Threat-Inducing Violent Events Exacerbate Social Desirability Bias in Survey Responses

Shane P. Singh
202 Herty Drive
University of Georgia
Athens, GA 30602
singh@uga.edu
517-214-3400

Jaroslav Tir
UCB 333
University of Colorado Boulder
Boulder, CO 80309
jtir@colorado.edu
303-492-7871

Keywords: Social Desirability Bias; Threat; Survey Methods; Turnout

Running Title: Threat and Social Desirability Bias

Acknowledgments: For helpful comments, we wish to thank members of the Political Science Department at the University of Colorado Boulder, and Jennifer Wolak in particular.

Abstract

A key challenge in survey research is social desirability bias: respondents feel pressured to report acceptable attitudes and behaviors. Building on established findings, we argue that threat-inducing violent events are a heretofore unaccounted for driver of social desirability bias. We probe this argument by investigating whether fatal terror attacks lead respondents to overreport past electoral participation, a well-known and measurable result of social desirability bias. Using a cross-national analysis and natural and survey experiments, we show that fatal terror attacks generate turnout overreporting. This highlights that threat-inducing violent events induce social desirability; that researchers need to account for the timing of survey fieldwork vis-à-vis such events; and that some of the previously reported post-violent conflict increases in political participation may be more apparent than real.

Word Count: 9,822

Verification Materials: The data and materials required to verify the computational reproducibility of the results, procedures and analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network, at:
https://doi.org/10.7910/DVN/XOCNK1

In this paper, we uncover a heretofore unaccounted for driver of social desirability bias in survey responses, violent events. Social desirability bias, which has long been linked to an array of personality traits, predispositions, and demographic characteristics (Edwards 1953; Crowne and Marlowe 1960), harms the quality of survey data because it induces respondents to report subscribing to attitudes or engaging in behaviors that are deemed socially appropriate when they do not actually hold such beliefs or engage in such behaviors.

Among a variety of types of crises a society may face (e.g., economic breakdown or constitutional crisis), violent events are unique in that they generate an acute sense of existential threat and harm. Accordingly, a long line of studies has shown that threat-inducing violent events have attitudinal and behavioral consequences. Focusing on events such as terrorism or war, researchers report impacts on, for example, vote choice, incumbent support, and attitudes toward minorities and women (see, e.g., Aldrich, Sullivan and Borgida 1989; James and Rioux 1998; Schildkraut and Furia 2003; Lawless 2004; Davis 2007; Norpoth and Sidman 2007; Hutchison and Gibler 2007; Kam and Ramos 2008; Berinsky 2009; Williams, Brulé and Koch 2010; Holman, Merolla and Zechmeister 2011; Tir and Singh 2013, 2015; Singh and Tir 2018, 2019). Furthermore, a number of studies have found specific impacts on political participation. For example, terrorist attacks in Spain significantly increased individuals' reported intent to participate in a future democratic election (Balcells and Torrats-Espinosa 2018); family members and residential neighbors of victims of the September 11th attacks have become more active in politics (Hersh 2013); experience of violence in northern Uganda increased claimed political engagement (Blattman 2009); and individuals whose households experienced more intense violence during in the civil war in Sierra Leone became more likely to report voting (Bellows and Miguel 2009). Although a number of studies find a link between violence and increased levels of subsequent political participation, the survey-based data used in many of these works leaves room for skepticism as to whether reported behaviors are more apparent than real.

Therefore, while we expect violent events to affect responses to a variety of survey questions on attitudes and behaviors that are subject to social desirability bias, in this article we focus on electoral participation. Through it, a citizen not only chooses among various candidates and policy preferences, but also demonstrates broader support for the political system on which society is based. Among survey respondents, social desirability bias can thus lead to deliberate overreporting of past electoral participation (e.g., Sigelman 1982; Bernstein, Chadha and Montjoy 2001; Ansolabehere and Hersh 2012; Hansen and Tyner forthcoming). We argue that violent events present a threat to society, which, in turn, prompts individuals to feel pressured to report engaging in behaviors that are perceived as being supportive of it, thus signalling that they are acting in society-supporting ways. That is, the sense of threat accompanying violent events has the capacity to elicit a well-known—but heretofore unassociated—outcome, social desirability bias. Observably, increased social desirability bias boosts the likelihood that an individual will report having voted when they did not.

To probe our expectations, we leverage what are arguably the most attention-grabbing, high-profile, sudden, and common types of violent events nowadays: deadly terrorist attacks. Our investigation consists of three separate studies. The first is a cross-national examination in which we compare answers to questions about recalled electoral turnout among respondents in countries with and without fatal terror attacks in the time frame between an election and the beginning of a post-election survey. In the second study, we use a regression discontinuity approach to leverage a natural experiment that took place in the Netherlands, in which a survey was interrupted by the murder of the Dutch filmmaker Theo van Gogh. In the third study, we design and implement a survey experiment in India that leverages a deadly 2019 bomb attack on Indian national police.

Across all three studies, the estimated impact of exposure to violent attacks on overreported turnout is positive and substantively and statistically significant. In the first, we find that being surveyed after a fatal terror attack is associated with about a seven percentage point increase in the probability of reporting turnout. Because we only consider post-election terror attacks and hold actual turnout rates constant, differences in recalled participation are attributable to overreporting rather than true variation in turnout. In the second study, we find that the van Gogh attack boosted reported turnout by nearly eight percentage points when comparing individuals who responded to the survey after as opposed to before the murder. Because every individual was interviewed *after* the election, this increase is attributable to overreporting. In the third study, we find that respondents who were shown a news report about the deadly terror attack on Indian national police were about ten percentage points more likely to report having voted than those in who were asked to read an innocuous news story.

A series of follow-up analyses help verify our proposed causal pathway of the sense of threat driving social desirability bias in survey responses. We show that effects of fatal terror attacks are stronger when there are more attacks or casualties, for those living nearest to an attack, and for demographic groups thought to be most prone to social desirability bias in claimed electoral participation. Furthermore, we show that such attacks impact responses to other survey questions known to be influenced by social desirability bias. Finally, we assess observationally equivalent alternate scenarios through which either genuine increases in intentions to vote or societal in-group bonding drive our findings. We fail to detect evidence for either.

Our findings identify threat-inducing violent events as a heretofore unaccounted driver of social desirability bias. This has notable substantive implications for survey-based research that links violence and subsequent heightened political participation. We offer a cautionary note, in that at least some of this increase may be more apparent than real, as survey respondents feel pressured to provide society-supporting responses whether or not they actually engaged in specified activities. In the conclusion, we discuss the substantive and methodological implications of our findings and give practical advice for survey researchers.

Threat-Inducing Violent Events and Social Desirability Bias

Sociologists such as Simmel (1898) and Coser (1956) long ago argued that individuals' attitudes, behavior, and predispositions are affected when the security of their society is threatened. Importantly, violent events such as terror attacks do not only invoke a potential feeling of threat to the individuals specifically targeted in the event. Rather, they are more broadly interpreted as an affront to the society as a whole. In the wake of events such as the September 11th attacks, for example, individuals are consequently thought to rally around their society (e.g., Schildkraut and Furia 2003; Davis 2007; Kam and Ramos 2008; see also Huddy, Feldman and Weber 2007; Tir and Singh 2013).

The threat associated with violent events creates in individuals the sense that the security of society is at risk. Individuals thus feel pressure to signal that they behave in ways that are supportive of their embattled country or are otherwise valuable members of society. For survey respondents, this will manifest as a feeling of pressure to provide socially acceptable answers to questions about socially-sensitive behaviors or personal qualities. That is, threat from violent events activates or strengthens social desirability bias in survey responses. By focusing on threat-driven pressure, we depart from the logic that threat may work indirectly by first creating a more genuine feeling of in-group bonding, where the in-group is generally defined by the majority ethnicity, race, or religion. Our argument is instead that the sense of threat to society directly leads survey respondents to feel like they need to pretend to have acted in socially desirable ways or to be valuable citizens to *signal* their support for society.[1]

Although we expect this pressure to manifest itself in a variety of questions about behaviors and attitudes subject to social desirability bias, we focus on a specific but highly important act: voting. Through electoral participation, citizens of democracies not only choose among various candidates, parties, and policies on offer, but also indicate their support for the political system and their society more broadly (see, e.g., Rahn, Brehm and Carlson 1999; Huddy and Khatib 2007; Schildkraut 2007; Wright, Citrin and Wand 2012). Tellingly, those who report more favorability toward the democratic system and its institutions are more likely to take part in it (Grönlund and Setälä 2007; Karp and Milazzo 2015).

Because voting is integral to democratic functioning, turnout overreporting, when a survey respondent reports having voted when they did not, is one of the most commonly studied and detected consequences of social desirability bias in democracies.[2] Existing research finds that individuals who belong to social or demographic groups that are "expected" to vote at high rates are more susceptible to turnout overreporting (e.g., Sigelman 1982; Bernstein, Chadha and Montjoy 2001; Ansolabehere and Hersh 2012; Dahlgaard et al. 2019; Hansen and Tyner forthcoming). Moreover, overreporting is especially pronounced in salient and well-attended elections—those in which abstention is a particularly socially deviant behavior (Karp and Brockington 2005; Górecki 2011). Additional evidence for the link between social desirability bias and turnout overreporting comes from survey experiments that randomize question wording and response options: constructions that frame abstention as socially acceptable reduce overreporting (e.g., Belli, Moore and VanHoewyk 2006; Morin-Chassé et al. 2017).

We argue that the sense of threat to society associated with violent events creates in individuals a feeling of pressure to provide socially acceptable answers to survey questions. With our focus on

[1] In follow-up analyses reported below, we investigate whether in-group bonding has merit as an alternate causal pathway.

[2] Nonvoters may also decline to participate in surveys at a higher rate than voters. Both nonresponse among abstainers and overreporting can contribute to overestimated turnout rates (Dahlgaard et al. 2019).

turnout overreporting, we thus put forth the following hypothesis:

> **Hypothesis**: Threat-inducing violent events increase the likelihood that individuals will overreport voting in past elections.

Study 1: Cross-National Associations

In this initial test of our hypothesis, we compare answers to questions about voter turnout among respondents in countries with and without violent events in the time between an election and their survey interview date. We focus on terrorist attacks involving fatalities, as these are arguably the most attention-grabbing, high-profile, sudden, and common types of threat-inducing violent events in recent years. The fatality condition constitutes a common restriction in research on violent conflict (see e.g., Oneal and Tir 2006) and helps us identify events that are likely to be, per our theoretical mechanism, capable of inducing a sense of threat. Our data on terror events come from the Global Terrorism Database (GTD),[3] which records the details of terror events around the world from 1970.

Our survey data come from Modules I-IV of the Comparative Study of Electoral Systems (CSES),[4] a collection of harmonized post-election surveys from dozens of countries beginning in 1996. We cross-referenced CSES post-election interview timing with the dates of fatal terror events recorded in the GTD, subsequently creating a measure of the number of fatal terror events that took place *after* the most recent election but before a respondent's interview date.[5] The sequencing of events is as follows:

> most recent election at t_0 $\rightarrow$ potential fatal terror attack(s) at t_1 $\rightarrow$ CSES post-election survey fieldwork begins at t_2

We also collected information on the number of fatalities in these terror events. Because the CSES inquires about recalled turnout in the recent elections, it is not possible for these post-

[3]The GTD is housed at the University of Maryland (`https://www.start.umd.edu/gtd/`).

[4]The Comparative Study of Electoral Systems is housed at the University of Michigan and the GESIS-Leibniz Institute for the Social Sciences (`http://www.cses.org/`).

[5]We exclude from the sample CSES studies that do not include respondents' interview dates. Further, about 20 percent of CSES respondents were allowed to mail back or complete online all or a portion of their survey questionnaires. Because these survey modes hamper our ability to identify a respondent's actual interview date, we only retain respondents who were interviewed in person or by telephone. Across the election surveys in our sample, the median time from the election until a respondent's interview is about 25 days.

election terror events to have directly influenced electoral behavior. Instead, per our expectations, they should increase the extent to which respondents *report* having voted.

We take as our initial dependent variable responses to a question about turnout in the national election[6] immediately preceding the survey, coding those who reported participating in the relevant lower house contest as 1 and reported abstainers as 0. Many CSES surveys also ask about turnout in an election that took place before the election held immediately prior to survey fieldwork. These are generally the penultimate nationwide election. As a second dependent variable, we gauge reported turnout in these previous lower house elections, again coding reported voters as 1 and reported abstainers as 0. It is again impossible that terror events that took place after the most recent election directly affected electoral behavior in some prior contest. However, per our theory, we expect that survey respondents are more likely to report having voted in these prior elections when their country was the subject of a recent attack.

We create three explanatory variables. The first is a binary measure that differentiates respondents in countries that experienced at least one fatal terror event in the time between the most recent election and their survey interview from respondents whose post-election interviews took place in countries without a post-election, pre-survey interview terror attack. The second two measures are a count of fatal terror events in the election-to-interview period and the number of fatalities due to these events. These allow us to probe our theorized threat mechanism, as a higher number of attacks and more fatalities should be more threat-inducing. We take the log of both count variables under the assumption that increases in the number of terror events and fatalities have a diminishing relationship with reported turnout. To account for zeros, we added one to all observations before taking the logarithms.

The post-election surveys from the CSES included in the sample are listed in Table 1. In three of these studies (Mexico 1997 and 2015; Thailand 2011) the terror events took place during fieldwork. However, because the number of respondents in the sample who were interviewed either before or after the terror events was very low, we were unable to leverage this quasi-random assignment. We leverage a more amenable interrupted survey in Study 2.

[6] The wording of the turnout questions varies across countries and follows national standards.

Table 1: CSES Surveys Included in Study 1

Argentina 2015	Hungary 1998	*Poland 2011*
Austria 2008	*Hungary 2002*	Romania 1996
Austria 2013	Iceland 1999	*Romania 2004*
Belarus 2008	*Iceland 2003*	*Romania 2012*
Brazil 2002	*Iceland 2009*	***Serbia 2012***
Brazil 2006	*Iceland 2013*	*Slovakia 2010*
Brazil 2010	*Ireland 2007*	*Slovakia 2016*
Brazil 2014	**Israel 1996**	Slovenia 1996
Bulgaria 2001	*Israel 2003*	*Slovenia 2004*
Bulgaria 2014	***Israel 2006***	*Slovenia 2008*
Canada 2008	*Israel 2013*	*Slovenia 2011*
Canada 2015	Japan 1996	***South Africa 2014***
Chile 2005	**Kenya 2013**	South Korea 2000
Chile 2009	*Latvia 2011*	*South Korea 2004*
Croatia 2007	*Latvia 2014*	Spain 1996
Czech Republic 1996	**Mexico 1997**	Spain 2000
Czech Republic 2006	Mexico 2000	*Spain 2004*
Czech Republic 2010	Mexico 2009	Spain 2008
Czech Republic 2013	*Mexico 2012*	Sweden 1998
Denmark 2007	**Mexico 2015**	*Sweden 2002*
Estonia 2011	*Montenegro 2012*	*Sweden 2006*
Finland 2007	Netherlands 1998	*Sweden 2014*
Finland 2011	*Netherlands 2006*	Switzerland 1999
France 2007	Norway 1997	*Switzerland 2011*
Germany 1998	*Norway 2005*	Taiwan 1996
Germany 2002	*Norway 2009*	*Taiwan 2001*
Germany 2009	*Norway 2013*	*Taiwan 2012*
Germany 2013	Peru 2016	***Thailand 2011***
Greece 2009	***Philippines 2004***	***Turkey 2011***
Greece 2015	***Philippines 2016***	***Turkey 2015***
Hong Kong 1998	Poland 1997	Ukraine 1998
Hong Kong 2000	*Poland 2001*	United States 1996
Hong Kong 2004	*Poland 2005*	*United States 2004*
Hong Kong 2008	*Poland 2007*	United States 2012
Hong Kong 2012		

Note: Entries identify the election surveys from the Comparative Study of Electoral Systems included in the sample analyzed in Study 1. Boldface font indicates a fatal terror event recorded by the GTD took place after the listed election but before survey interviews were completed. Nonitalicized font indicates that the survey is included only in the models used in the creation of the left-hand panels of Figures 1, 2, and 3. Italicized font indicates the survey is included in the models used in the creation of the left- and right-hand panels of the figures.

Statistical Models and Results

We estimate multilevel logistic regression models in which we allow the intercepts to vary randomly over CSES studies. In each model, we include controls for common individual-level correlates of turnout.[7] While we do not expect that these covariates associate with the existence or character

[7]These are: a binary female/male gender variable; age in years and its square (Smets and van Ham 2013); education on a five-point ordinal scale; and income on a five-point ordinal scale. See Appendix F for question wording.

of post-election terror events, their presence in the model helps increase precision. We also control for the number of days between the most recent election and a respondent's interview date. This is necessary because longer election-to-interview gaps increase the likelihood of a fatal terror attack occurring before a respondent's interview, and it is possible that those who are interviewed longer after an election are more likely to misrecall electoral behavior.

At the election level, we control for the actual turnout rate in the election in question.[8] This is a critical control variable, as its inclusion allows us to estimate the impact of post-election terror attacks on turnout reporting while effectively holding actual turnout rates constant. We also control for the level of democratic development using the Polity score[9] (Marshall and Jaggers 2012) in the year of the contest. This helps account for geographic variation in respondents' propensity to give acquiescent responses to survey questions (see, e.g., Franzen and Vogl 2013).

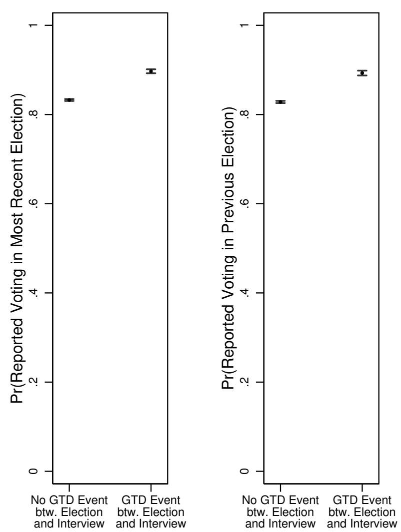

Figure 1: The Probability of Reported Voting Across Countries With and Without Recent Fatal Terror Attacks

Note: Point estimates represent the predicted probability of reported turnout, as predicted with multilevel logistic regression models shown in Table A1. The covariates are held at their observed values for each case in the estimation sample. Brackets indicate 90 percent confidence intervals.

Figure 1 illustrates the results of the models that take the binary terror measure as the ex-

[8]The turnout rate is 100 times the number of voters divided by the number of individuals in the voting age population. Data are from the Varieties of Democracy Project (V-Dem) (Coppedge 2018). Data from V-Dem were unavailable for a small number of elections, for which we used data from the CSES, the International Institute for Democracy and Electoral Assistance, and official sources.

[9]Following Persson and Tabellini (2003, p. 76), we use information from Freedom House to approximate Polity scores in the few countries unrated on the Polity index.

planatory variable. As is clear from the figure, accounting for the covariates, survey respondents are more likely to report having voted if their interview took place after a fatal post-election terror attack. Per the left-hand panel, the probability that a respondent who was interviewed after an election that was not followed by at least one subsequent pre-interview fatal terror attack reported having voted in that election is about 0.83. For those whose interviews took place after one or more fatal post-election terror events, the probability of reported turnout is about 0.90. The two-sided p-value associated with this difference is smaller than 0.001. The pattern in the right-hand panel, in which the outcome is the probability of reported participation in a national election held before the most recent contest, is similar. For those in the no-fatal-terror-attack scenario, the probability of reported turnout in a previous election is about 0.83. For those interviewed after post-election fatal terror events, it is about 0.89, and the two-sided p-value associated with this difference is again smaller than 0.001.

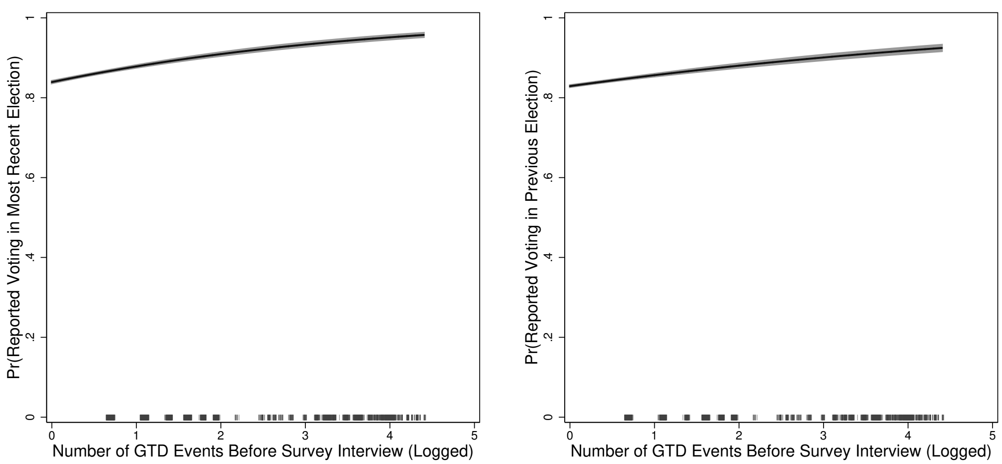

Figure 2: The Probability of Reported Voting Across Countries Vis-à-Vis the Number of Recent Fatal Terror Attacks

Note: The solid curves represent the predicted probability of reported turnout, as predicted with multilevel logistic regression models shown in Table A2. The covariates are held at their observed values for each case in the estimation sample. For respondents whose interviews were preceded by at least one post-election fatal terror attack, the horizontal axis tick marks represent locations on the independent variable, jittered to reduce overlap. Shaded areas indicate 90 percent confidence intervals.

Leveraging variation in terror-induced threat, Figure 2 illustrates the results of the models that take the number of post-election, pre-interview fatal GTD attacks as the explanatory variable. The patterns suggest that not only the existence of violent events, but also their frequency, associates with turnout overreporting. The left-hand panel tracks reported turnout in election held most recently before CSES fieldwork relative to the number of post-election fatal terror events. The right-hand panel tracks reported turnout in a prior election according to this same count. In both panels, the probability of reporting turnout increases along with the number of post-election, pre-interview fatal terror events, and the relationships are nontrivial. Per the left-hand panel, the probability of reported turnout in the most recent election is about 0.84 for those whose interview was preceded by no post-election fatal terror events. As the number of post-election fatal terror events increases one logged unit, this probability increases by, on average, just over four percentage points (two-sided p-value < 0.001). With regard to previous elections, the average increase in the probability of reported participation associated with a unit increase in logged post-election terror events is nearly three percentage points (two-sided p-value < 0.001).

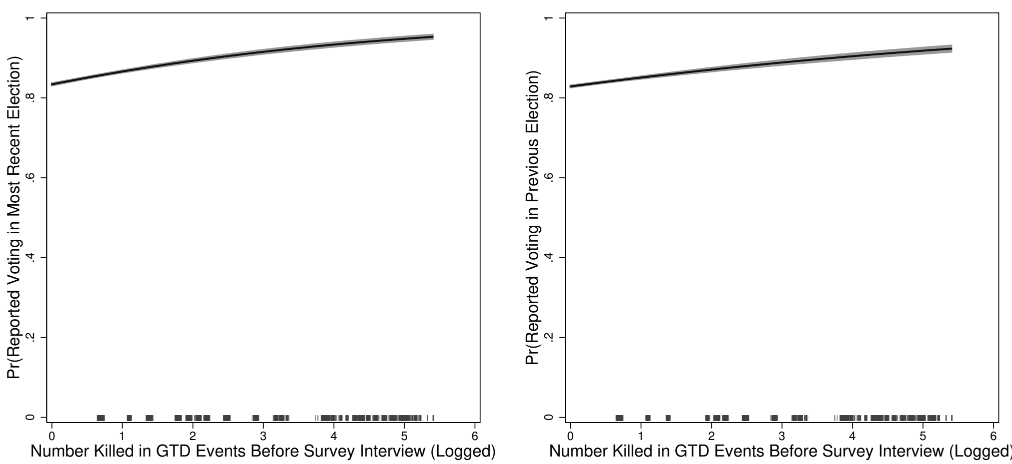

Figure 3: The Probability of Reported Voting Across Countries Vis-à-Vis the Number Killed in Recent Fatal Terror Attacks

Note: The solid curves represent the predicted probability of reported turnout, as predicted with multilevel logistic regression models shown in Table A3. The covariates are held at their observed values for each case in the estimation sample. For respondents whose interviews were preceded by at least one post-election fatal terror attack, the horizontal axis tick marks represent locations on the independent variable, jittered to reduce overlap. Shaded areas indicate 90 percent confidence intervals.

Again accounting for variation in terror-induced threat, Figure 3 illustrates the results of the models in which the number of post-election, pre-interview deaths in fatal terror attacks is the explanatory variable. The patterns shown in the figure indicate that the post-election death count also positively associates with reported turnout. This is true with regard to recalled turnout in both the most recent election (left-hand panel) and in contests held before the most recent election (right-hand panel). Per the left-hand panel, as the number of post-election deaths in fatal terror events increases one logged unit, the probability of reported turnout increases, on average, by just over three percentage points (two-sided p-value < 0.001). With regard previous elections, the average increase in the probability of reported participation associated with a unit increase in logged post-election terror-related deaths is about two percentage points (two-sided p-value < 0.001).

These cross-national results lend credence to our theoretical proposition that the sense of threat drives vote overreporting: the greater the threat (measured as the number of attacks or casualties), the greater the likelihood of overreporting. Importantly, because the recorded terror events took place *after* the elections in question, they could not have affected a respondent's actual turnout behavior. Moreover, because our models account for actual turnout rates, it is not the case that our results are a byproduct of elevated turnout in elections that were followed by terror attacks. Thus, we can be confident that the substantively and statistically important associations we uncover are due to overreporting rather than actual variation in voting behavior.

Study 2: A Natural Experiment in The Netherlands

Determining whether threat-inducing violent events induce turnout overreporting using cross-country data structures, as in Study 1, presents challenges related to identification; detected patterns in overreporting could stem from unobserved factors related to violent events and social desirability. To help overcome these challenges, we follow Finseraas, Jakobsson and Kotsadam (2011)[10] and leverage the murder of Theo van Gogh, a prominent critic of Islam who was shot to death and nearly decapitated by a fundamentalist Muslim on a street in Amsterdam on November 2, 2004. The attack was extremely high profile and sparked public outrage and debate in the Netherlands, while also increasing public concerns about death and perceptions of threat (Boomgaarden and de Vreese 2007; Das et al. 2009)

Importantly for our design, the murder took place in the middle of Dutch fieldwork for Round 2 of the European Social Survey (ESS),[11] which ran from September 11, 2004 through February 19, 2005. A total of 1,075 respondents were interviewed before the day of the van Gogh murder, and 806 were interviewed on or after the day of the murder. All respondents were interviewed in person and were asked whether they voted in the most recent Dutch general election prior to the murder,

[10] Finseraas, Jakobsson and Kotsadam (2011) examine immigration policy preferences as their outcome variable.

[11] The European Social Survey Systems is housed at City University of London (`https://www.europeansocialsurvey.org/`).

held on May 15, 2002.[12] We code those who reported participating as 1 and reported abstainers as 0. The sequencing of events is as follows:

> May 15, 2002: Dutch general election $\rightarrow$ September 11, 2004: Dutch ESS survey fieldwork starts $\rightarrow$ November 2, 2004: van Gogh murder $\rightarrow$ February 19, 2005: Dutch ESS survey fieldwork ends

The van Gogh murder represents an opportunity to cleanly identify the impact of a violent event on turnout overreporting. Given the attack's extremely high-profile nature and immediate and abundant news coverage, we assume that respondents interviewed after the murder were aware of it, though we eliminate those interviewed on the day of the murder to account for the possibility that the news took a day to spread.[13]

We make three additional identifying assumptions. First, we assume that respondents did not manipulate their interview dates in a way that systematically relates to their propensity to report having voted. Second, we assume that the ESS rolled out their fieldwork in a way that did not select respondents on either side of the murder in a way that correlates with their propensity to report turning out to vote. We find it very unlikely that respondents or survey enumerators deliberately sorted interviews to a particular side of the (unpredictable) murder date in a way that correlated with a respondent's propensity to answer the turnout question one way or another.[14] We also conducted an extensive review of the ESS fieldwork documents and found no evidence to suggest relevant imbalances across the murder threshold. Balance tests reported on pages 6-7 of Appendix C also suggest that respondents interviewed before and after the murder do not meaningfully differ on background covariates.

Our third assumption is that nothing else happened around the time of the murder that could have shaped respondents' propensity to report turning out to vote. This is supported by the fact that the murder is the only such event that took place in Netherlands during the survey period

[12]The wording of the question is: "Some people don't vote nowadays for one reason or another. Did you vote in the last Dutch national election in May, 2002?"

[13]This meant dropping 19 respondents. The murder took place at about 9AM, and all interviews that day began at or after that time. Given that mobile internet connectivity and online news was not yet prominent in 2004, we decline to assume that these individuals had already learned of the attack.

[14]Respondents who took longer for enumerators to contact are, in our sample, unsurprisingly more likely to be interviewed after the murder. This could bias our estimated treatment effect if more elusive respondents are also systematically more or less likely to overreport turnout (Muñoz, Falcó-Gimeno and Hernández 2020, pp. 192-193). As detailed in Appendix C on pages 7-8, the trend in the number of contact attempts with regard to time is smooth at the murder, and our main model is robust to a control for contact attempts.

according to the GTD. We also empirically interrogate the plausibility of this assumption below.

We provide a visualization of reported turnout and respondents' survey interview dates in Figure 4. The ESS surveys began 53 days before the day subsequent to the van Gogh murder (again, we drop those interviewed on the day of the attack) and ended 108 days later. To equalize the bandwidth of observations on either side of the murder threshold, we exclude respondents interviewed more than 53 days after the day following the murder. There is a clear discontinuity at the time of the murder, with reported turnout jumping about six percentage points at the threshold. Under our identifying assumptions, this boost is attributable to overreporting. Further suggesting that such an attribution is accurate, the reported turnout rate for those interviewed just before the murder hovers around the official turnout rate of 76.8 percent[15] in the 2002 election. Shortly after the murder, respondents report turning out at a rate substantially higher than the official figure.

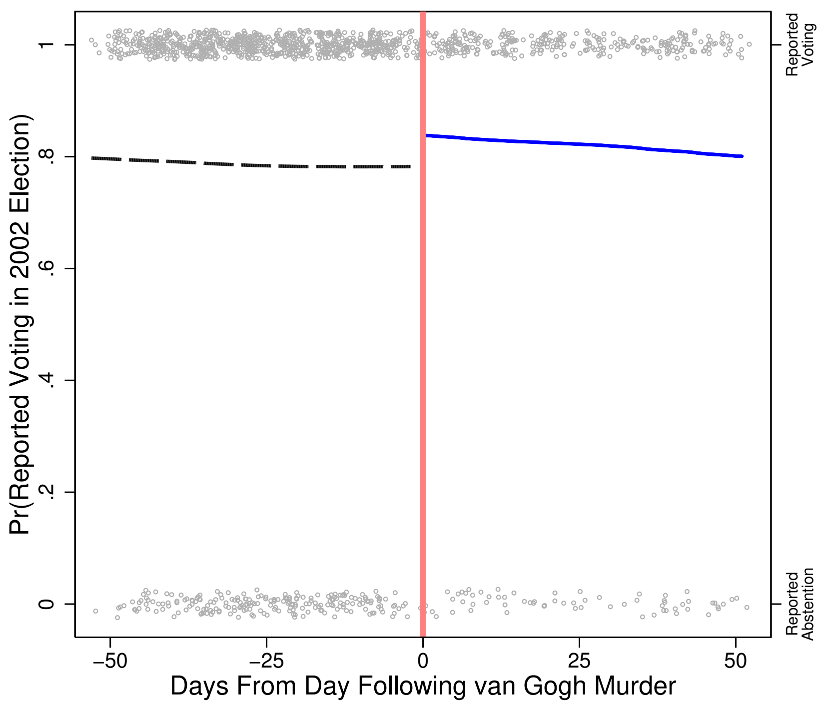

Figure 4: Reported Turnout as a Function of Survey Interview Date

Note: Jittered open circles represent respondents. Those near the top of the figure reported voting in the 2002 general election in the Netherlands. Those near the bottom reported abstaining. The dashed black curve tracks predicted values from a locally smoothed polynomial regression fit to observations recorded before the murder of Theo van Gogh. The solid blue curve tracks predicted values from a locally smoothed polynomial regression fit to observations recorded after the day following the murder.

[15]This is 100 times the number of voters divided by the number of individuals in the voting age population. Data are from the International Institute for Democracy and Electoral Assistance: `https://www.idea.int/`.

Statistical Models and Results

To more rigorously assess the effect of the van Gogh murder on overreporting, we use a regression discontinuity (RD) design. This allows us to estimate the effect of a binary treatment precisely determined by the value of a forcing variable—in this case, the timing of one's survey interview relative to the day after the murder. Survey respondents interviewed after the murder are "forced" from a pre-murder "control" condition into a post-murder "treatment" condition. We adjust for the same covariates employed in Study 1.[16]

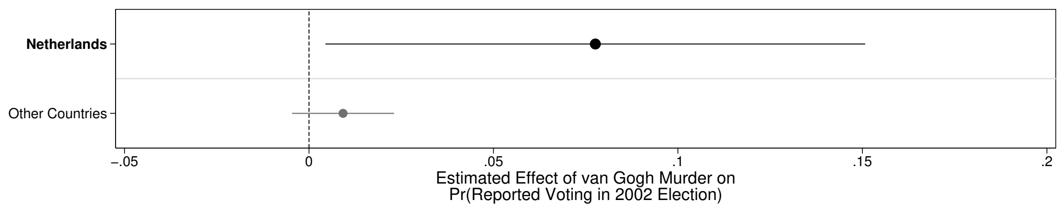

Figure 5: Results of Regression Discontinuity Models

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on the probability of reporting participation in the most recent national election. The effects are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. Both models employ a uniform kernel. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. The number of observations in the analyses is: Netherlands, 1,230; Other Countries, 16,557.

Results of our primary regression discontinuity model are shown in Figure 5. The effect demarcated by the solid black circle in the top row is the estimated average treatment effect (ATE) of being interviewed after the day following the van Gogh murder on the probability of reporting participation in the 2002 general election in the Netherlands. The estimated ATE is just under 0.08, and its associated two-sided p-value of 0.081 indicates that it is unlikely a result of chance. The van Gogh attack boosted reported turnout by nearly eight percentage points.

The bandwidth of observations used in our primary RD model is 53 days, as determined by ESS fieldwork dates. On pages 8-10 of Appendix C, we examine the sensitivity of our results to bandwidth size, finding that the estimated ATE decreases in size and precision as bandwidth decreases. In interrupted survey designs, narrower bandwidths will decrease precision but not necessarily bias (Muñoz, Falcó-Gimeno and Hernández 2020, pp. 195-196). Thus, we are confident that larger

[16]This both improves precision and accounts for slight pre- and post-murder imbalances reported on pages 6-7 of Appendix C. The covariates are gender, age, age squared, education, and income. See Appendix F for question wording.

bandwidths are preferable.

As noted above, we assume that nothing else happened around the time of the murder that could have shaped respondents' propensity to report turning out to vote. Could an international event that happened at the time of the murder (e.g., the 2004 U.S. presidential election) have led to the boost in reported turnout, thus invalidating our third assumption? If there was some event external to the Netherlands that boosted turnout reporting, this should also be evident in comparable countries. To probe the influence of such international events, we estimate an RD model using data from 14 other countries included in Round 2 of the ESS that had surveys in the field at the time of the van Gogh attack.[17] We again estimate an RD model, pooling the data across the 14 countries and including fixed effects for each, and we again eliminate from the sample those interviewed on the day of the murder. As in our analysis of the Dutch data, for each country the bandwidth is the number of days between the beginning of fieldwork and the murder threshold or the threshold and the end of fieldwork, whichever is smaller.

Results of the cross-national RD analysis are shown in the second row of Figure 5. The estimated average treatment effect of being interviewed after the day following the murder in other countries is essentially zero, suggesting no internationally-salient event was responsible for the observed bump in turnout reporting in the Netherlands. This further indicates that the van Gogh attack, and it alone, is responsible for the observed discontinuity in recalled participation. On pages 7-8 of Appendix C, we also show that "placebo" thresholds before and after the murder have no discernible impact on turnout reporting.

The interrupted ESS survey provides an opportune setting to further probe the role of threat in driving overreporting. In studies of war, a key variable is geographic proximity, which captures the fact that threats that are physically closer are more salient (see, e.g., Vasquez 1995). In the context of the van Gogh murder, this logic suggests that survey respondents who were geographically closest to the murder would experience threat most acutely. Consistent with this notion, we established that the assassination impacted Dutch survey respondents but not those further afield in Europe (see Figure 5). This pattern is investigable on a smaller scale within the Netherlands, and in Appendix

[17] These are: Belgium; Czechia; Denmark; Finland; Germany; Luxembourg; Norway; Poland; Slovakia; Slovenia; Spain; Sweden; Switzerland; and the United Kingdom. Estonia and Portugal also had surveys in the field at the time of the murder. We exclude Estonia due to missing data on the income variable, and we exclude Portugal because only 40 survey interviews took place before the attack.

D we show that respondents who live in or nearby Amsterdam, where the attack took place, are much more strongly impacted by the murder than those who do not.

In sum, the findings from this study provide robust evidence that respondents in the 2004 ESS survey conducted in the Netherlands were more likely to report having voted in the 2002 elections if they were interviewed after the murder of Theo van Gogh. Because there were no contemporaneous events that also affected turnout reporting, and because survey respondents were effectively randomly assigned to the pre- and post-murder conditions, observed differences in responses to the turnout question were likely caused by the van Gogh attack and were not influenced by genuine differences in voting behavior. That is, the van Gogh attack increased the likelihood of a respondent claiming to have voted when they did not. Additional analyses of respondents' geographic proximity to the attack suggest that the perceived sense of threat was the driver of the related social desirability bias, as we theorize.

Study 3: A Survey Experiment in India

As a final test of the causal link between violent events and turnout overreporting, we designed and implemented a survey experiment in India meant to gauge turnout overreporting in a bygone national election.[18] Our experiment, fielded in India from April 15-27, 2020, leverages a deadly February 2019 bomb attack on Indian national police. The attack, responsibility for which was claimed by the Islamist militant group Jaish-e-Mohammed, garnered weeks of headline media attention. The sequencing of events is as follows:

> February 14, 2019: fatal terror attack in India → April 11-May 19, 2019: Indian General Election → April 15-27, 2020: Survey fieldwork

The survey was administered online using Qualtrics, and we used Amazon Mechanical Turk (MTurk) to obtain respondents. Numerous studies demonstrate that MTurk-based samples rival those from other sources in quality and usefulness in experimental research (e.g., Berinsky, Huber and Lenz 2012; Clifford, Jewell and Waggoner 2015; Coppock 2019). Regarding the use of MTurk with respondents based in India, Litman, Robinson and Rosenzweig (2015) show that it can be used to gather high quality survey data when pay exceeds US$1.00 per hour, which is well over

[18] The University of Georgia Institutional Review Board (IRB) approved this experiment in Study ID: PROJECT00000815/VERSION00000412.

the minimum wage in any Indian region or job sector. We paid MTurk Workers US$0.35 or its rupee equivalent for completing the survey, which had a median completion time of about about six minutes. Hence, our pay rate corresponds to about US$3.50 per hour for the median participant. Our MTurk task attracted 1,091 unique participants.[19]

Participants read a consent form and answered a variety of demographic questions. Before randomly assigning participants to treatment and control conditions, we presented them with an instructional manipulation check, which had a passage rate of 88.9 percent. (See Appendix F for question wording.) Following Litman, Robinson and Rosenzweig (2015), who show that the quality of MTurk data is significantly higher among Indian respondents who pass such checks, we restrict our sample to those who correctly answered this question.

The treatment condition contained a vignette with information about the terror attack and was designed to prime a threat-related emotional reaction. The control condition presented a similar-length but innocuous news story about the Recycling of Ships Bill, 2019, which was introduced to align India with global practices in ship recycling. The vignettes were accompanied by a photograph of wreckage resulting from the attack or of a container ship floating near shore, respectively. Full content of both conditions is shown in Appendix G of the Supplemental Information.

Immediately after displaying the vignettes, we asked participants whether they voted in the 2019 general elections. We used a preamble that presented abstention as a legitimate choice, coupled with "face saving" response options. This construction has been shown to reduce turnout overreporting due to social desirability across countries (e.g., Belli, Moore and VanHoewyk 2006; Morin-Chassé et al. 2017) and sets the bar high for detecting treatment effects on vote overreporting. The wording was:

> In the last Indian general election, held during April-May, 2019, about 1 in 3 people were not able to vote. People are sometimes unable to vote because they are not registered, they are sick, or they do not have time. Which of the following statements best describes you?
>
> - I did not vote in the 2019 Indian general election.
> - I thought about voting in the 2019 Indian general election but did not.
> - I usually vote but didn't in the 2019 Indian general election.
> - I am sure I voted in the 2019 Indian general election.
> - Don't know

[19] Of these, 32 respondents dropped out before being assigned to a treatment condition. Per IRB guidelines, we also removed 21 respondents who requested their data be erased after being informed at the end of the survey that participants were randomly shown different news stories. Additionally, we dropped two respondents who do not live in India. Finally, we dropped 345 respondents who were assigned to a treatment group used in an unrelated study.

- Prefer not to say

We code those who selected the "I am sure I voted" option as reported voters (1); those who selected the initial three response options as reported abstainers (0); and those who selected "Don't know" or "Prefer not to say" as missing. The question construction appears to have helped mitigate overreporting. In a pilot survey, we asked respondents whether they turned out to vote using a question without the preamble and two simple response options: "I voted" and "I did not vote." With this construction, the reported turnout rate was an untenable 94.6 percent.[20] With the mitigating construction, 75.7 percent of participants claimed to have voted, which is closer to the official rate of 68.8 percent.[21]

At the end of the survey, we asked a treatment-relevant factual manipulation check (FMC-TR) (Kane and Barabas 2019), which inquired about information provided in the experimental conditions. (See Appendix F for question wording.) The correct answer was selected by 72.1 percent of those in the terror group and 70.9 percent of those in the shipping group.[22]

Statistical Models and Results

We estimate both intention to treat (ITT) and complier average causal effects (CACEs). ITT effects are conservative in that they assess the impact of the experimental treatment on the outcome variable, regardless of whether or not a respondent who was assigned to receive the treatment was effectively manipulated. We calculate the ITT effect by regressing reported turnout on a dummy capturing assignment to the terror group.

We expect that the sense of threat induced by violent attacks heightens social desirability, and the CACE helps us investigate this threat mechanism. As compared to those in the control condition, respondents in the treatment condition who were effectively reminded of the terror attack should most strongly experience a threat-related emotional reaction. We assume these are the treated participants who passed the the FMC-TR. At the same time, treated participants who were not effectively manipulated, who we assume are those in the treatment group who failed the FMC-TR, will experience a relatively weak reaction to treatment. The CACE provides the average

[20]After restricting the sample as described in footnote 19, the sample size used to calculate this percentage is 56.

[21]The official turnout rate is 100 times the number of voters divided by the number of individuals in the voting age population. Data are from the International Institute for Democracy and Electoral Assistance: `https://www.idea.int/`.

[22]These proportions are not statistically different from one another. Participants were not able to revisit the treatment vignette after seeing the FMC-TR question.

treatment effect only among those in the treatment group who passed the FMC-TR. We estimate the CACE with two-stage least squares regressions of reported turnout on passage of the FMC-TR, using randomization into the terror group as an instrument for FMC-TR passage (Imbens and Angrist 1994). To increase precision, we adjust for the same demographic covariates employed in Studies 1 and 2 in the estimation of the ITT effect and the CACE.[23]

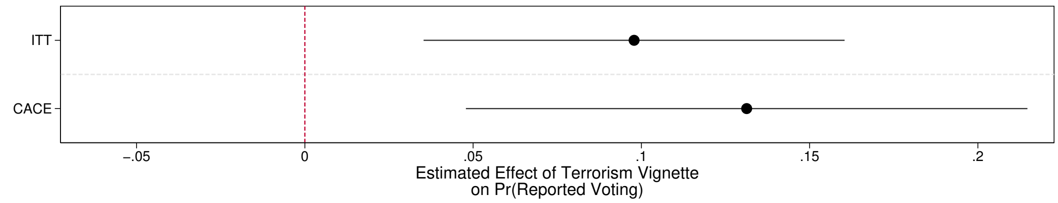

Figure 6: Estimated Effects of the Terror Vignette on the Probability of Reporting Past Turnout

Note: Point estimates represent the intention to treat (ITT) effect or complier average causal effect (CACE). Horizontal lines indicate 90 percent confidence intervals. The number of observations in the underlying models is 499.

The results shown in Figure 6 indicate that being exposed to the terror vignette increases the probability of claiming to have voted by about ten percentage points, relative to those in the shipping condition (two-sided p-value = 0.010). For those who were effectively manipulated by the vignette, the increase is about 13 percentage points (two-sided p-value = 0.010). This indicates that the effect was biggest for those who were manipulated by the terror treatment and thus more likely to have experienced a threat-based emotional reaction. The results of the experiment thus lend additional evidence to our expectation that threatening violent events increase the likelihood that an individual will incorrectly report having participated in bygone election because of threat-induced social desirability bias.

Exploring the Theoretical Mechanism

We argue that societal threat from violent events puts pressure on individuals to *signal* that they behave in ways that are supportive of their society, such as saying that they have voted even when they have not in fact done so. On pages 12-14 and 17-20 of Appendix E of the Supplemental Information, we discuss and empirically interrogate two alternate but theoretically plausible causal

[23]These are: a binary female/male gender variable (no respondents reported an "other" gender); age in years and its square; education on a nine-point ordinal scale; and income on a 13-point ordinal scale. See Appendix F for question wording. While these controls may also help adjust for any covariate imbalances across groups, we are less concerned about such "unlucky draws," as the statistical significance of the estimated effects already takes into account the possibility that the groups will be uneven due to chance after randomization (Mutz and Pemantle 2015).

pathways that could produce an observationally equivalent relationship between violent events and turnout overreporting: increased intentions to vote and in-group bonding. We also investigate whether fatal terror attacks associate more strongly with subsequent increases in turnout overreporting among individuals thought to be particularly prone to social desirability bias and cause increases in measures of social desirability bias beyond turnout overreporting. As discussed on pages 14-17 of Appendix E, we find much empirical support for the latter associations but little for the alternate causal pathways.

Conclusion, Implications, and Practical Advice for Researchers

While the literature has done much to identify the individual-level forerunners and consequences of social desirability bias in survey responses, we contend that its conclusions are incomplete because this bias is not driven only by personal attributes, but is also impacted by events external to the individual. In this article demonstrate that fatal terror attacks exacerbate social desirability bias in survey reports of past voting. One key substantive implication of this is that any findings that terror attacks or wars increase subsequent levels of political participation based on respondent-reported behavior should be viewed with caution. At least some of this increased participation may well be more apparent than real, as survey respondents feel pressured to provide society-supporting responses whether or not they actually engaged in specified activities. More broadly, evidence presented on pages 14-17 of Appendix E suggests that threat-induced social desirability bias has the potential to affect outcomes beyond reported voter turnout. Nevertheless, future research should investigate the scope of survey responses affected by threat-induced social desirability bias.

Our results also indicate that the effect of threat-inducing, violent events on social desirability bias in survey responses is not trivial. This is evident with reference to the impact of the most consistent predictor of social desirability bias in responses to questions about electoral participation, education (Ansolabehere and Hersh 2012, p. 441). Education's effects on overreporting are substantial: comparing reported and validated turnout data using several surveys from the United States, Ansolabehere and Hersh (pp. 448-450) find that a one-unit increase in a five-point education scale corresponds with a seven or eight percentage point increase in the probability of a nonvoter claiming to have voted, an effect that overshadows nearly all other predictors in their models. Hansen and Tyner (forthcoming) reach similar conclusions in an analysis of more recent data. We find the effect of fatal terror attacks is comparably large, at about seven to ten percentage points across our three studies—and education is accounted for in each.

We also add to a growing list of studies showing that the timing of survey fieldwork vis-à-vis salient events affects response patterns (e.g., Eifert, Miguel and Posner 2010; Banducci and Stevens 2015; Michelitch and Utych 2018; Singh and Thornton 2019). The message from these and our study is that the common practice of ignoring interview dates can threaten the quality of inferences. For example, a survey interrupted by a threat-generating event such as a fatal terror attack is likely to yield different responses on questions sensitive to social desirability bias, such as those about voter turnout, depending on whether the respondent was interviewed prior to or after the event. Indeed, the commonness of fatal terror attacks suggests that researchers cannot simply assume that their surveys are not interrupted by such events. In designs that combine surveys from many countries or time periods (including panel surveys), responses may not be comparable without accounting for the presence of threatening events preceding or interrupting a portion of the surveys.

Moreover, survey designers cannot simply rely on preambles and response options meant to "excuse" abstention to prevent the undue influence of social desirability. The ESS turnout question employed in Study 2 contains an abstention-validating preamble, while the turnout question we use in Study 3 has both a preamble and face saving response options. In both cases, we still find evidence of turnout overreporting due to a threatening event.

Given our findings, we see at least two broad strategies researchers can employ to help guard against making faulty inferences from survey data. The first is an exploration akin to the analyses depicted in the Supplemental Information in Figure A4, in which we examine the ESS data for potential time-based discontinuities vis-à-vis reported electoral participation. This could be done for every variable in an analysis, though researchers should pay particular attention to those created from survey responses that are susceptible to social desirability bias. Those interested in causal associations should check whether *the relationships* between their key causal and outcome variables differ before and after the external event. In either case, if demographic covariates are mostly balanced on either side of a discovered discontinuity, this would suggest an external event may be responsible.[24] This should prompt researchers to search relevant databases and news sources

[24] Even responses to demographic questions could be subject to social desirability bias. Thus, for such an investigation, it would be best to rely on demographic factors recorded directly by survey interviewers or matched to public records.

for potentially threatening external events that took place during survey fieldwork. The second approach we recommend reverses the order of operations: researchers should reference event-based data sets to investigate whether some salient event occurred during survey fieldwork, and, upon discovering such an event, check whether it created a time-based discontinuity in relevant survey variables or the relationships between them.

So, what should survey researchers do if they find that an important variable or relationship is affected by a violent event that took place during fieldwork? For those interested in assessing trends in responses to questions that may be sensitive to the event, we recommend not utilizing post-event observations. Similarly, if there is evidence that estimated effects differ on either side of the external event, we again recommend that researchers not use the observations collected after the event in their model estimations. The resulting reduction in sample size and any loss in statistical power that comes with it would, we believe, be preferable to biased inferences.

Our results also have implications for established findings. It may be worthwhile to revisit influential survey-based conclusions to verify that they are not impacted by the timing of survey fieldwork relative to violent events. Furthermore, our study may help explain some of the inconsistencies in findings across published articles on the same subject conducted with different survey data. Some of the surveys employed may have hidden time-based discontinuities that influence authors' conclusions. Looking into this could help harmonize inconsistent findings and thus advance or resolve ongoing debates.

References

Aldrich, John H., John L. Sullivan and Eugene Borgida. 1989. "Foreign Affairs and Issue Voting: Do Presidential Candidates 'Waltz before a Blind Audience?'." *American Political Science Review* 83(1):123–141.

Ansolabehere, Stephen and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20(4):437–459.

Balcells, Laia and Gerard Torrats-Espinosa. 2018. "Using a Natural Experiment to Estimate the Electoral Consequences of Terrorist Attacks." *Proceedings of the National Academy of Sciences* 115(42):10624–10629.

Banducci, Susan and Daniel Stevens. 2015. "Surveys in Context: How Timing in the Electoral Cycle Influences Response Propensity and Satisficing." *Public Opinion Quarterly* 79(S1):214–43.

Belli, Robert F., Sean E. Moore and John VanHoewyk. 2006. "An Experimental Comparison of Question Forms Used to Reduce Vote Overreporting." *Electoral Studies* 25(4):751–59.

Bellows, John and Edward Miguel. 2009. "War and Local Collective Action in Sierra Leone." *Journal of Public Economics* 93(11-12):1144–1157.

Berinsky, Adam J. 2009. *In Time of War: Understanding American Public Opinion from World War II to Iraq*. Chicago: University of Chicago Press.

Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–68.

Bernstein, Robert, Anita Chadha and Robert Montjoy. 2001. "Overreporting Voting: Why It Happens and Why It Matters." *Public Opinion Quarterly* 65(1):22–44.

Blattman, Christopher. 2009. "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review* 103(2):231–247.

Boomgaarden, Hajo G. and Claes H. de Vreese. 2007. "Dramatic Real-World Events and Public Opinion Dynamics: Media Coverage and Its Impact on Public Reactions to an Assassination." *International Journal of Public Opinion Research* 19(3):354–366.

Clifford, Scott, Ryan M. Jewell and Philip D. Waggoner. 2015. "Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?" *Research & Politics* 2(4):1–9.

Coppedge, Michael, et al. 2018. *V-Dem [Country-Year/Country-Date] Dataset V8.* Varieties of Democracy (V-Dem) Project.

Coppock, Alexander. 2019. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* 7(3):613–28.

Coser, Louis A. 1956. *The Functions of Social Conflict.* New York: Free Press.

Crowne, Douglas P. and David Marlowe. 1960. "A New Scale of Social Desirability Independent of Psychopathology." *Journal of Consulting Psychology* 24(4):349–354.

Dahlgaard, Jens Olav, Jonas Hedegaard Hansen, Kasper M. Hansen and Yosef Bhatti. 2019. "Bias in Self-Reported Voting and How It Distorts Turnout Models: Disentangling Nonresponse Bias and Overreporting among Danish Voters." *Political Analysis* 27(4):590–98.

Das, Enny, Brad J. Bushman, Marieke D. Bezemer, Peter Kerkhof and Ivar E. Vermeulen. 2009. "How Terrorism News Reports Increase Prejudice against Outgroups: A Terror Management Account." *Journal of Experimental Social Psychology* 45(3):453–459.

Davis, Darren W. 2007. *Negative Liberty: Public Opinion and the Terrorist Attacks on America.* New York: Russell Sage.

Edwards, Allen. 1953. "The Relationship between the Judged Desirability of a Trait and the Probability That the Trait Will Be Endorsed." *Journal of Applied Psychology* 37(2):90–93.

Eifert, Benn, Edward Miguel and Daniel N. Posner. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science* 54(2):494–510.

Finseraas, Henning, Niklas Jakobsson and Andreas Kotsadam. 2011. "Did the Murder of Theo van Gogh Change Europeans Immigration Policy Preferences?" *Kyklos* 64(3):396–409.

Franzen, Axel and Dominikus Vogl. 2013. "Acquiescence and the Willingness to Pay for Environmental Protection: A Comparison of the Issp, Wvs, and Evs." *Social Science Quarterly* 94(3):637–59.

Górecki, Maciej A. 2011. "Electoral Salience and Vote Overreporting: Another Look at the Problem of Validity in Voter Turnout Studies." *International Journal of Public Opinion Research* 23(4):544–557.

Grönlund, Kimmo and Maija Setälä. 2007. "Political Trust, Satisfaction and Voter Turnout." *Comparative European Politics* 5(4):400–22.

Hansen, Eric R. and Andrew Tyner. forthcoming. "Educational Attainment and Social Norms of Voting." *Political Behavior* .

Hersh, Eitan D. 2013. "Long-term Effect of September 11 on the Political Behavior of Victims' Families and Neighbors." *Proceedings of the National Academy of Sciences* 110(52):20959–20963.

Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2011. "Sex, Stereotypes, and Security: A Study of the Effects of Terrorist Threat on Assessments of Female Leadership." *Journal of Women, Politics & Policy* 32(3):173–92.

Huddy, Leonie and Nadia Khatib. 2007. "American Patriotism, National Identity, and Political Involvement." *American Journal of Political Science* 51(1):63–77.

Huddy, Leonie, Stanley Feldman and Christopher Weber. 2007. "The Political Consequences of Perceived Threat and Felt Insecurity." *Annals of the American Academy of Political and Social Science* 614(1):131–153.

Hutchison, Mark L. and Doug M. Gibler. 2007. "Political Tolerance and Territorial Threat: A Cross-National Study." *Journal of Politics* 69(1):128–142.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–75.

James, Patrick and Jean-Sebastien Rioux. 1998. "International Crises and Linkage Politics: The Experiences of the United States, 1953–1994." *Political Research Quarterly* 51(3):781–812.

Kam, Cindy D. and Jennifer M. Ramos. 2008. "Joining and Leaving the Rally: Understanding the Surge and Decline in Presidential Approval Following 9/11." *Public Opinion Quarterly* 72(4):619–650.

Kane, John V. and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science* 63(1):234–49.

Karp, Jeffrey A. and Caitlin Milazzo. 2015. "Democratic Scepticism and Political Participation in Europe." *Journal of Elections, Public Opinion and Parties* 25(1):97–110.

Karp, Jeffrey A. and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *Journal of Politics* 67(3):825–840.

Lawless, Jennifer L. 2004. "Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57(3):479–490.

Litman, Leib, Jonathan Robinson and Cheskie Rosenzweig. 2015. "The Relationship between Motivation, Monetary Compensation, and Data Quality among US- and India-Based Workers on Mechanical Turk." *Behavior Research Methods* 47(2):519–28.

Marshall, Monty G. and Keith Jaggers. 2012. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010.".

Michelitch, Kristin and Stephen Utych. 2018. "Electoral Cycle Fluctuations in Partisanship: Global Evidence from Eighty-Six Countries." *Journal of Politics* 80(2):412–27.

Morin-Chassé, Alexandre, Damien Bol, Laura B. Stephenson and Simon Labbé St-Vincent. 2017. "How to Survey About Electoral Turnout? The Efficacy of the Face-Saving Response Items in 19 Different Contexts." *Political Science Research and Methods* 5(3):575–84.

Muñoz, Jordi, Albert Falcó-Gimeno and Enrique Hernández. 2020. "Unexpected Event During Survey Design: Promise and Pitfalls for Causal Inference." *Political Analysis* 28(2):186–206.

Mutz, Diana C. and Robin Pemantle. 2015. "Standards for Experimental Research: Encouraging a Better Understanding of Experimental Methods." *Journal of Experimental Political Science* 2(2):192–215.

Norpoth, Helmut and Andrew H. Sidman. 2007. "Mission Accomplished: The Wartime Election of 2004." *Political Behavior* 29(2):175–195.

Oneal, John R. and Jaroslav Tir. 2006. "Does the Diversionary Use of Force Threaten the Democratic Peace? Assessing the Effect of Economic Growth on Interstate Conflict, 1921-2001." *International Studies Quarterly* 50(4):755–779.

Persson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Cambridge: MIT Press.

Rahn, Wendy M., John Brehm and Neil Carlson. 1999. National Elections as Institutions for Generating Social Capital. In *Civic Engagement in American Democracy*, ed. Theda Skocpol and Morris P. Fiorina. Washington, DC: Brookings Institution Press pp. 111–160.

Schildkraut, Deborah J. 2007. "Defining American Identity in The Twentyfirst Century: How Much There Is There?" *American Journal of Political Science* 69(3):597–615.

Schildkraut, Deborah and Peter Furia. 2003. Patriotism. In *The Encyclopedia of Community*, ed. Karen Christensen and David Levinson. Thousand Oaks, CA: Sage.

Sigelman, Lee. 1982. "The Nonvoting Voter in Voting Research." *American Journal of Political Science* 26(1):47–56.

Simmel, Georg. 1898. "The Persistence of Social Groups. II." *American Journal of Sociology* 3(6):829–836.

Singh, Shane P. and Jaroslav Tir. 2018. "Partisanship, Militarized International Conflict, and Electoral Support for the Incumbent." *Political Research Quarterly* 71(1):172–183.

Singh, Shane P. and Jaroslav Tir. 2019. "The Effects of Militarized Interstate Disputes on Incumbent Voting Across Genders." *Political Behavior* 41:975–999.

Singh, Shane P. and Judd R. Thornton. 2019. "Elections Activate Partisanship across Countries." *American Political Science Review* 113(1):248–53.

Smets, Kaat and Carolien van Ham. 2013. "The Embarrassment of Riches? A Meta-Analysis of Individual-Level Research on Voter Turnout." *Electoral Studies* 32(2):344–259.

Tir, Jaroslav and Shane P. Singh. 2013. "Is it the Economy or Foreign Policy, Stupid? The Impact of Foreign Crises on Leader Support." *Comparative Politics* 46(1):83–101.

Tir, Jaroslav and Shane P. Singh. 2015. "Get off My Lawn: Territorial Civil Wars and Subsequent Social Intolerance in the Public." *Journal of Peace Research* 52(4):478–491.

Vasquez, John. 1995. "Why Do Neighbors Fight?: Proximity, Interaction, and Territoriality." *Journal of Peace Research* 32(3):277–293.

Williams, Laron K., David J. Brulé and Michael Koch. 2010. "War Voting: Interstate Disputes, the Economy, and Electoral Outcomes." *Conflict Management and Peace Science* 27(5):442–460.

Wright, Matthew, Jack Citrin and Jonathan Wand. 2012. "Alternative Measures of American National Identity: Implications For The Civic Ethnic Distinction." *Political Psychology* 33(4):469–482.

SUPPLEMENTAL INFORMATION

Threat-Inducing Violent Events Exacerbate Social Desirability Bias in Survey Responses

Shane P. Singh
University of Georgia

Jaroslav Tir
University of Colorado Boulder

Contents

A Study 1: Numerical Results

This appendix contains numerical results of the models used in the production of Figures 1, 2, and 3.

Table A1: Reported Voting Across Countries With and Without Recent Fatal Terror Attacks, Multilevel Logit Estimations

	Most Recent Election	**Previous Election**
Variable	Coefficient (Std. Err.)	Coefficient (Std. Err.)
	Fixed Components	
Surveyed After GTD Event	0.593	0.619
	(0.032)	(0.041)
Female	-0.049	0.006
	(0.016)	(0.022)
Education	0.262	0.272
	(0.008)	(0.010)
Income	0.130	0.103
	(0.006)	(0.008)
Age	0.767	1.662
	(0.026)	(0.035)
Age^2	-0.056	-0.129
	(0.003)	(0.004)
Days Until Interview	-0.001	-0.002
	(<0.001)	(<0.001)
Actual Turnout	0.029	0.028
	(0.001)	(0.001)
Democratic Development	-0.001	0.052
	(0.004)	(0.004)
intercept	-3.409	-6.215
	(0.078)	(0.108)
	Random Components	
var(intercept)	0.227	0.206
	(0.009)	(0.011)
number of observations	117,495	70,320
number of surveys	103	68
log-likelihood	-48182.103	-28198.534

Note: Results are from multilevel logit models with a survey-level random intercept. The observations are survey respondents, who are clustered in election surveys. The dependent variable in the **Most Recent Election** column is reported participation in the most recent national election prior to their survey interview date. The dependent variable in the **Previous Election** column is reported participation in a national election held before the most recent national election prior to their survey interview date. The key independent variable, *Surveyed After GTD Event*, is a binary measure that differentiates respondents in countries that experienced at least one fatal terror event in the time between the most recent election and their survey interview from respondents whose post-election interviews took place before any terror attacks. *Female* is a binary variable coded 1 for reported females and 0 for reported males. *Age* is measured in tens of years. *Education* is measured on a five-point ordinal scale. *Income* is measured on a five-point ordinal scale. *Days Until Interview* is the number of days between the most recent election and a respondent's survey interview. *Actual Turnout* is measured as 100 times the number of voters divided by the number of individuals in the voting age population. *Democratic Development* is measured with the Polity and Freedom House indices and has a theoretical range of -10 to 10. Information on terror attacks is from the Global Terrorism Database. Survey data are from the Comparative Study of Electoral Systems. Included surveys are listed in Table 1 of the main text. Results are shown graphically in Figure 1 of the main text.

Table A2: Reported Voting Across Countries Vis-à-Vis the Number of Recent Fatal Terror Attacks

	Most Recent Election	**Previous Election**
Variable	Coefficient (Std. Err.)	Coefficient (Std. Err.)
	Fixed Components	
Number of GTD Events Before Survey	0.345	0.235
	(0.019)	(0.018)
Female	-0.053	0.005
	(0.017)	(0.022)
Education	0.279	0.272
	(0.009)	(0.010)
Income	0.124	0.101
	(0.007)	(0.008)
Age	0.777	1.671
	(0.026)	(0.035)
Age^2	-0.057	-0.129
	(0.003)	(0.004)
Days Until Interview	-0.001	-0.001
	(<0.001)	(<0.001)
Actual Turnout	0.030	0.028
	(0.001)	(0.001)
Democratic Development	0.006	0.057
	(0.004)	(0.004)
intercept	-3.562	-6.267
	(0.092)	(0.109)
	Random Components	
var(intercept)	0.234	0.203
	(0.009)	(0.010)
number of observations	117,495	70,320
number of surveys	103	68
log-likelihood	-48164.899	-28205.715

Note: Results are from multilevel logit models with a survey-level random intercept. The observations are survey respondents, who are clustered in election surveys. The dependent variable in the **Most Recent Election** column is reported participation in the most recent national election prior to their survey interview date. The dependent variable in the **Previous Election** column is reported participation in a national election held before the most recent national election prior to their survey interview date. The key independent variable, *Number of GTD Events Before Survey*, is the logged count of fatal terror events that took place between a the most recent election and a respondent's survey interview in his or her country of residence. To account for zeros, we added one to all observations before taking logarithms. *Female* is a binary variable coded 1 for reported females and 0 for reported males. *Age* is measured in tens of years. *Education* is measured on a five-point ordinal scale. *Income* is measured on a five-point ordinal scale. *Days Until Interview* is the number of days between the most recent election and a respondent's survey interview. *Actual Turnout* is measured as 100 times the number of voters divided by the number of individuals in the voting age population. *Democratic Development* is measured with the Polity and Freedom House indices and has a theoretical range of -10 to 10. Information on terror attacks is from the Global Terrorism Database. Survey data are from the Comparative Study of Electoral Systems. Included surveys are listed in Table 1 of the main text. Results are shown graphically in Figure 2 of the main text.

Table A3: Reported Voting Across Countries Vis-à-Vis the Number Killed in Recent Fatal Terror Attacks

	Most Recent Election	**Previous Election**
Variable	Coefficient (Std. Err.)	Coefficient (Std. Err.)
	Fixed Components	
Number Killed in GTD Events Before Survey	0.270 (0.014)	0.188 (0.014)
Female	-0.048 (0.017)	0.005 (0.022)
Education	0.264 (0.009)	0.272 (0.010)
Income	0.128 (0.006)	0.101 (0.008)
Age	0.766 (0.026)	1.671 (0.035)
Age^2	-0.056 (0.003)	-0.129 (0.004)
Days Until Interview	-0.001 (<0.001)	-0.001 (<0.001)
Actual Turnout	0.029 (0.001)	0.028 (0.001)
Democratic Development	0.001 (0.004)	0.056 (0.004)
intercept	-3.434 (0.080)	-6.266 (0.109)
	Random Components	
var(intercept)	0.221 (0.009)	0.202 (0.010)
number of observations	117,495	70,320
number of surveys	103	68
log-likelihood	-48171.432	-28201.993

Note: Results are from multilevel logit models with a survey-level random intercept. The observations are survey respondents, who are clustered in election surveys. The dependent variable in the **Most Recent Election** column is reported participation in the most recent national election prior to their survey interview date. The dependent variable in the **Previous Election** column is reported participation in a national election held before the most recent national election prior to their survey interview date. The key independent variable, *Number Killed in GTD Events Before Survey*, is the logged count of deaths in fatal terror events that took place between a the most recent election and a respondent's survey interview in his or her country of residence. To account for zeros, we added one to all observations before taking logarithms. *Female* is a binary variable coded 1 for reported females and 0 for reported males. *Age* is measured in tens of years. *Education* is measured on a five-point ordinal scale. *Income* is measured on a five-point ordinal scale. *Days Until Interview* is the number of days between the most recent election and a respondent's survey interview. *Actual Turnout* is measured as 100 times the number of voters divided by the number of individuals in the voting age population. *Democratic Development* is measured with the Polity and Freedom House indices and has a theoretical range of -10 to 10. Information on terror attacks is from the Global Terrorism Database. Survey data are from the Comparative Study of Electoral Systems. Included surveys are listed in Table 1 of the main text. Results are shown graphically in Figure 3 of the main text.

B Study 1: Mean Reported Turnout in CSES Surveys

Figure A1: Mean Reported Turnout in CSES Surveys

Note: The most recent national elections prior to the beginning of CSES survey fieldwork are indicated on the vertical axis. The solid circles represent the proportion of respondents who reported participating in these elections. The open circles represent the proportion of respondents who reported participating in a national election held before the most recent national election. Labels that appear in bold indicate that the country experienced at least one fatal terror attack between the most recent election and the final CSES interview date. The CSES surveys are also listed in Table 1 of the main text.

C Study 2: Tests of Assumptions and Model Sensitivity

As discussed in the presentation of Study 2 in the main text, in attributing differences in reported turnout to the Theo van Gogh murder, we assume that respondents did not manipulate their interview dates in a way that systematically relates to their propensity to report having voted. We also assume that the European Social Survey (ESS) rolled out fieldwork in a way that did not select respondents on either side of the murder in a way that correlates with their propensity to report turning out to vote.

These identifying assumptions suggest that pre- and post-murder respondents will not systematically differ on demographic factors. To probe this, we search for pre- and post-murder imbalances in demographic variables that are related to turnout but are less likely to themselves be affected by the van Gogh murder. These are the same covariates employed in Studies 1 and 3: gender, age, education, and income. Gender is coded 1 for those identifying as female and 0 for males; age is measure in years; education is measured with five-point ordinal scale; and income is measured with a 12-point ordinal scale. We rescale the age, education, and income variables to range from 0 to 1 for comparability. The wording of the survey questions used to measure these variables is provided in Appendix F.

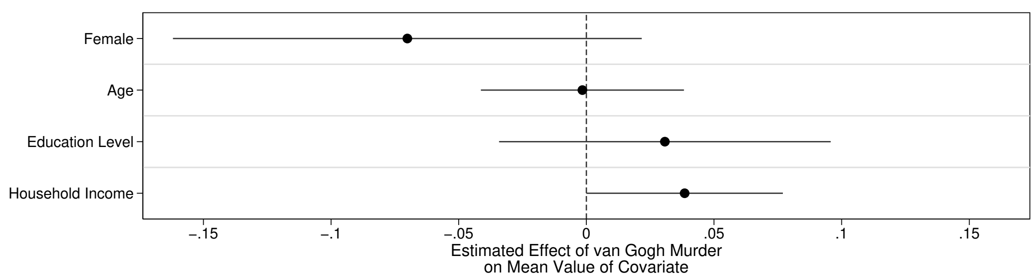

Figure A2: Balance of Demographic Covariates

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on the mean value of the covariates. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. Each comparison includes respondents interviewed up to 53 days before or after the day following the van Gogh murder. Female is coded 1 for those identifying as women and 0 for those identifying as men. The age, education, and income variables are rescaled to range from 0 to 1 to facilitate comparison. On their original metrics, age is measured in years, education is a five-point ordinal scale, and income is a 12-point ordinal scale. The number of observations in the analyses is: gender, 1,479; age, 1,478; education, 1,479; income, 1,271. Data are from Round 2 of the European Social Survey.

Results of the balance tests, which are shown in Figure A2, indicate that the demographic covariates are largely balanced across those interviewed before and after the murder. There is some evidence that women are less likely to be interviewed post-murder, and those interviewed after the murder have, on average, slightly higher household incomes. Because these differences are small or statistically uncertain, we are confident in our identifying assumptions. Nevertheless, we adjust for these covariates in each of the models reported here and in the main text of Study 2 to improve precision and account for the slight imbalances reported in the figure.

Also, as Muñoz, Falcó-Gimeno, and Hernández (2020, pp. 192-193) point out, harder-to-reach subjects are more likely to be interviewed later in the fieldwork period. This would be a problem for our research design if the number of contact attempts for a respondent is also correlated with their propensity to overreport turnout. If this were the case, the number of contact attempts would operate as a confounder.

Reassuringly, there is no statistically discernible relationship between reported turnout and the number of contact attempts in our sample.[1] More importantly, there is no imbalance in the number of contact attempts at the murder threshold.[2] Moreover, our controls for gender, age, education, and income should help account for the characteristics that make respondents easier or harder to contact. Nevertheless, we reestimate the main model of Study 2 with the inclusion of a control for the number of contact attempts, in addition to the usual adjustments for gender, age, age squared, education, and income. As shown in Figure A3, adding this control has almost no impact on the estimated effect of the van Gogh murder on the probability of reporting having turned out to vote.

We next investigate the possibility that, in choosing the van Gogh murder as our discriminating event, we happened to land upon one of several meaningful discontinuities. To probe whether this is the case, we estimate six regression discontinuity (RD) models using fake “placebo” thresholds that are far away from the true threshold and thus unlikely to pick up any of its influence. We use placebos of 15, 20, and 25 days above and below the true threshold, which we take to be the day following the murder. If there are in fact many consequential discontinuities, then turnout reporting should also jump or drop off at one or more of the placebo thresholds.

[1] In a linear regression of reported turnout on the number of contact attempts, the two-sided p-value equals 0.283.

[2] We estimated a regression discontinuity model akin to those reported in Figure A2, with the number of contact attempts as the outcome variable. The estimated effect of being interviewed after the day after the murder on contact attempts is -0.14, with an associated two-sided p-value of 0.432.

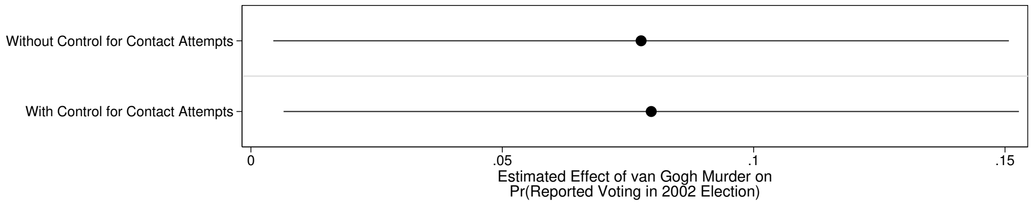

Figure A3: Results of Regression Discontinuity Models With and Without a Control for the Number of Contact Attempts

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on the probability of reporting participation in the most recent national election. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. Both models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). In both models, the ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. The number of observations in the analyses is 1,230.

For analyses that use placebo thresholds situated before the actual murder, the bandwidth is truncated because survey fieldwork started 53 days before the true cutoff. For example, the "20 Days Before" placebo threshold occurs 33 days after fieldwork started. As such, we use a 33-day bandwidth to maintain symmetry. For analyses that use the placebo thresholds located beyond the actual murder date, we are able to maintain equal 53-day bandwidths on either side of the pretend cutoffs, as fieldwork did not end for over three months following the murder. The results of the placebo analyses are shown in Figure A4. None of the RD estimates leveraging the placebo thresholds are nearly as large or precisely estimated as that using the true threshold (see Figure 5 of the main text).

Finally, we investigate the sensitivity of our results to the bandwidth of survey dates used to select respondents. As discussed in the main text, in our comparisons of respondents interviewed before and after the day of the van Gogh murder, we opted for a bandwidth of 53 days based on the timing of survey fieldwork in the Netherlands. Here we reestimate our primary RD model (results of which are depicted in the top panel of Figure 5 in the main text) using a range of bandwidths spanning 5 to 53 days. Results are depicted in Figure A5. The solid black connected line plots estimated average treatment effects from an RD model estimated with each bandwidth depicted on the horizontal axis, and the spikes at the bottom of the plot indicate the number of observations in each corresponding RD model.

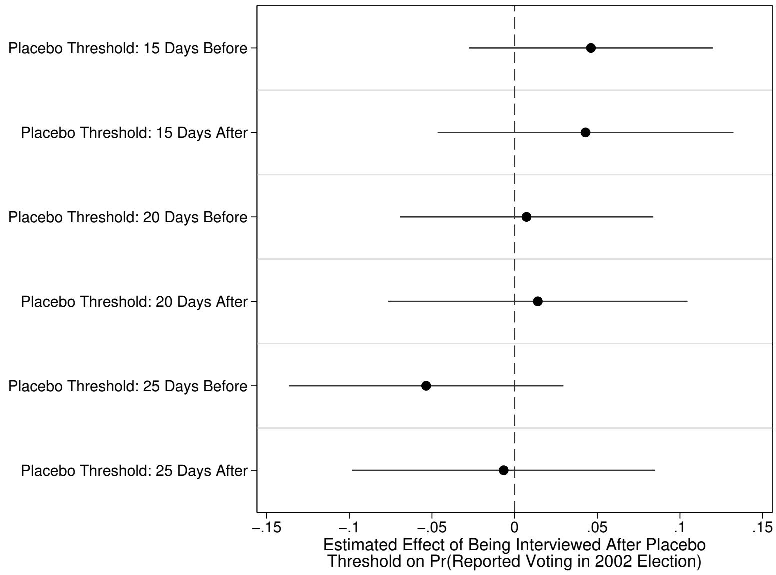

Figure A4: Results of Regression Discontinuity Models with Placebo Thresholds

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after a placebo threshold on the probability of reporting participation in the most recent national election. These are calculated with regression discontinuity models that include a linear trend of the number of days from the indicated threshold on either side of that threshold. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. The placebo thresholds are arbitrarily placed 15, 20, or 25 days above or below the actual day following the murder. Each model includes equivalent bandwidths on either side of the indicated threshold, with a maximum bandwidth size of 53 days, as survey fieldwork began 53 days prior to the threshold and ended 108 days subsequent. The number of observations in the analyses is: 15 Days Before, 1,107; 15 Days After, 1,097; 20 Days Before, 1,042; 20 Days After, 1,001; 25 Days Before, 946; 25 Days After, 988. Data are from Round 2 of the European Social Survey.

The ATE estimates tend to increase in both size and precision as bandwidth increases. At the same time, as bandwidth gets smaller, the ATEs tend to decrease in both size and precision. In typical RD designs, the forcing variable has a theoretical relationship with the outcome variable. In interrupted survey designs like ours, in which the forcing variable is time, Muñoz, Falcó-Gimeno, and Hernández (2020, pp. 195-196) point out that individuals interviewed close to the day of the event are not necessarily more similar to each other than they are to those interviewed further from the event. In this case, narrower bandwidths will decrease precision but not necessarily bias. Thus, we are confident that larger bandwidths are preferable considering the nature of our data and design.

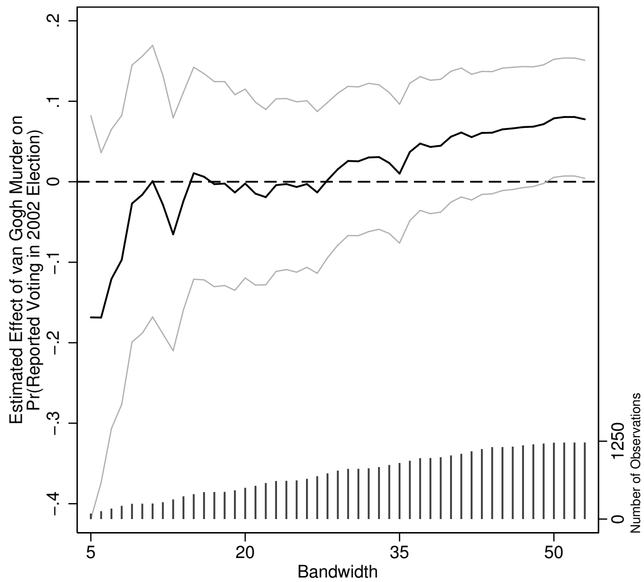

Figure A5: Regression Discontinuity Results with Varying Bandwidth Size

Note: The connected solid black line represents the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on the probability of reporting participation in the most recent national election. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. The connected grey lines indicate corresponding 90 percent confidence intervals, which are calculated with robust standard errors. The horizontal axis tracks the bandwidth of observations associated with each RD model, and the spikes that run along it indicate the number of observations included in each model. Data are from Round 2 of the European Social Survey.

D Study 2: The Conditioning Impact of Geographic Location

Here we test whether respondent location, which itself is very unlikely to be affected by the van Gogh murder,[3] conditions the impact of the van Gogh murder on reported turnout. The ESS codes respondents' region of residence, though it does not give precise information on respondent location. We thus compare respondents living in or closest to Amsterdam (the provinces of North Holland, Flevoland, and Utrecht) to those in the rest of the country.

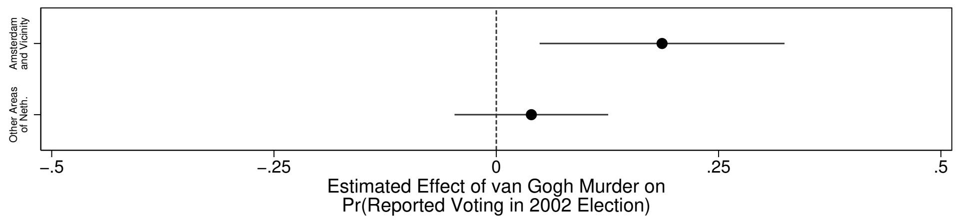

Figure A6: Results of a Regression Discontinuity Model Conditioned by Respondent Location

Note: Point estimates represent conditional average treatment effects (CATEs) of being interviewed after the day following the murder of Theo van Gogh on the probability of reporting participation in the most recent national election. These are calculated with a regression discontinuity model that includes a linear trend of the number of days from the day after the murder on either side of that day. To calculate CATEs, the linear trend and the indicator for treatment are interacted with the variable capturing respondent location. The model employs a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The CATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. The model has a bandwidth of 53 days. The number of observations is 1,230. Data are from Round 2 of the European Social Survey.

To test for the conditioning effect of location, we modify the regression discontinuity approach described in the main text and illustrated in Figure 5, employing an interaction with location to assess the conditional impact of the van Gogh attack for people near and far away from the scene of the crime. Results are reported in Figure A6. As is clear from the figure, for those living nearer to the site of the attack, the van Gogh murder has a stronger positive impact on the probability of reporting turnout. While only the effect for those living near the attack is statistically significant, the difference between the two point estimates does not quite reach conventional levels of statistical significance (the two sided p-value is 0.14).

[3] Respondents could not dishonestly report their location, given that interviews were conducted in person and location was coded directly by the interviewer. It is also not likely that respondents seeking to move because of the murder would have had time to establish residence in a different part of the country in the short period during which remaining survey fieldwork was conducted.

E Studies 1, 2, and 3: Empirical Explorations of the Theoretical Mechanism

We argue that societal threat from violent events puts pressure on individuals to report behaving in socially desirable ways, such as participating in democratic elections. As noted in the "Exploring the Theoretical Mechanism" section of the main text, there are other theoretically plausible mechanisms that could produce an observationally equivalent relationship between violent events and turnout overreporting. Here, we provide detail about our investigations of increased genuine intentions to vote and in-group bonding as potential mechanisms. We also describe our tests of the relationship between violent events and subsequent increases in measures of social desirability bias aside from reported turnout.

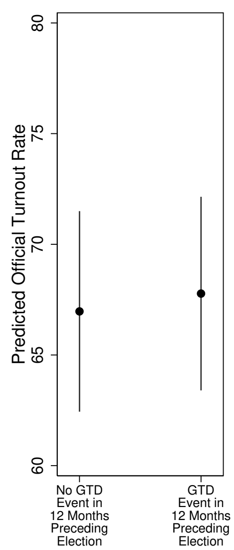

Figure A7: Do Violent Events Increase Actual Electoral Participation? Official Turnout and Pre-Election Fatal Terror Attacks

Note: The independent variable is a binary measure that differentiates elections included in the Comparative Study of Electoral Systems (CSES) in which the host country experienced at least one fatal terror event in 12 months before the election from those that did not. Point estimates represent predicted official turnout rates, as predicted with a linear regression model with country clustered standard errors. Vertical lines indicate 90 percent confidence intervals. Information on terror attacks is from the Global Terrorism Database. The turnout rate is calculated as 100 times the number of voters divided by the number of individuals in the voting age population. Turnout data are primarily from from the Varieties of Democracy Project and supplemented with data from the CSES. There are 170 elections clustered in 55 countries in the analysis.

First, it is possible that violent events boost respondents' legitimate intentions to vote, which then ostensibly drive up their likelihood of reporting past turnout (cf. van Elsas, Miltenburg and

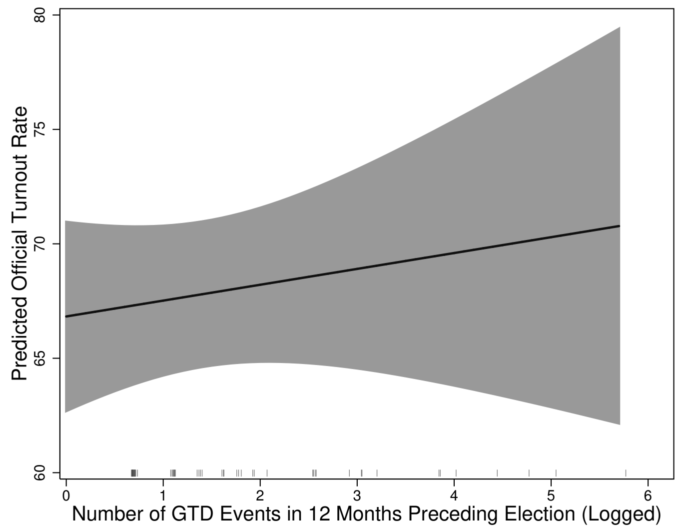

Figure A8: Do Violent Events Increase Actual Electoral Participation? Official Turnout and the Number of Pre-Election Fatal Terror Attacks

Note: The independent variable is a logged count of fatal terror events that that took place in countries included in the Comparative Study of Electoral Systems (CSES) in the 12 months before the election. To account for zeros, we added one to all observations before taking logarithms. The tick marks along the horizontal axis represent elections' locations on the independent variable. These are jittered to reduce overlap. No tick marks are plotted for cases in which there was not at least one fatal terror attack in the 12 months before voting day. The solid line represents predicted official turnout rates, as predicted with a linear regression model with country clustered standard errors. The shaded area indicates the 90 percent confidence interval. Information on terror attacks is from the Global Terrorism Database. The turnout rate is calculated as 100 times the number of voters divided by the number of individuals in the voting age population.Turnout data are primarily from from the Varieties of Democracy Project and supplemented with data from the CSES. There are 170 elections clustered in 55 countries in the analysis.

van der Meer 2016; Balcells and Torrats-Espinosa 2018). If it were the case that violent events increase individuals' intentions to vote, they should also have a positive association with actual voter turnout in subsequent elections. To test this, we gathered official turnout figures from the Varieties of Democracy Project (Coppedge et al. 2018) for the elections included in the CSES (see Study 1). We also gathered data on all fatal terror attacks that took place in the 12 months preceding each election from the Global Terrorism Database (also used in Study 1).

We use linear regression to estimate the relationship between actual turnout and pre-election GTD events. As shown in Figures A7, A8, and A9, actual turnout does not have a substantive or statistically significant relationship with the existence of a pre-election GTD event, the number of pre-election GTD events, or the number of fatalities in these events. The two-sided p-values associated with the coefficient on the GTD variable in each of the three analyses are, respectively, 0.820, 0.538, and 0.545. Results are similar when we use three- or six-month pre-election windows.

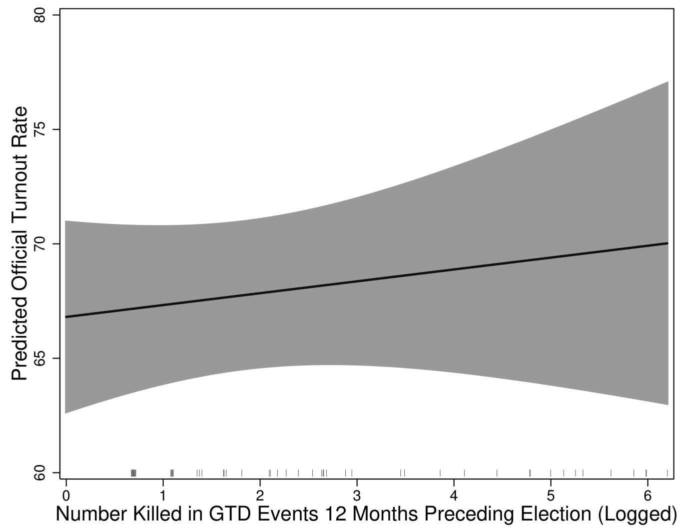

Figure A9: Do Violent Events Increase Actual Electoral Participation? Official Turnout and the Number Killed in Pre-Election Fatal Terror Attacks

Note: The independent variable is a logged count of deaths in fatal terror that that took place in countries included in the Comparative Study of Electoral Systems (CSES) in the 12 months before the election. To account for zeros, we added one to all observations before taking logarithms. The tick marks along the horizontal axis represent elections' locations on the independent variable. These are jittered to reduce overlap. No tick marks are plotted for cases in which there was not at least one death in a terror attack in the 12 months before voting day. The solid line represents predicted official turnout rates, as predicted with a linear regression model with country clustered standard errors. The shaded area indicates the 90 percent confidence interval. Information on terror attacks is from the Global Terrorism Database. The turnout rate is calculated as 100 times the number of voters divided by the number of individuals in the voting age population. Turnout data are primarily from from the Varieties of Democracy Project and supplemented with data from the CSES. There are 170 elections clustered in 55 countries in the analysis.

Second, not finding evidence of fatal terror attacks impacting actual subsequent turnout, we substantiate the role of such attacks in inducing social desirability bias. To this end, we gathered a slate of variables from the ESS that relate to subjective well-being: self assessed general health; life satisfaction; cheerfulness; calmness and relaxation; energy and vigorousness; feeling fresh and rested; and interest in daily life. All variables are coded so that higher values mean a respondent reported a higher level of well-being. (See Appendix F for question wording.) Answers to these types of questions have been strongly and positively linked to survey respondents' scores on social desirability scales (e.g., Kozma and Stones 1988; Caputo 2017). If individuals feel pressured by threatening events to signal social desirability, as we argue, the van Gogh attack, which we leveraged along with the ESS in Study 2, should make them more likely to report a rosy personal situation. In doing so, they signal that they are not a drag on society and standing in the way of the country dealing with larger issues such as the terror threat.

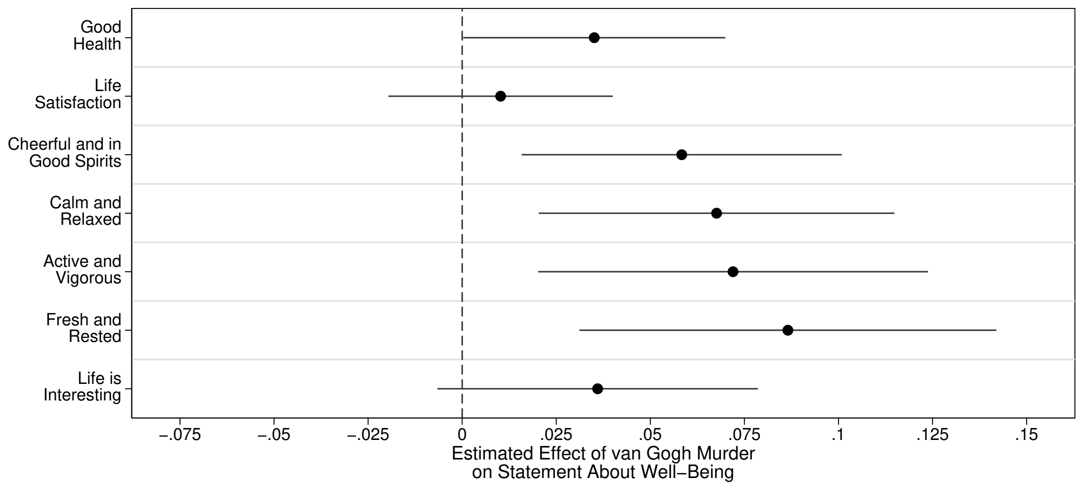

Figure A10: Results of Regression Discontinuity Models with Personal Well-Being Outcome Variables

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on various dimensions of subjective well-being. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. Each model has bandwidth of 53 days. The well-being variables are scaled to range from 0 to 1. The number of observations in the analyses, by outcome variable, is: good health, 1,270; life satisfaction, 1,269; cheerful and in good spirits, 1,270; calm and relaxed, 1,269; active and vigorous, 1,269; fresh and rested, 1,270; life is interesting, 1,270. Data are from Round 2 of the European Social Survey.

Using the same regression discontinuity approach described in Study 2 of the main text and illustrated in Figure 5, we assess the effect of the van Gogh murder on answers to these items. As shown in Figure A10, the murder has a positive and statistically significant impact on five of the seven outcomes. For the items about life satisfaction and having an interesting life, the effects are also positive but rather imprecisely estimated.

The ESS data also allow us to probe whether demographic factors that should associate with social desirability bias in turnout reporting condition the impact of the van Gogh attack. We focus on two demographic factors that are particularly well-suited for this investigation: gender and age. These generally immutable factors should not themselves be affected by the van Gogh murder, and there is evidence that men and older people are more susceptible to overreporting turnout (Sigelman 1982; Ansolabehere and Hersh 2012; Dahlgaard et al. 2019). Moreover, because the ESS is conducted in person, these factors would be more difficult to misreport than other correlates of social desirability bias (and gender is coded directly by interviewers). Thus, identifying their conditioning effects should be relatively insusceptible to post-treatment bias.

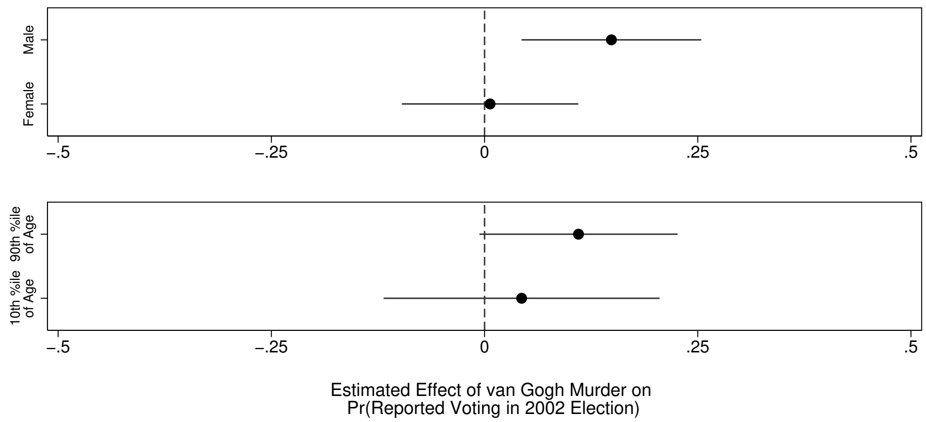

Figure A11: Results of Regression Discontinuity Models with Conditioning Variables

Note: Point estimates represent conditional average treatment effects (CATEs) of being interviewed after the day following the murder of Theo van Gogh on the probability of reporting participation in the most recent national election. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. To calculate CATEs, the linear trend and the indicator for treatment are interacted with the conditioning variables. Both models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The CATEs are adjusted by: a binary female/male gender variable (itself a conditioner in the top model); age in years (itself a conditioner in the bottom model) and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. Each model has a bandwidth of 53 days. The number of observations in each model is 1,230. Data are from Round 2 of the European Social Survey.

To test for the conditioning effects, we modify the regression discontinuity approach described in the main text, employing interactions to assess the impact of the van Gogh attack at different values of both moderators. Results are reported in Figure A11. As is clear from the figure, for men and older people, the van Gogh murder has a stronger positive impact on the probability of reporting turnout. To be sure, the differences between the point estimates reported in each panel of the figure do not themselves reach conventional levels of statistical significance (though that pertaining to gender is close, with a two sided p-value of 0.11), but the conditional effects themselves are only (near) significant for the groups most susceptible to social desirability in turnout reporting, men and older people.

Finally, in the experiment we conducted in India, described in Study 3, we asked a post-treatment question sourced from the influential social desirability scale of Crowne and Marlowe (1960) and found by Greenwald and Satow (1970) to perform well in capturing a respondent's over-all level of "positive"[4] social desirability which, like turning out to vote in democracies, represents a behavior deemed valuable by society. In true or false format, the item inquires whether respondents are always courteous, even to people who are disagreeable. Respondents who answer the item as "true" are considered to be influenced by social desirability bias, while those who selected "false" are not. In our sample, 81.8 percent of respondents chose the "true" option.

To test whether responses to this question are affected by exposure to terror, we use the same approach described in Study 3, estimating both the intention to treat (ITT) and complier average causal effect (CACE) of the terror treatment relative to the shipping condition. Despite the high proportion of courtesy claimers, there is some evidence that the terror treatment has a positive effect. While estimated effects are not significant at conventional levels, they are in the expected direction and substantively nontrivial.

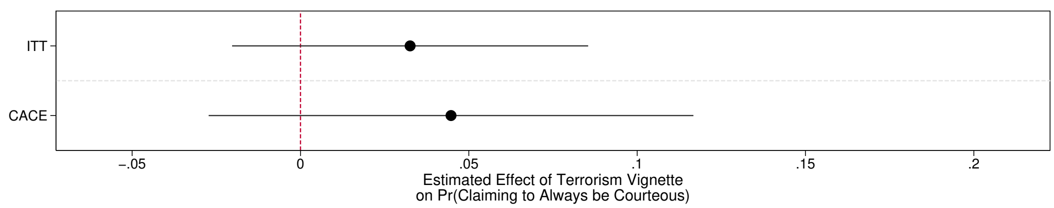

Figure A12: Estimated Effects of the Terror Vignette on the Probability of Reporting Always Being Courteous

Note: Point estimates represent the intention to treat (ITT) effect or complier average causal effect (CACE). The ITT effect gives the difference in the probability of reporting always being courteous for those in the terrorism treatment group relative to those in the shipping condition. The CACE gives the difference in the same probability for those in the terrorism treatment group who were manipulated by treatment relative to those in the shipping condition. The ITT effect and CACE are adjusted by: a binary female/male gender variable; age in years and its square; education on a nine-point ordinal scale; and income on a 13-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals. The number of observations in the underlying models is 572. Data are from an original survey experiment conducted in India.

Third, having amassed additional evidence that threat from violent events induces social desirability, we next consider the possibility that it does so by first triggering in-group bonding. In this scenario, individuals would feel a genuine up-swell of positive attitudes toward their society and country after an attack. While this logic may not explain specifically why one may misreport past voting behavior, the idea would be that in-group bonding generates an uptick in social desirability

[4]Crowne and Marlowe (1960) also include "negatively keyed" items, to which dissenting answers are taken as evidence of social desirability. While not directly related to this study, we included one such item, finding some evidence that those in the terror group were more likely than those in the shipping group to admit to sometimes feeling resentful. We cannot rule out this being a genuine effect, in that an understandable emotional response to a terror attack would be resentment.

bias and, subsequently, attendant vote overreporting.

The logic of in-group bonding dictates that suspicions of those who are perceived not to be the part of the in-group (i.e., the out-group) should increase when the in-group is thought to be under threat (cf. Simmel 1898; Coser 1956). While there are different ways of defining an out-group, immigrants are often perceived as not being the core members of a society. This operational definition of the out-group is sensible in the context of the van Gogh murder, given that he was a victim of an immigrant. It is thus reasonable to expect that immigration attitudes soured in the wake of the assassination if the in-group bonding mechanism has merit.

Using the same regression discontinuity approach described in Study 2, we assess the effect of the van Gogh murder on several variables related to attitudes toward immigrants and immigration available in the ESS. All variables are coded so that higher values mean a respondent is more positive toward immigration or immigrants. (See Appendix F for question wording.) As shown in Figure A13, there is no detectable impact of the attack on attitudes toward allowing more immigrants from different racial or ethnic groups or non-European countries or the belief that immigrants are good for the economy or cultural life. There is a negative impact on the belief that immigrants improve the country, but it is small and not estimated precisely.[5]

If in-group bonding is at work, we should also be able to detect increased nationalistic or patriotic feelings in the wake of a fatal terror attack. Examining this directly in our framework is complicated by the fact that neither the CSES nor the ESS, which we employ in Studies 1 and 2, respectively, contain survey questions that speak directly to such feelings. The ESS, nevertheless, includes questions that should pick up feelings related to increased nationalistic sentiment. Since nationalism concerns favorable feelings toward one's own country, respondents who experience an uptick in nationalism should become more supportive of items that represent or embody their country, such as national institutions. And, since heightened nationalist feelings are simultaneously antagonistic toward representations of foreign influence, international institutions should be evaluated less favorably.

The ESS asks about confidence in national and international institutions. The national institutions are national parliament, the legal system, the police, and political parties; the international

[5]Our insignificant findings mirror those reported by Finseraas, Jakobsson and Kotsadam (2011) in their study of immigration attitudes in the wake of the van Gogh murder.

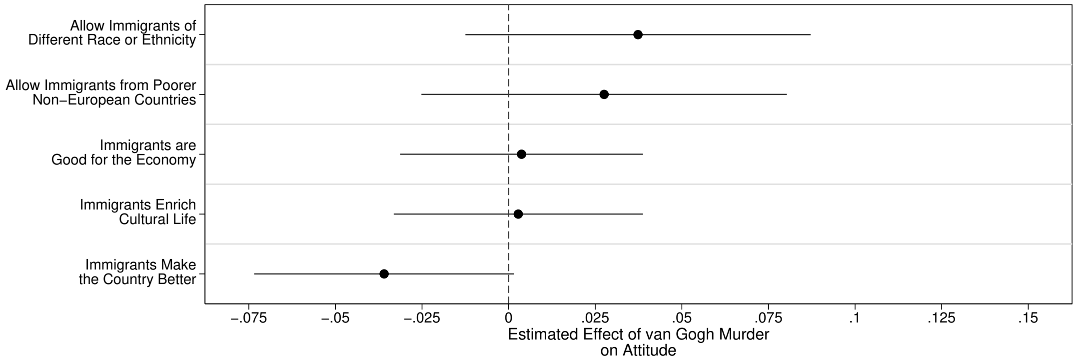

Figure A13: Results of Regression Discontinuity Models with Immigration Attitudes Outcome Variables

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on various attitudes toward immigration and immigrants. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. Each model has bandwidth of 53 days. The immigration attitudes variables are scaled to range from 0 to 1. The number of observations in the analyses, by outcome variable, is: different race or ethnicity, 1,257; poorer, non-European countries, 1,253; immigrants good for economy, 1,247; immigrants enrich cultural life, 1,258; immigrants make country better, 1,257. Data are from Round 2 of the European Social Survey.

institutions are the European Parliament and the United Nations. (See Appendix F for question wording.) We again use the same regression discontinuity approach described in Study 2. Results are shown in Figure A14. As is clear from the figure, we fail to detect a significant effect of the murder on any of the institutional confidence variables. Only concerning attitudes toward national parliament do we see any evidence at all of a nontrivial effect of the van Gogh murder.

To address the CSES and ESS's omission of a direct item about nationalistic or patriotic feelings, we asked a post-treatment question that directly asked respondents whether being an Indian is personally important in the experiment employed in Study 3. (See Appendix F for question wording.) If the in-group bonding mechanism drives the link between violent conflict and social desirability, the terror treatment we employ in Study 3 should alter responses to this question in the positive direction.

To assess the impact of treatment on responses to this question, we use the same approach described in the main text, estimating both the ITT effect and CACE of the terror treatment on nationalism. As shown in Figure A15, we find no evidence that exposure to the terrorism vignette

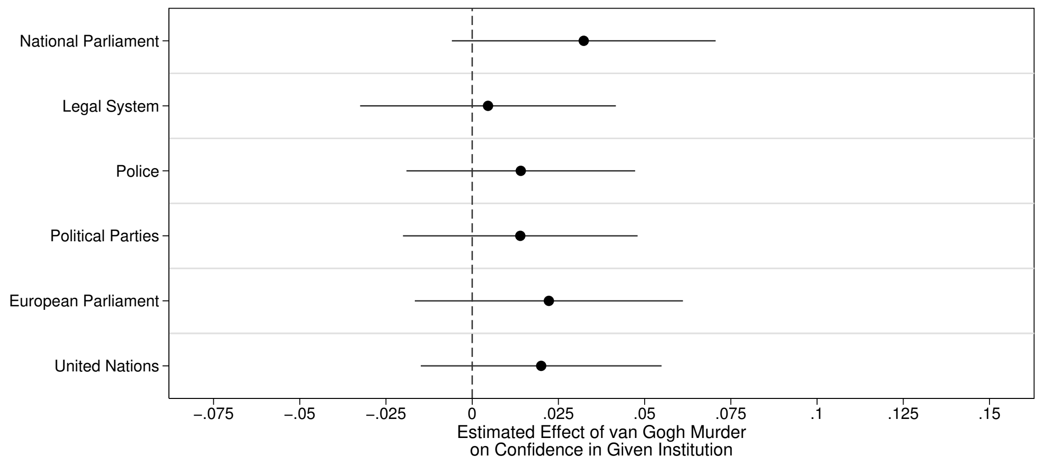

Figure A14: Results of Regression Discontinuity Models with Institutional Confidence Outcome Variables

Note: Point estimates represent the average treatment effect (ATE) of being interviewed after the day following the murder of Theo van Gogh on various dimensions of institutional confidence. These are calculated with regression discontinuity models that include a linear trend of the number of days from the day after the murder on either side of that day. All models employ a uniform kernel, which gives equal weight to all observations that fall within the window of included observations (and no weight to observations outside the window). The ATEs are adjusted by: a binary female/male gender variable; age in years and its square; education on a five-point ordinal scale; and income on a 12-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals, which are calculated with robust standard errors. Each model has bandwidth of 53 days. The institutional confidence variables are scaled to range from 0 to 1. The number of observations in the analyses, by outcome variable, is: national parliament, 1,264; the legal system, 1,253; police, 1,260; political parties, 1,253; European Parliament, 1,125; United Nations, 1,193. Data are from Round 2 of the European Social Survey.

causes an increase in nationalistic sentiment relative to the shipping vignette. The estimated ITT effect and CACE are not only substantively and statistically insignificant, but also run in the counter to the direction we would expect given in-group bonding.

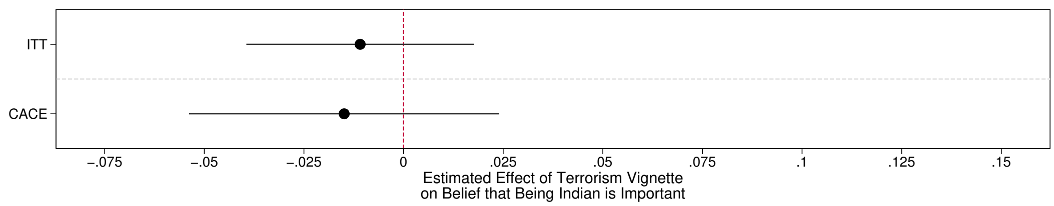

Figure A15: Estimated Effects of the Terror Vignette on Predicted Nationalism

Note: Point estimates represent the intention to treat (ITT) effect or complier average causal effect (CACE). The ITT effect gives the difference in predicted nationalism for those in the terrorism treatment group relative to those in the shipping condition. The CACE gives the difference in the same probability for those in the terrorism treatment group were manipulated by treatment relative to those in the shipping condition. The ITT effect and CACE are adjusted by: a binary female/male gender variable; age in years and its square; education on a nine-point ordinal scale; and income on a 13-point ordinal scale. Horizontal lines indicate 90 percent confidence intervals. The nationalism variable is scaled to range from 0 to 1. The number of observations in the underlying models is 569. Data are from an original survey experiment conducted in India.

F Studies 1, 2, and 3: Survey Questions

Study 1

This section provides the wording of the survey questions from the Comparative Study of Electoral Systems (CSES) used in the construction of the control variables employed in Study 1. Note that wording varies slightly across CSES surveys.

Gender: "Are you male or female?" In some surveys, gender was coded by the interviewer.

Age: "What is your age?" In some surveys, age was determined with respondents' birth dates.

Education: "What is the highest level of education you have completed?" Response options vary across surveys and are harmonized as follows:
1. None (No Education)/Illiterate
2. Primary Education/Lower Secondary Education
3. Higher Secondary Education
4. Post-Secondary (Non-University) Education
5. University Education

Income: "Which of these options represents the total income of your household from all sources before tax including benefits, savings, and so on?" Respondents are grouped into quintiles within surveys.

Study 2

This section provides the wording of the survey questions from Round 2 of the European Social Survey employed in the Netherlands and used in the construction of the covariates employed in Study 2 and the variables used this appendix.

Gender: Gender was coded by the interviewer.

Age: "In what year were you born?"

Education: "What is the highest level of education you have achieved?" Response options vary across surveys and are harmonized as follows:
1. No qualifications
2. CSE grade 2-5/GCSE grades D-G or equivalent
3. CSE grade 1/O-level/GCSE grades A-C or equivalent
4. A-level, AS-level or equivalent
5. Degree/postgraduate qualification or equivalent

Minority Status: Do you belong to a minority ethnic group in the Netherlands?

- Yes
- No

- Don't know

Income: "Using this card, if you add up the income from all sources, which letter describes your household's total net income? If you don't know the exact figure, please give an estimate. Use the part of the card that you know best: weekly, monthly or annual income." Respondents were shown 12 response options. All responses were coded in terms of monthly income.

Health: "How is your health in general?"
- Very good
- Good
- Fair
- Bad
- Very bad

Life Satisfaction: "All things considered, how satisfied are you with your life as a whole nowadays? Please answer where 0 means extremely dissatisfied and 10 means extremely satisfied."

Recent Feelings About Life: "I am going to read out a list of statements about how you may have been feeling recently. For each statement, using this card, I would like you to say how often you have felt like this over the last two weeks." Respondents were shown a card with the options "all of the time," "most of the time," "more than half of the time," "less than half of the time," "some of the time," and "none of the time."
- I have felt cheerful and in good spirits.
- I have felt calm and relaxed.
- I have felt active and vigorous.
- I have woken up feeling fresh and rested.
- My daily life has been filled with things that interest me.

Immigration: To what extent do you think the Netherlands should allow people [of a different race or ethnic group from most Dutch people/from poorer countries outside Europe] to come and live here?
- Allow many to come and live here
- Allow some
- Allow a few
- Allow none

Immigrants: "On a scale of 0-10, would you say [it is generally bad or good for the Dutch economy that people come to live here from other countries/Dutch cultural life is generally undermined or enriched by people coming to live here from other countries/the Netherlands is made a worse or better place to live by people coming to live here from other countries]?"

Institutional Confidence: "On a scale of 0-10, how much confidence you personally have in each of the institutions I read out. 0 means you do not have any confidence in an institution at all, and 10 means you have complete confidence."
- The national parliament
- The legal system
- The police
- Political parties
- The European Parliament
- The United Nations

Study 3

This section provides the wording of the survey questions used in the construction of the covariates and manipulation checks employed in Study 3 and the nationalism variable used in this appendix. We fielded the survey in India from April 15-April 27, 2020.

Gender: "What is your gender?"

- Male
- Female
- Other

Age: "What is your age in years?"

Education: "What is the highest level of education you have successfully completed??"

1. No formal education
2. Incomplete primary school
3. Completed primary school
4. Middle pass
5. 10th pass
6. 11th pass, not completed intermediate
7. 12th pass/Intermediate
8. Undergraduate, still in college
9. B.A. and Higher degrees

Income: "What is your total monthly household income in rupees—putting together the income of all members of the household?"

1. Less than ₹5,000
2. ₹5,000 to ₹9,999
3. ₹10,000 to ₹14,999
4. ₹15,000 to ₹19,999
5. ₹20,000 to ₹24,999
6. ₹25,000 to ₹29,999
7. ₹30,000 to ₹34,999
8. ₹35,000 to ₹39,999
9. ₹40,000 to ₹44,999
10. ₹45,000 to ₹49,999
11. ₹50,000 to ₹54,999
12. ₹55,000 to ₹59,999
13. ₹60,000 or more

Instructional Manipulation Check: "You probably have a favourite colour. But we are more interested in making sure you're doing the survey carefully, so please just select the colour purple here."

- Orange
- Blue
- Green

- Purple
- Red

Treatment-Relevant Factual Manipulation Check: "If you do not know the answer to the following question, it is perfectly acceptable to respond with 'don't know.' What was discussed in the news story that we asked you to read?"

- The coronavirus pandemic
- A terror attack
- Ship recycling
- Panchayat health care delivery
- A major traffic accident
- Ocean cleanup
- Don't know

Courtesy to Others: I am always courteous, even to people who are disagreeable.

- True
- False

Resentfulness: I sometimes feel resentful when I don't get my own way.

- True
- False

Nationalism: "How important is being an Indian to you? Please answer using a score of 0 to 10, where 0 means not at all important and 10 means it is the most important thing in your life."

G Study 3: Vignettes

Here we display the content of the treatment and control conditions to which participants were randomly assigned in Study 3. Those in the terror treatment group saw the following vignette:

Before you continue answering questions, we would like you to please carefully read this short news story:

On February 14, 2019, a convoy of vehicles carrying Central Reserve Police Force (CRPF) personnel was attacked by a suicide bomber in the Pulwama district of Jammu and Kashmir. The attack took place about 20 km from Srinagar and caused the deaths of 40 CRPF personnel. The Islamist militant group Jaish-e-Mohammed has claimed responsibility for the attack. It was the deadliest terror attack on India's state security personnel in Kashmir since 1989.

Please click the arrow to continue.

Those in the shipping control group saw the following vignette:

Before you continue answering questions, we would like you to please carefully read this short news story:

On December 9, 2019, Parliament passed the "Recycling of Ships Bill, 2019." The Bill, upon becoming Act, will regulate the recycling process of ships and the protection of yard workers. Ships to be recycled in India will need to obtain a "Ready for Recycling Certificate" in accordance with international agreements. Existing facilities need to apply for authorisation within 60 days of the commencement of the Act.

Please click the arrow to continue.

References

Ansolabehere, Stephen and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20(4):437–459.

Balcells, Laia and Gerard Torrats-Espinosa. 2018. "Using a Natural Experiment to Estimate the Electoral Consequences of Terrorist Attacks." *Proceedings of the National Academy of Sciences* 115(42):10624–10629.

Caputo, Andrea. 2017. "Social Desirability Bias in Self-Reported Well-Being Measures: Evidence from an Online Survey." *Universitas Psychologica* 16(2):245–255.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Agnes Cornell, Sirianne Dahlum, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Rmer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. 2018. *V-Dem [Country-Year/Country-Date] Dataset V8*. Varieties of Democracy (V-Dem) Project.

Coser, Louis A. 1956. *The Functions of Social Conflict*. New York: Free Press.

Crowne, Douglas P. and David Marlowe. 1960. "A New Scale of Social Desirability Independent of Psychopathology." *Journal of Consulting Psychology* 24(4):349–354.

Dahlgaard, Jens Olav, Jonas Hedegaard Hansen, Kasper M. Hansen and Yosef Bhatti. 2019. "Bias in Self-Reported Voting and How It Distorts Turnout Models: Disentangling Nonresponse Bias and Overreporting among Danish Voters." *Political Analysis* 27(4):590–98.

Finseraas, Henning, Niklas Jakobsson and Andreas Kotsadam. 2011. "Did the Murder of Theo van Gogh Change Europeans Immigration Policy Preferences?" *Kyklos* 64(3):396–409.

Greenwald, Herbert J. and Yoichi Satow. 1970. "A Short Social Desirability Scale." *Psychological Reports* 27(1):131–35.

Kozma, Albert and M. J. Stones. 1988. "Social Desirability in Measures of Subjective Well-Being: Age Comparisons." *Social Indicators Research* 20(1):1–14.

Muñoz, Jordi, Albert Falcó-Gimeno and Enrique Hernández. 2020. "Unexpected Event During Survey Design: Promise and Pitfalls for Causal Inference." *Political Analysis* 28(2):186–206.

Sigelman, Lee. 1982. "The Nonvoting Voter in Voting Research." *American Journal of Political Science* 26(1):47–56.

Simmel, Georg. 1898. "The Persistence of Social Groups. II." *American Journal of Sociology* 3(6):829–836.

van Elsas, Erika J., Emily M. Miltenburg and Tom W. van der Meer. 2016. "If I Recall Correctly. An Event History Analysis of Forgetting and Recollecting Past Voting Behavior." *Journal of Elections, Public Opinion and Parties* 26(3):253–272.