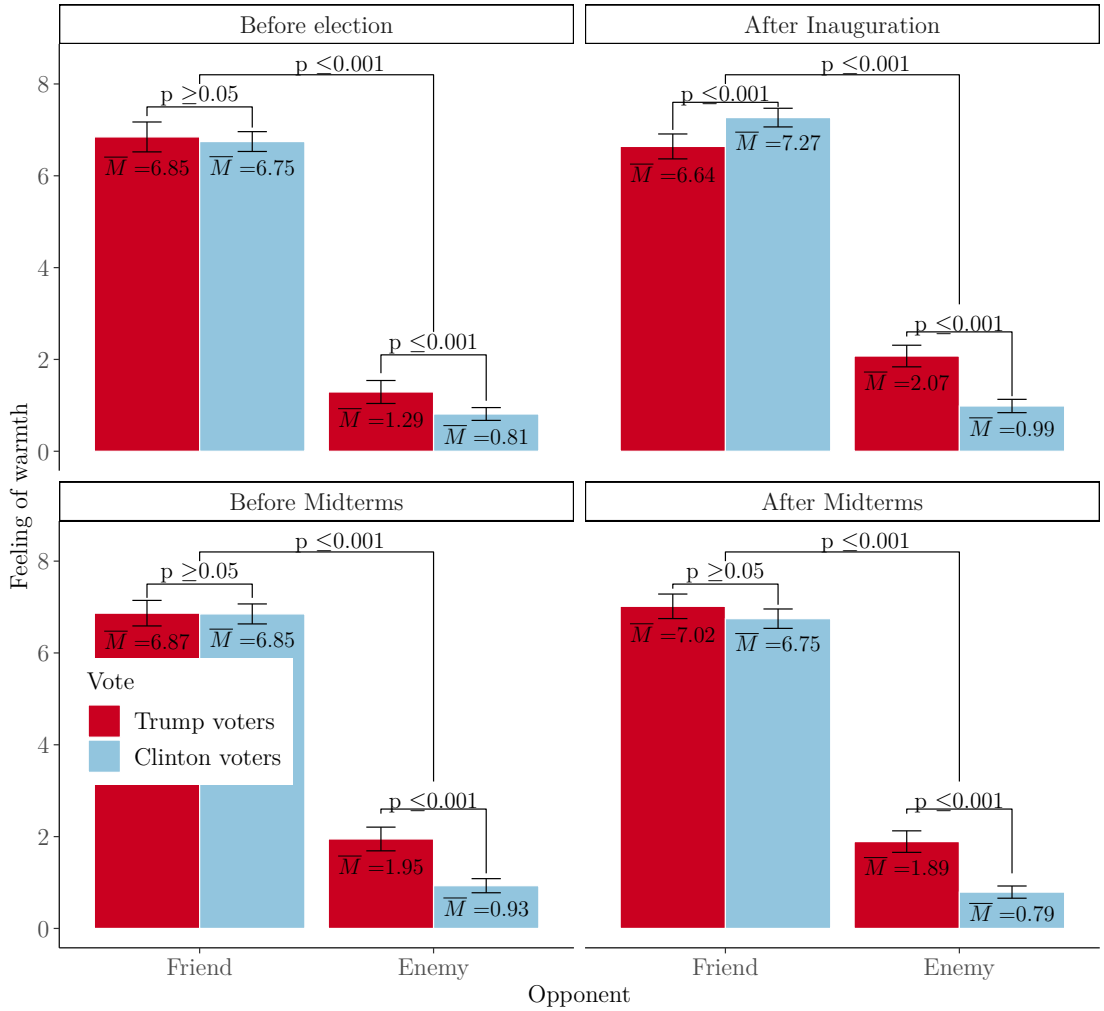


Figure 7: Feeling of warmth and opponent.
The figure depicts how warm participants report to feel towards their friends (i.e., coinciding voter) and enemies (i.e., opposing voter). The left two columns show how warm participants felt towards their friends (i.e., coinciding voters) while the two right columns indicate how warm participants felt towards their enemies (i.e., opposing voters). Red bars denote the attitudes of Trump voters, while blue bars denote the attitudes of Clinton voters. Tie fighters denote 95% confidence intervals. P-values obtained using Wilcoxon signed-rank tests for within comparisons and Wilcoxon-Mann-Whitney-Test for between comparisons. The upper left panel shows attitudes before the election. The upper right panel shows attitudes after the inauguration. The lower left and right panel depict the attitudes before and after the midterms, respectively.

Table 13: Regression of feeling of warmth.

	Difference in feelings towards friend and enemy		Feeling of warmth	
	(1)	(2)	(3)	(4)
Constant	6.06^{***} (0.07)	6.03^{***} (0.12)	0.88^{***} (0.05)	1.03^{***} (0.07)
Enemy			6.03^{***} (0.07)	6.03^{***} (0.07)
Trump voter	-0.98^{***} (0.11)	-0.97^{***} (0.12)	0.97^{***} (0.08)	0.96^{***} (0.08)
Trump voter x Enemy			-1.03^{***} (0.11)	-1.03^{***} (0.11)
Model	OLS	OLS	Mixed Effects	Mixed Effects
Wave fixed effects	×	✓	×	✓
Observations	2,616	2,616	5,232	5,232

Notes: $^{\cdot}$p<0.10;*p<0.05;**p<0.01;***p<0.001;
Enemy denotes a dummy with value zero if the opponent is a friend and one if the opponent is an enemy. Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Models (1) and (2) regress (using OLS) on the difference in the feeling of warmth towards friends and enemies. Models (3) and (4) regresses simply on the feeling of warmth towards the opponent using a mixed-effects model. Models (2) and (4) account for wave specific effects. The omitted category is after the inauguration. Standard errors are in parenthesis.

C.2.2. Time and Social distance

Concerning the social distance Table, 14 reports the attitudes participants indicated to have accounting for wave specific effects. We can see that accounting for wave specific effects does not change the results substantially and that all main results prevail. In addition, Figure 8 depicts the reported social distance in each wave. We can also see that, as reported before, typically Clinton and Trump voters do not differ in their attitudes towards friends; that Clinton and Trump voters have substantially and significantly less positive attitudes towards their opposing voters and that Clinton voters are also substantially and significantly more negative towards their enemies than Trump voters are. We can see that the attitudes do vary slightly between waves, which, however, does not change the overall pattern nor any results.

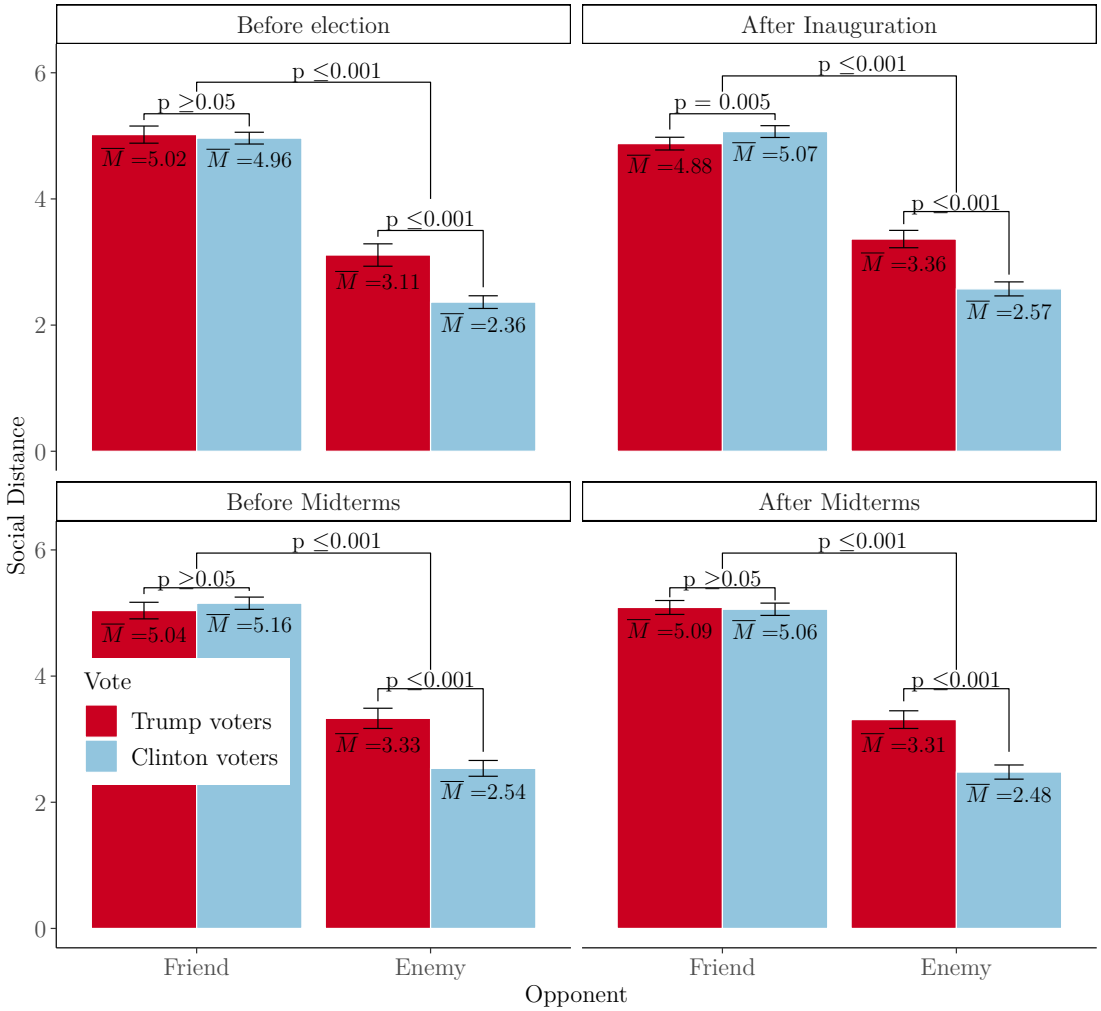


Figure 8: Social distance and opponent.
The figure depicts how close participants report to feel towards their friends (i.e., coinciding voter) and enemies (i.e., opposing voter). The left two columns show how close participants felt towards their friends (i.e., coinciding voters) while the two right columns indicate how close participants felt towards their enemies (i.e., opposing voters). Red bars denote the attitudes of Trump voters, while blue bars denote the attitudes of Clinton voters. Tie fighters denote 95% confidence intervals. P-values obtained using Wilcoxon signed-rank tests for within comparisons and Wilcoxon-Mann-Whitney-Test for between comparisons. The upper left panel shows attitudes before the election. The upper right panel shows attitudes after the inauguration. The lower left and right panel depict the attitudes before and after the midterms, respectively.

Table 14: Regression of the social distance.

	Difference in social distance between friend and enemy		Social distance	
	(1)	(2)	(3)	(4)
Constant	2.59*** (0.04)	2.47*** (0.06)	2.49*** (0.03)	2.51*** (0.04)
Enemy			2.57*** (0.04)	2.57*** (0.04)
Trump voter	−0.78*** (0.06)	−0.78*** (0.06)	0.80*** (0.04)	0.80*** (0.04)
Trump voter x Enemy			−0.86*** (0.06)	−0.86*** (0.06)
Model	OLS	OLS	Mixed Effects	Mixed Effects
Wave fixed effects	×	✓	×	✓
Observations	2,616	2,616	5,232	5,232

Notes: ·p<0.10;*p<0.05;**p<0.01;***p<0.001;
Enemy denotes a dummy with value zero if the opponent is a friend and one if the opponent is an enemy. Trump voter denotes a dummy with value one if the deciding participant is a Trump voter and zero if the deciding participant is a Clinton voter. Models (1) and (2) regress (using OLS) on the difference in the social distance towards friends and enemies. Models (3) and (4) regresses simply on the social distance towards the opponent using a mixed-effects model. Models (2) and (4) account for wave specific effects. The omitted category is after the inauguration. Standard errors are in parenthesis.

D. Path analysis

We have seen that participants' attitudes towards their opponent depend on whether the other is a friend or an enemy (result A1). We also have seen that Clinton voters, exhibit stronger emotions towards their enemies (result A2). Additionally, Clinton voters show more spite behavior towards enemies than do Trump voters (result B1). Hence, it seems plausible that attitudes might mediate the decision to be spiteful.

To estimate the indirect relationship of the opponent for Trump voters and Clinton voters (see Figure 10 for the path-diagram) on the spite behavior via the attitude towards the opponent, we use a structural equation model (SEM), a common approach, especially in psychology. However, this approach is not very common in political science and economics, and hence, we discuss the results and the method only in the appendix.

First, we will compare several structural equation models (SEMs) and use the best to derive insight about the indirect effect of partisanship on behavior.

D.1. Model comparisons SEM

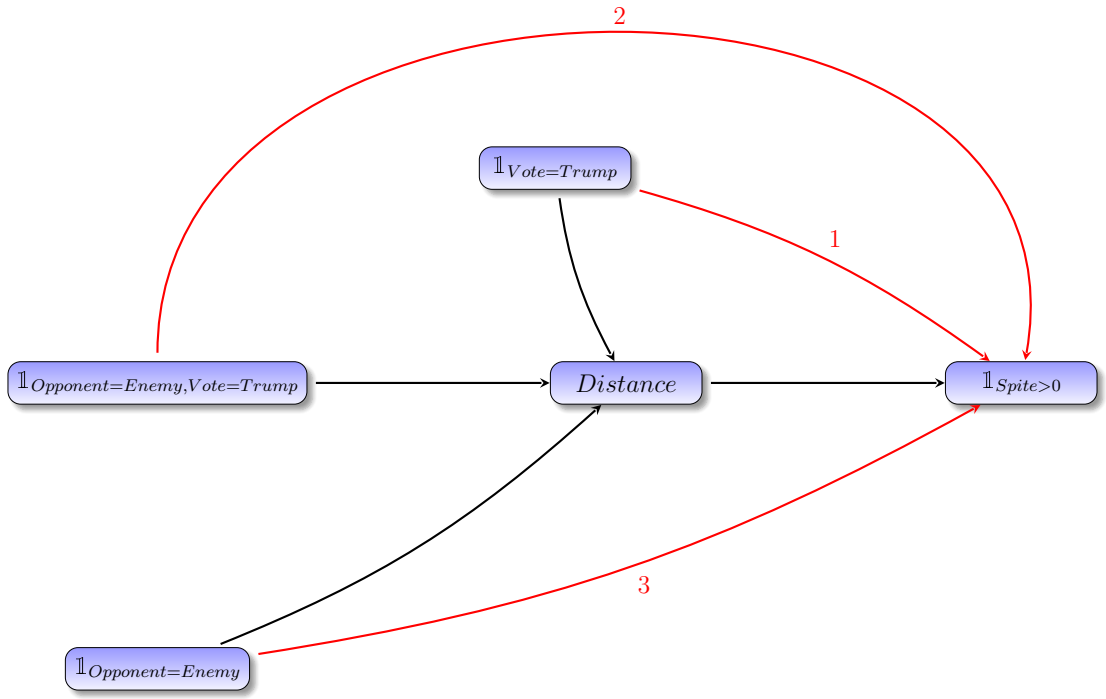


Figure 9: Possible structural equation models on the effect of the binary decision to behave spitefully or not, via social distance.

In this section, we investigate which structural model performs best in describing the effect on spite via social distance and also via feeling of warmth. As we were not able to find a canned version in R of a SEM with a zero-inflated beta regression, we used only the binary decision to be spiteful. A SEM with a standard OLS of the spite score leads to substantially identical results.

For that purpose, we compare the basic structural equation shown in Figure 9 as the model with black edges with extended models, including further direct effects. Red edges are extensions of the basic model, including, for example, the edge 1 denotes the direct effect of vote on spite. Table 15 shows the relevant model comparisons for the SEM via social distance and Table 16 shows the relevant model comparisons for the SEM via feeling of warmth. It can be seen that the best model is the full model (i.e., the basic model plus all the direct effect on spite) for a structural equation model via social distance. The best model for a structural equation model via feeling of warmth is the basic model plus the direct effects of vote and the interaction on spite.

Hence, the structural equation model we use has a direct and indirect effect of

vote on the decision to behave spitefully at all and in addition, it has an indirect (and direct, for the social distance model) effect of the opponent and a direct and indirect effect of the interaction of opponent and vote on the decision to behave spitefully. The resulting model can be seen in Figure 10 for the social distance and Figure 11 for the feeling of warmth model.

Table 15: Comparing the possible structural equation models on the effect of spite via social distance.

Comparison	DF	Chisq diff	P-value
Model:Basic + 1 vs. Model:Basic	1	16.91	$\leq$0.001***
Model:Basic + 2 vs. Model:Basic	1	5.07	= 0.024*
Model:Basic + 3 vs. Model:Basic	1	3.06	= 0.08.
Model:Basic + 1 + 2 vs. Model:Basic	2	18.59	$\leq$0.001***
Model:Basic + 1 + 3 vs. Model:Basic	2	20.38	$\leq$0.001***
Model:Basic + 2 + 3 vs. Model:Basic	2	8.10	= 0.017*
Model:Basic + 1 + 2 + 3 vs. Model:Basic	3	21.59	$\leq$0.001***
Model:Basic + 1 + 3 vs. Model:Basic + 1	1	3.38	= 0.066.
Model:Basic + 1 + 2 vs. Model:Basic + 1	1	5.11	= 0.024*
Model:Basic + 1 + 2 +3 vs. Model:Basic + 1 + 2	1	4.07	= 0.044*

The basic model is a model with direct effects of vote, opponent, and the interaction of opponent and vote on the social distance plus a direct effect of social distance on spite. Model "Basic + 1" is the basic model plus a direct effect of vote on spite etc.

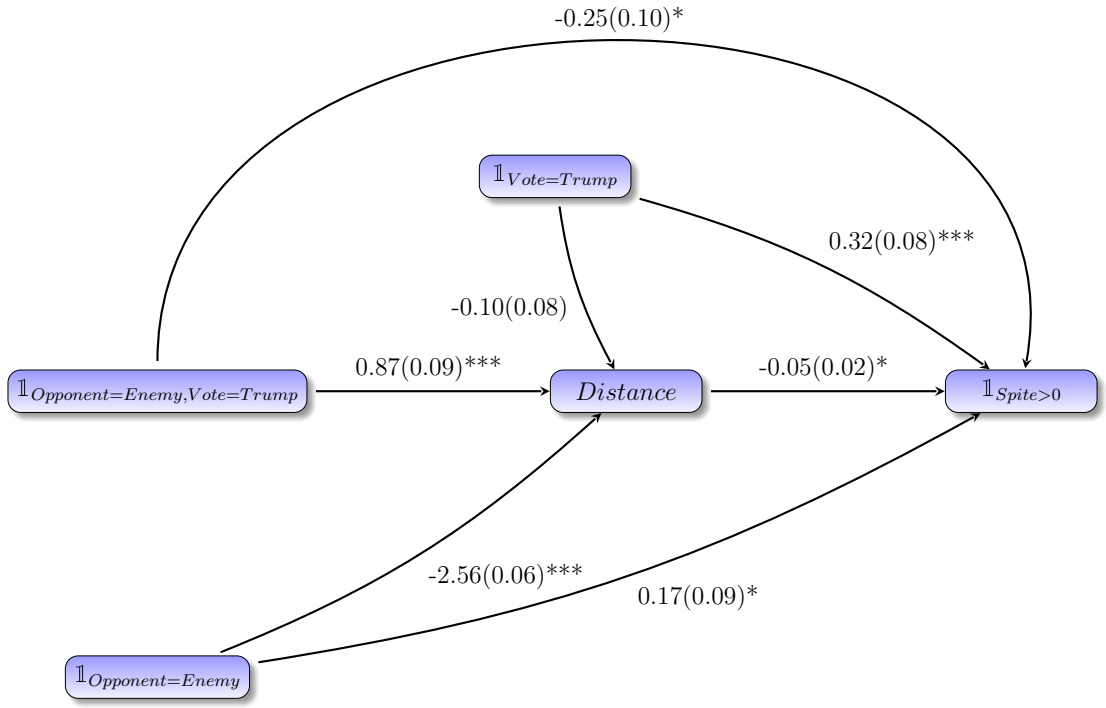
Table 16: Comparing the possible structural equation models on the effect of spite via feeling of warmth.

Comparison	DF	Chisq diff	P-value
Model:Basic + 1 vs. Model:Basic	1	17.39	≤0.001***
Model:Basic + 2 vs. Model:Basic	1	6.42	= 0.011*
Model:Basic + 3 vs. Model:Basic	1	2.20	≥0.10
Model:Basic + 1 + 2 vs. Model:Basic	2	19.80	≤0.001***
Model:Basic + 1 + 3 vs. Model:Basic	2	19.82	≤0.001***
Model:Basic + 2 + 3 vs. Model:Basic	2	8.54	= 0.014*
Model:Basic + 1 + 2 + 3 vs. Model:Basic	3	22.30	≤0.001***
Model:Basic + 1 + 3 vs. Model:Basic + 1	1	2.30	≥0.10
Model:Basic + 1 + 2 vs. Model:Basic + 1	1	6.42	= 0.011*
Model:Basic + 1 + 2 +3 vs. Model:Basic + 1 + 2	1	2.88	= 0.09.

The basic model is a model with direct effects of vote, opponent, and the interaction of opponent and vote on the feeling of warmth plus a direct effect of the feeling of warmth on the decision to behave spitefully. Model "Basic + 1" is the basic model plus a direct effect of vote on spite etc.

D.2. Results of the SEM

D.2.1. SEM and social distance



Notes: ·p<0.10;*p<0.05;**p<0.01;***p<0.001;

Figure 10: SEM of spite via social distance.
Structural Equation Model on the effect on the binary decision to behave spitefully or not, via social distance. Numbers in brackets denote standard errors.

As Figure 10 shows and as seen before, the opponent and the interaction of Trump voters and opponent had a statistically significant influence on the social distance ($p_{\text{Opponent}} \leq 0.001$, $p_{\text{Interaction}} \leq 0.001$) while Trump voters did not show significantly different attitudes in the social distance towards friends compared to Clinton voters ($p_{\text{Vote}} \geq 0.05$).

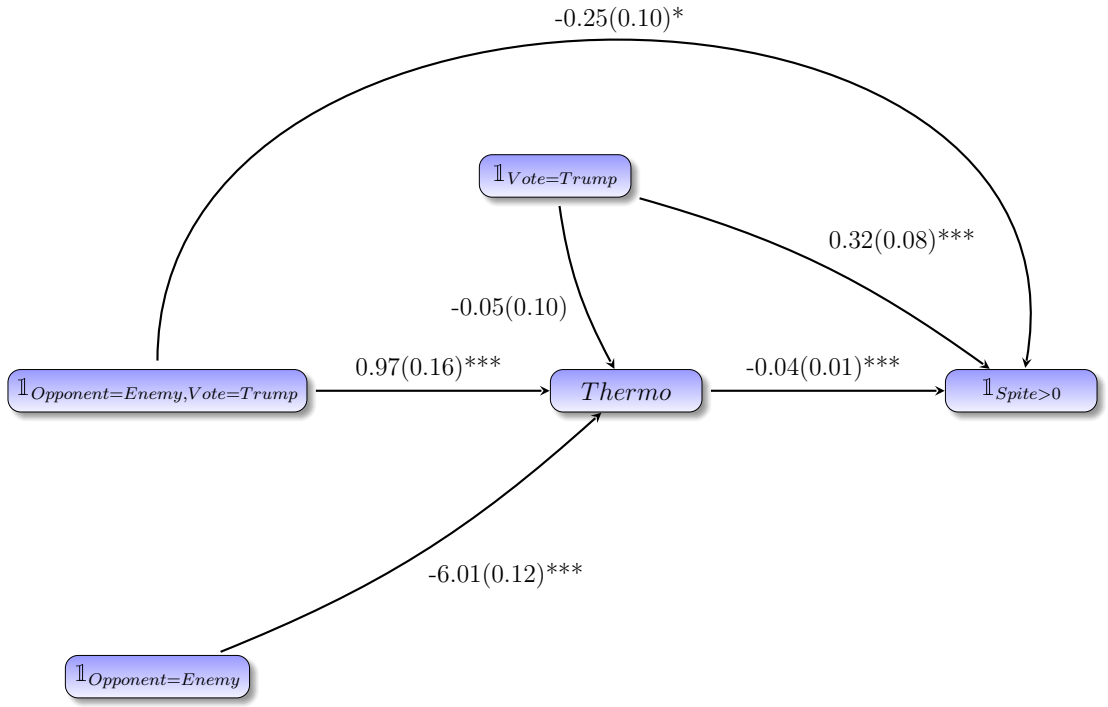
Additionally, all direct effects statistically significantly influence the decision to behave spitefully ($p_{\text{Distance}} = 0.041$, $p_{\text{Vote.Direct}} \leq 0.001$, $p_{\text{Interaction.Direct}} = 0.018$, $p_{\text{Opponent.Direct}} = 0.043$).

Using 10,000 bootstrapped samples we obtain the 95% confidence intervals for the indirect effects of the opponent ([0.005,0.230]), the vote ([-0.004,0.012]) and the interaction of opponent and vote ([-0.079,-0.001]). As the confidence intervals for the indirect effects of Opponent and Interaction did not include the zero, we can reject the null hypothesis of no indirect effect. However, we cannot show

an indirect effect of Trump voters on spite via social distance. The total effects (combining both the indirect and direct effects) are statistically significant for the opponent ([0.166,0.414]), the vote ([0.171,0.471]), and the interaction of opponent and vote ([-0.486,-0.085]).

Thus, we have seen that Trump voters have a higher odd of behaving spitefully, that Clinton voters have a higher odd of behaving spitefully towards their enemies, that Trump voters have a lower odd of behaving spitefully towards their enemies than Clinton voters, and that an increased feeling of closeness leads to a decrease in the odds of behaving spitefully. We also have seen that the opponent strongly influences the feeling of closeness and that Clinton and Trump voters differ in their attitudes towards their opponents. We also have seen that this indirectly influences the decision to be spiteful. Overall, we can see that Trump voters have higher odd of behaving spitefully, which is not mediated through the social distance, and that Clinton voters have a higher odd of behaving spitefully towards their enemies which is mediated through social distance and that Trump voters have a lower odd of behaving spitefully towards their enemies than Clinton voters which also is mediated through social distance.

D.2.2. SEM and feeling of warmth



Notes: ·p<0.10;*p<0.05;**p<0.01;***p<0.001;

Figure 11: SEM of spite via feeling of warmth.
Structural Equation Model on the effect on the binary decision to behave spitefully or not, via the feeling of warmth. Numbers in brackets denote standard errors.

Using the same model for the indirect effect of feeling of warmth leads to similar results. The opponent and the interaction of Trump voters and opponent had a statistically significant influence on the social distance ($p_{\text{Opponent}} \leq 0.001$, $p_{\text{Interaction}} \leq 0.001$) while Trump voters did not show significantly different attitudes in the social distance towards friends compared to Clinton voters ($p_{\text{Vote}} \geq 0.05$).

Additionally, all direct effects statistically significantly influence the decision to behave spitefully ($p_{Distance} \leq 0.001$, $p_{\text{Vote.Direct}} \leq 0.001$, $p_{\text{Interaction.Direct}} = 0.011$).

Using again 10,000 bootstrapped samples we obtain the 95% confidence intervals for the indirect effects of the opponent ([0.126,0.318]), the vote ([-0.006,0.009]) and the interaction of opponent and vote ([-0.055,-0.017]). As the confidence intervals for the indirect effects of Opponent and Interaction did not include the zero, we can reject the null hypothesis of no indirect effect. As with the social distance model, we cannot reject the null hypothesis of no indirect effect of the vote. However, the total effects (combining both the indirect and direct effects) are statistically signif-

icant for the opponent ([0.126,0.318]), the vote ([0.171,0.471]), and the interaction of opponent and vote ([-0.486,-0.085]).

Thus, we have seen that Trump voters have a higher odd of behaving spitefully, that Trump voters have a lower odd of behaving spitefully towards their enemies than Clinton voters, and that an increased feeling of warmth leads to a decrease in the odds of behaving spitefully. We also have seen that the opponent strongly influences the feeling of warmth and that Clinton and Trump voters differ in their attitudes towards their opponents. We also have seen that this indirectly influences the decision to be spiteful. Overall, we can see that Trump voters have higher odd of behaving spitefully, which is not mediated through the social distance, and that Clinton voters have a higher odd of behaving spitefully towards their enemies which is mediated through social distance and that Trump voters have a lower odd of behaving spitefully towards their enemies than Clinton voters which also is mediated through social distance.

Thus, we obtain basically the same results if feeling of warmth is used as the measure of attitudes instead of the feeling of closeness.

E. Spite questionnaire

To have a better understanding of the general spite tendencies of participants, we elicited the spitefulness of participants by using the questionnaire by Marcus et al. (2014), which is based on 17 items. Examples of the 17 questions are the following:[48]

- If I am checking out at a store and I feel like the person in line behind me is rushing me, then I will sometimes slow down and take extra time to pay.
- I would rather no one get extra credit in a class if it meant that others would receive more credit than me.

Participants were asked to indicate their agreement on a scale between 1 and 5. Higher scores on the scale indicate higher spitefulness.

Figure 12 shows the classification of how spiteful participants are over time for Clinton and Trump voters. We can see that Clinton and Trump voters do not differ substantially[49] and that there is no change over time in attitudes. As the

[48] All questions are shown in related materials.

[49] Trump voters seem to be classified as slightly more spiteful ($M = 2.36$, $SD = 1$) than Clinton voters($M = 2.31$, $SD = 0.91$), which, however, is not significant at the 5 % level, $t(1948.3)=$

measure does not rate the spite *towards* others but rather the individual spite attitude there are no differences in spite attitudes between those participants who have been assigned an enemy as the opponent and those who have been assigned a friend.[50] Given that there are no differences between Clinton and Trump voters, the results are omitted to the appendix and are reported solely for reasons of transparency.

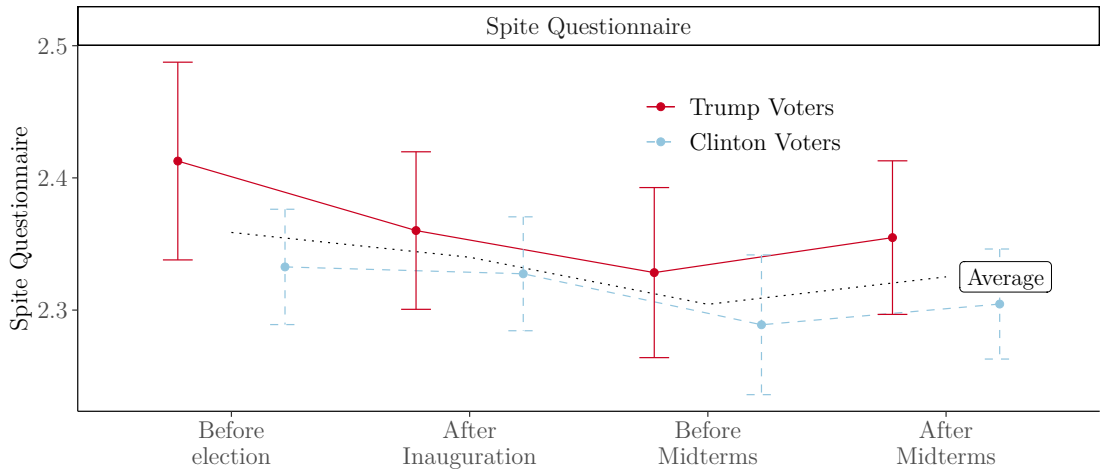


Figure 12: Spite questionnaire
The figure depicts how spiteful participants were classified following the spite questionnaire in each of the four waves. Red solid lines denote the general spite attitudes of Trump voters while blue dashed lines denote the general spite attitudes of Clinton voters. The black dotted line depicts the average spite attitude. Tie fighters represent standard errors.

-1.2, $p \geq 0.05$

[50] Friends: $M = 2.34$, $SD = 0.93$; Enemies $M = 2.33$, $SD = 0.95$. The difference is not significant, $t(1948.3) = -1.2$, $p \geq 0.05$.

F. Moral attitudes

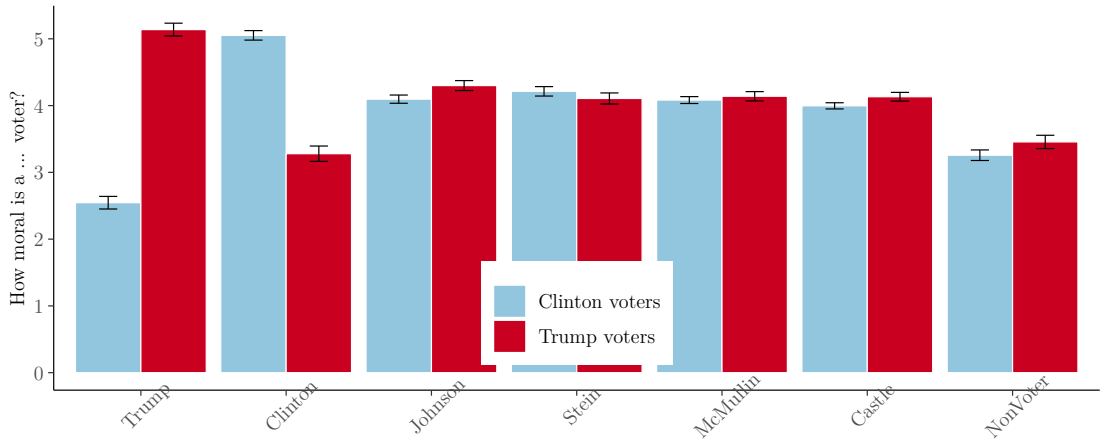


Figure 13: How moral participants report to consider different voters.
The figure depicts how moral participants consider Trump, Clinton, Johnson, Stein, McMullin, and Castle voters as well as people who did not vote during the presidential election of 2016. The scale goes from 1 (not moral at all) to 7 (very moral). Red bars denote the attitudes of Trump voters, while blue bars denote the attitudes of Clinton voters. Tie fighters denote 95% confidence intervals.

To further investigate whether the perception of morality might drive the difference between Clinton and Trump voters, we elicit their moral attitudes during the third and fourth wave of the experiment (late October 2018 and early November 2018). In particular, we asked participants how moral they consider a Trump, Clinton, Johnson, Stein, McMullin, and Castle voter as well as people who did not vote during the presidential election 2016. The results are shown in Figure 13.

The moral attitudes of both Clinton and Trump voters are very similar and rather neutral for Johnson, Stein, McMullin, and Castle voters. Both Clinton and Trump voters considered non-voters significantly less moral than Johnson, Stein, McMullin, and Castle voters.[51] Interestingly, Clinton voters considered Clinton voters as moral as Trump voters considered Trump voters[52], which was significantly bet-

[51] In particular, non-voters were considered to be $M = 3.34$ points moral on a scale from one to seven while Johnson, Stein, McMullin, and Castle voters were considered on average to be $M = 4.13$ points moral on a scale from one to seven, a highly significant difference ($t(1353) = 21.4$, $p \leq 0.001$).

[52] In particular, Clinton voters considered fellow Clinton voters to be $M = 5.05$ points moral on a scale from one to seven while Trump voters considered fellow Trump voters to be $M = 5.14$ points moral on a scale from one to seven, $t(1083.1) = -1.4$, $p \geq 0.05$.

ter than Johnson, Stein, McMullin, and Castle voters ($t(1353)$= -27.2, $p \leq 0.001$).

More importantly, we can see that Trump voters considered Clinton voters significanlty less moral than fellow Trump voters[53], than Johnson, Stein, McMullin, Castle voters[54] and even less than non-voters[55].

The same pattern can be found for Clinton voters who considered Trump voters significantly less moral than fellow Clinton voters[56], than Johnson, Stein, McMullin, Castle voters[57] and even less than non-voters[58]. Even more interestingly, the difference in morality between friends and enemies is much more pronounced for Clinton voters, who considered Trump voters significantly less moral than Trump voters considered Clinton voters.[59]

The results provide suggestive evidence that morality might, in fact, be driving the difference in attitudes of participants and, in turn, might influence spite behavior.

[53] In particular, Trump voters considered fellow Trump voters to be $M = 5.14$ points moral on a scale from one to seven while Clintons voters were considered to be $M = 3.28$ points moral on a scale from one to seven, a highly significant difference $t(543)$= 22.1, $p \leq 0.001$.

[54] In particular, Trump voters considered Johnson, Stein, McMullin, Castle on average to be $M = 4.17$ points moral on a scale from one to seven while Clintons voters were considered to be $M = 3.28$ points moral on a scale from one to seven, a highly significant difference $t(543)$= 13.2, $p \leq 0.001$.

[55] In particular, Trump voters considered non-voters on average to be $M = 3.46$ points moral on a scale from one to seven while Clintons voters were considered to be $M = 3.28$ points moral on a scale from one to seven, a highly significant difference $t(543)$= -2.7, p= 0.008.

[56] In particular, Clinton voters considered fellow Clinton voters to be $M = 5.05$ points moral on a scale from one to seven while Trump voters were considered to be $M = 2.55$ points moral on a scale from one to seven, a highly significant difference $t(809)$= -36.9, $p \leq 0.001$.

[57] In particular, Clinton voters considered Johnson, Stein, McMullin, Castle on average to be $M = 4.10$ points moral on a scale from one to seven while Trump voters were considered to be $M = 2.55$ points moral on a scale from one to seven, a highly significant difference $t(809)$= 29.1, $p \leq 0.001$.

[58] In particular, Clinton voters considered non-voters on average to be $M = 3.26$ points moral on a scale from one to seven while Clintons voters were considered to be $M = 2.55$ points moral on a scale from one to seven, a highly significant difference $t(809)$= -13.8, $p \leq 0.001$.

[59] In particular, Clinton voters considered Trump voters to be $M = 2.55$ points moral on a scale from one to seven while Trump voters considered Clintons voters to be $M = 3.28$ points moral on a scale from one to seven, a highly significant difference $t(1165.1)$= -9.7, $p \leq 0.001$.

G. Instructions

In this section, we show the instructions for the own spite measure, the spite questionnaire, the feeling thermometer and the social distance questionnaire. Further, we show the manipulation and the attention check. The procedure of the experiment can be seen in Figure 14.

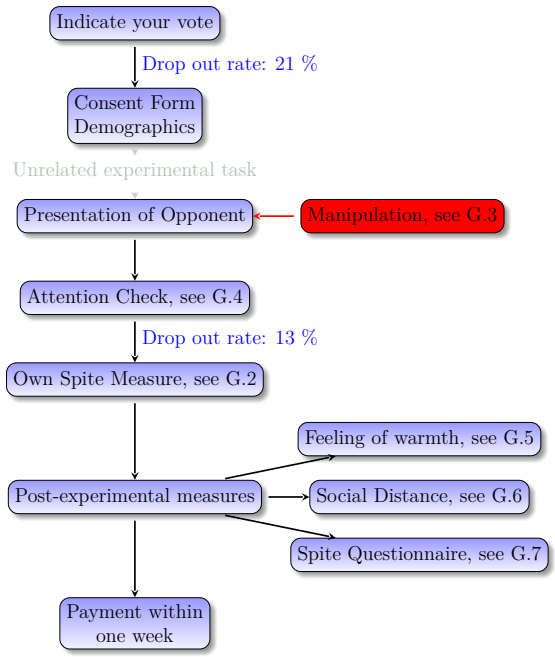


Figure 14: Procedure of the experiment.
The figure shows the procedure of the experiment.

G.1. Welcome

Welcome to this experiment in the economics of market decision making. If you follow these instructions carefully and make good decisions you will earn a considerable amount of money that will be paid to you within one week to your MTurk account. We ask that you pay close attention to the instructions.

Note that one of the main guidelines in the experimental economics is that we do NOT deceive participants (see for example https://en.wikipedia.org/wiki/Experimental_economics). Hence, all rules and restrictions will indeed be implemented in the way we describe them. We go to great lengths to ensure that assignments, randomization of variables and rules are implemented exactly in the

way they are presented here to you!

In this experiment, you will be assigned an opponent. Your payoff will depend on his/her decisions and his/her payoff may depend on your decisions. Typically every person is assigned, one opponent.

To comply with the non-deception-rules of economics we also need to inform you about a technical issue: It may happen that more than one person is assigned to another person. In such a (rather rare) case it will be randomly decided whose decision will be payoff relevant to this other person. Thus, your payoff will always depend on the decision of somebody else. Your decision will influence the payoff of your assigned partner in most cases. It, however, may happen that your decision does not impact the payoff of your partner as somebody else's choice has been determined payoff-relevant for your partner.

G.2. Own spite measure

In this task, you will be paired with another player, whom we will refer to as the opponent. All of your choices will be confidential. After you take your decisions this task will not be repeated and there is no further interaction with your opponent.

You will be making a series of decisions about allocating resources between you and your opponent. For each of the following questions, please indicate the distribution you prefer most by selecting the button below the payoff allocations. You can only make one selection for each question. Your decisions will yield money for both yourself and your opponent.

Each point shown is worth 0.2 cents (100 points = 20 cents).

In the example below, a person has chosen to distribute the payoff so that he/she receives 50 points (=10 cents), while his opponent receives 40 points (=8 cents).

There are no right or wrong answers, this is all about personal preferences. After you have made your decision, select the resulting distribution of money by clicking on the button below your choice. As you can see, your choices will influence both

the amount of money you receive as well as the amount of money your opponent receives.

At the end of the experiment, a computer program will randomly pick either you or your opponent as the payoff-relevant decision maker.

Only one of the following decisions will be payoff relevant. Which decision will be paid will be determined by a random process at the end of the experiment. Hence, you have to take all decisions seriously as any of those can be chosen by the random process with equal probability.

Please indicate your choice for each of the following distributions.
Note: These decisions are payoff relevant and will influence your payment!

[[Participants had to make choices as shown in figure 1]]

G.3. Assignment

[[Treatment 1:]]
Individuals who have indicated to vote for Donald Trump at the beginning of the experiment were assigned to a group called "red".
Individuals who have indicated to vote for Hillary Clinton at the beginning of the experiment were assigned to a group called "blue".

Your assigned opponent indicated to vote for Hillary Clinton. Hence, your opponent was assigned to be a member of the group "blue".

[[Treatment 2:]]
Individuals who have indicated to vote for Donald Trump at the beginning of the experiment were assigned to a group called "red".
Individuals who have indicated to vote for Hillary Clinton at the beginning of the experiment were assigned to a group called "blue".

Your assigned opponent indicated to vote for Donald Trump. Hence, your opponent was assigned to be a member of the group "red".

G.4. Attention Check

Which group does your opponent belong to?

- (a) "Red"
- (b) "Blue"

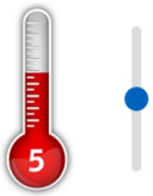
- (c) "Green"

G.5. Feeling thermometer

This measure is called a "feeling thermometer" as it measures your feeling towards groups.

Here is how it works:
If you do not know too much about a group, or don't feel particularly warm or cold toward them, then you should place them in the middle, at the 5-degree mark.
If you have a warm feeling toward a group or feel favorably toward it, you would give it a score somewhere between 5 and 10 depending on how warm your feeling is toward the group.
On the other hand, if you don't feel very favorable toward some of these groups -if there are some you don't care for too much - then you would place them somewhere between 0 and 5.
Please indicate below your feeling towards the named group of people.

Where would you put the [[republicans]][[democrats]][[people who will vote for Hillary Clinton]][[people who will vote for Donald Trump]] on the thermometer?



G.6. Social distance questionnaire

Please rate the following statements. The person in question is a person who has indicated to vote for [[Hillary Clinton]][[Donald Trump]].

	Strongly agree	Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Disagree	Strongly disagree
This appears to be a likeable person.	○	○	○	○	○	○	○
I would like this person to be a close personal friend.	○	○	○	○	○	○	○
I would like this person to move into my neighborhood.	○	○	○	○	○	○	○
I would like this person to come and work at the same place I do.	○	○	○	○	○	○	○
This is a person who is similar to me.	○	○	○	○	○	○	○
I would like to have this person marry into my family.	○	○	○	○	○	○	○
This is the kind of person that I tend to avoid.	○	○	○	○	○	○	○

G.7. Spite questionnaire

The questions of the questionnaire according to Marcus et al. (2014) included the following questions:

- I would be willing to take a punch if it meant that someone I did not like would receive two punches.
- I would be willing to pay more for some goods and services if other people I did not like had to pay even more.
- If I was one of the last students in a classroom taking an exam and I noticed that the instructor looked impatient, I would be sure to take my time finishing the exam just to irritate him or her.
- If my neighbor complained about the appearance of my front yard, I would be tempted to make it look worse just to annoy him or her.
- It might be worth risking my reputation in order to spread gossip about someone I did not like.
- If I am going to my car in a crowded parking lot and it appears that another driver wants my parking space, then I will make sure to take my time pulling out of the parking space.
- I hope that elected officials are successful in their efforts to improve my community even if I opposed their election. (reverse scored)
- If my neighbor complained that I was playing my music too loud, then I might turn up the music even louder just to irritate him or her, even if meant I could get fined.
- I would be happy receiving extra credit in a class even if other students received more points than me. (reverse scored)
- Part of me enjoys seeing the people I do not like fail even if their failure hurts me in some way.
- If I am checking out at a store and I feel like the person in line behind me is rushing me, then I will sometimes slow down and take extra time to pay.
- It is sometimes worth a little suffering on my part to see others receive the punishment they deserve.
- I would take on extra work at my job if it meant that one of my coworkers who I did not like would also have to do extra work.
- If I had the opportunity, then I would gladly pay a small sum of money to see a classmate who I do not like fail his or her final exam.
- There have been times when I was willing to suffer some small harm so that I could punish someone else who deserved it.
- I would rather no one get extra

credit in a class if it meant that others would receive more credit than me.

- If I opposed the election of an official, then I would be glad to see him or her fail even if their failure hurt my community.