Supporting Information

Why Don't Partisans Sanction Electoral Malpractice?

March 14, 2022

Contents

A Survey Items for Key Measures

Perceptions of electoral integrity (PEI) was measured using the following survey item.

> "Think back on the election that you just read about.
>
> How would you rate the election fairness on this scale from one to five, where one means that the election was conducted unfairly and five means that the election was conducted fairly?"
>
> 1 The election was conducted unfairly
>
> 2
>
> 3
>
> 4
>
> 5 The election was conducted fairly
>
> 6 Don't know/prefer not to answer

Among Danish respondents, the mean PEI score across experimental arms is 2.19 with a standard deviation of 1.55. Among Mexican respondents, the mean PEI score across experimental arms is 2.23 with a standard deviation of 1.48.

Government support was measured using the following survey item.

> "Think back on the election that you just read about. Would you support this government? Please indicate your support for this government on a scale from 0 to 10, where 0 means you dislike this government a great deal, and 10 means you like this government a great deal"
>
> 0 Dislike a great deal
>
> 1
>
> 2
>
> 3
>
> 4
>
> 5
>
> 6
>
> 7
>
> 8

9

10 Like a great deal

11 Don't know

12 Prefer not to answer

Among Danish respondents, the mean support score across experimental arms groups is 2.75 with a standard deviation of 3.18. Among Mexican respondents, the mean support score across experimental arms is 3.24 with a standard deviation of 3.40.

Perceived Legitimacy was measured using the following survey item.

"Think back on the election that you just read about.

To what extent would you think that this government has the right to make binding political decisions?"

1 Not at all

2 To a small extent

3 To some extent

4 To a large extent

5 To a very large extent

6 Don't know

7 Prefer not to answer

Among Danish respondents, the mean legitimacy score across experimental arms is 2.30 with a standard deviation of 1.43. Among Mexican respondents, the mean PEI score across experimental arms is 2.57 with a standard deviation of 1.35.

Compliance is used as a different conceptualization of legitimacy (cf. Levi et al. 2009; Tyler 1997, 2006) and was measured using the following survey item and

"Think back on the election that you just read about.

To what extent would you feel that people have an obligation to comply with this government's decisions even if they go against what they think is right?"

1 Not at all

2 To a small extent

3 To some extent

4 To a large extent

5 To a very large extent

6 Don't know

7 Prefer not to answer

Among Danish respondents, the mean legitimacy score across experimental arms is 2.68 with a standard deviation of 1.34. Among Mexican respondents, the mean PEI score across experimental arms is 2.50 with a standard deviation of 1.31.

As described in the main text, two measures were used to capture **partisanship**. First, respondents rated the major parties (which also featured in the vignettes) on a five-point like-dislike scale. The following item was used pre-treatment.

"We would like to understand your attitudes toward the major political parties. Please rate the following parties on a scale from 1 to 5, where 1 means that you dislike the party a great deal, and 5 means you like the party a great deal."

Scale:

1 Dislike a great deal

2 Dislike somewhat

3 Neither like nor dislike

4 Like somewhat

5 Like a great deal

6 Don't know

7 Prefer not to answer

Parties (Mexico)

- Movimiento de Regeneración Nacional (MORENA)
- Partido Accíon Nacional (PAN)
- Partido Revolucionario Institucional (PRI)

- Partido Revolucionario Democrático (PRD)

Parties (Denmark)

- Socialdemokratiet
- Venstre

Please note that the Danish survey asked respondents to rate the main "prime minister" parties ("statsministerpartier") instead of "major parties" because these two parties are traditionally leading government coalitions and are commonly seen as the only parties competing for the office of prime minister. I matched respondents like-dislike scores to the party mentioned in the vignettes they received, which gives a measure of prior attitudes towards the perpetrators of electoral malpractice. This measures ranges from 1 to 5, with higher values indicating stronger co-partisanship (Denmark: mean = 2.98, SD = 1.26; Mexico: mean = 2.45, SD = 1.42). This measure has two key advantages over simply probing whether respondents identify with a given party. First, this measure captures both positive and negative partisanship; it captures in-group and out-group affection (Abramowitz and Webster 2018). Because a partisan identifier may feel quite positive towards an out-group, simply coding whether people identify with a party disguises important differences in how they may feel towards other parties. Secondly, this measure is suited for cross-country comparisons (Wagner 2020). These advantages notwithstanding, I also used a different measure, relying on two questions: respondents preferred party and whether they identified as partisans. First, respondents were asked the following.

"If elections to the (Cámera de Diputados / Folketinget) were held today, which party or which party's candidate would you vote for?

Parties (Mexico)

1. Movimiento de Regeneración Nacional (MORENA)
2. Partido del Trabajo (PT)
3. Partido Accíon Nacional (PAN)
4. Partido Revolucionario Democrático (PRD)
5. Movimiento Ciudadano (MC)
6. Partido Revolucionario Institucional (PRI)
7. Partido Verde Ecologista (PVEM)
8. Other/None of the above

9. Don't know
10. Prefer not to answer

Parties (Denmark)

1. Socialdemokratiet (A)
2. Radikale Venstre (B)
3. Det Konservative Folkeparti (C)
4. Nye Borgerlige (D)
5. Socialistisk Folkeparti (F)
6. Veganerpartiet (G)
7. Liberal Alliance (I)
8. Kristendemokraterne (K)
9. Dansk Folkeparti (O)
10. Stram Kurs (P)
11. Venstre (V)
12. Enhedslisten (Ø)
13. Alternativet (Å)
14. Other/None of the above
15. Would not vote
16. Not eligible to vote
17. Don't know
18. Prefer not to answer

Subsequently, respondents answering between 1 and 7 (Mexico) or 1 and 13 (Denmark) were asked whether they identify as partisans of that party. The following item was used.

> "Many people consider themselves supporters of a specific party, while others do not feel particularly supportive of a specific party. Do you consider yourself a supporter of [*party from the previous question*]?"

1. Yes

2. No

3. Don't know

4. Prefer not to answer

I then coded respondents as either leaners, co-party identifiers, or opponents of the party that was mentioned in the vignettes they received. In Denmark, there were 2,061 opponents, 186 leaners, and 279 identifiers. In Mexico, there were 2,106 opponents, 160 leaners, and 262 identifiers.

B Survey, Vignettes and Ethics

The two boxes below show translated versions of the experimental vignettes used in the experiment.

Experimental Vignette (Mexico - Spanish Version)

"**Ahora, nos gustaría que imagine un escenario.**

Imagine que las elecciones presidenciales se celebran el próximo mes. El candidato que representa a una coalición liderada por [*PARTY*] es elegido presidente.

Después de la elección, queda claro que [*EXPERIMENTAL CONDITION*]."

Experimental conditions:

Control: hubo más debates presidenciales en la televisión de lo habitual para esta elección.

Treatment #1: Vote-buying: a muchos votantes se les pagó comida y dinero para votar por el candidato de [*PARTY*].

Treatment #2: Voter coercion: muchos votantes se enfrentaron a amenazas de perder el acceso a los programas sociales, como Prospera, si no votaban por el candidato de [*PARTY*].

Treatment #3: Ballot stuffing: se añadieron a las urnas votos falsos para el candidato de [*PARTY*].

Experimental Vignette (Denmark - Danish Version)

"**Nu skal du forestille dig et scenarie.**

Forestil dig, at der var Folketingsvalg næste måned. [*PARTY*] går frem ved valget og danner regering med støtte fra andre partier fra samme fløj.

Efter valget kommer det frem, at [*EXPERIMENTAL CONDITION*]."

Experimental conditions:

Control: der var flere statsministerdueller på TV end normalt der plejer at være.

Treatment #1: Vote-buying: mange vælgere var blevet betalt for at stemme på [*party*].

Treatment #2: Voter coercion: mange vælgere var blevet truet med fratagelse af velfærdsydelser, som f.eks. børnepenge, hvis ikke de stemte på [*PARTY*].

Treatment #3: Ballot stuffing: der blev tilføjet ekstra (falske) stemmer til [*PARTY*] i stemmeboksene.

The parties in Mexico were MORENA, PRI, and PAN. Parties in Denmark were Socialdemokratiet and Venstre.

Ethical Considerations. The experiment manipulates (a) whether electoral malpractice has taken place and (b) which party benefits from this violation. The vignettes provides fictitious information about a potential future scenario. This is, of course, a sensitive topic. Consequently, it is important to openly reflect and discuss the ethical concerns that may arise from such an experiment. The experiment was designed to comply with the "silver standard" for experimental ethics (cf. Teele 2021). This implies that survey participation was voluntary/consensual, but that participants did not know the exact aims of the experiment or whether the information they received was the treatment or the placebo. Giving participants such information (i.e. the "gold standard") would likely tamper with the workings of the experiment (Teele 2021). In this study, the key ethical concern when designing the treatments related to deception. The APSA (2020) principles and guidelines for human subject research notes that "political science researchers should carefully consider any use of deception and the ways in which deception can conflict with participant autonomy" (p. 7).

In this experiment, one may worry that respondents would believe the information about fraudulent electoral behavior from named parties from the "hypothetical future scenarios", and that this could potentially

impact respondents' real-world behavior and have electoral implications.

To limit the possibility this, two measures were taken. These ensure that participant autonomy was respected in line with APSA (2020) guidelines. First, the vignettes prompted the respondents to engage in a hypothetical scenario by highlighting that **"Now, we would like you to imagine a scenario."** Such prompts are oft-used in similar experiment (e.g., Anduiza et al. 2013; Graham and Svolik 2020; Gutiérrez-Romero and Lebas 2020; Reuter and Szakonyi 2021; Svolik 2019; Szakonyi 2021; Tomz and Weeks 2020). Such prompts function to ensure that respondents are made aware that the information they are about to read is hypothetical and does not contain factual information.

Second, to make sure that survey participants were not deceived, the survey included a debriefing as recommended by APSA (2020, p. 8). After completing the survey, respondents read the following debriefing to ensure that respondents were made aware that any mention of electoral malpractice was fictitious.

> "Thank you for participating in the survey!
>
> We want to make it clear that the text you read about elections was fictitious. The information you were given was an important part of our research. Although such information is fictitious, it is valuable for our research. We hope for your understanding.
>
> Thanks again!"

As respondents were prompted to imagine a scenario in the experimental vignette and were thoroughly debriefed to avoid deception, the experiments adhere to APSA (2020) guidelines and were conducted in line with the Declaration of Helsinki. Institutional review boards or committees are not mandatory for survey-based research in the author's country of employment.

However, one may still be concerned that these measures were inadequate, and that participants still changed real-world behavior. On the other hand, studies have shown that survey treatments have very short term effects (e.g., Druckman and Nelson 2003; Gaines et al. 2007; Luskin et al. 2002; Mutz and Reeves 2005; Sniderman 2018). Mutz and Reeves (2005), for example, show that while exposing respondents to uncivil political TV debates did have an immediate effect on trust, such effects does not last very long. By the time of their follow-up interview, there were no differences between treated and non-treated respondents. In other words, it is unlikely that participants in this experiments would still differ on e.g. perceived election fairness just shortly after completing the survey. To limit the potential of possible real-world consequences, the experiments were fielded in after elections and not in near proximity to coming elections. The most recent Mexican General Election was held in July, 2018. The most recent Danish General Election was held in June, 2019. These experiments were fielded in December, 2020, more than half a year before the 2021 legislative election in Mexico. Hence, for these vignettes to have real-world consequences in terms of voting

behavior (i.e. Mexican participants would sanction the party involved in hypothetical malpractice in these vignettes in real elections), they would have to still have an effect on attitudes more than six months after completing the surveys.

In the author's country of employment, institutional review boards or committees are not mandatory for survey-based research in the author's country of employment. However, it is worth noting that two recent studies using very similar vignettes have received IRB approval (Daxecker and Fjelde 2020; Szakonyi 2021). Daxecker and Fjelde (2021) manipulate information about election violence, whereas Szakonyi (2021) prompts respondents to imagine an election where some type of manipulation occurs. These designs are, therefore, very similar to one employed here.

Finally, YouGov collects informed consent from all participants in their panels. These respondents are informed that some surveys will contain experimental manipulation. Further, YouGov's panelists are experienced survey participants who are very good at engaging in surveys and experiments, but also very apt at putting information received in a survey context aside. This is important as APSA (2020) notes that researchers should consider power relations between them and research participants.

C Covariate Balance Across Treatment Conditions

Tables C1 and C2 below show that control and treatment groups are balanced on pre-treatment covariates.

Table C1: Covariate Balance in Danish Sample

	Control			Treatment			
	n	mean	sd	n	mean	sd	Diff
Gender	623	1.49	0.50	1903	1.49	0.50	0.002
Age category	623	6.81	3.58	1903	6.70	3.50	-0.111
Interest in politics	620	3.86	1.05	1888	3.83	1.08	-0.030
Party choice	482	5.03	4.13	1515	5.03	4.12	0.000
Attitude Socialdemokratiet	591	3.31	1.27	1808	3.29	1.32	-0.022
Attitude Venstre	587	2.65	1.15	1798	2.66	1.13	0.010
Economic satisfaction	584	6.19	1.87	1751	6.13	1.96	-0.065

Note: Table shows averages for baseline.
The Diff column is the coefficient of a simple regression of treatment status on the variable.
Stars indicate whether this difference is significant. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Randomization created groups that are balanced on gender (1: female, 2: male), age category (1-15), interest in politics (1-5), party choice (1-14 in Denmark and 1-10 in Mexico), attitudes towards the major parties (1-5) - which are used to create the measure of co-partisanship - and economic satisfaction (0-10).

Table C2: Covariate Balance in Mexican Sample

	Control			Treatment			
	n	mean	sd	n	mean	sd	Diff
Gender	630	1.50	0.50	1898	1.50	0.50	0.005
Age category	630	5.07	2.88	1898	4.88	2.88	-0.188
Interest in politics	618	3.65	1.19	1866	3.65	1.19	0.001
Party choice	630	4.87	3.39	1898	4.88	3.35	0.008
Attitude Morena	584	2.78	1.52	1799	2.78	1.53	0.006
Attitude PAN	578	2.55	1.34	1787	2.58	1.35	0.027
Attitude PRI	580	2.14	1.34	1794	2.05	1.29	-0.088
Economic satisfaction	615	4.02	3.05	1868	3.90	2.99	-0.123

Note: Table shows averages for baseline.
The Diff column is the coefficient of a simple regression of treatment status on the variable.
Stars indicate whether this difference is significant. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Table for Figure 1

Table D1 below presents the full model used to generate Figure 1 in the main text. The table shows unstandardized regressions estimates from OLS regression using robust standard errors.

Table D1: Effect of Electoral Malpractice on Perceptions of Electoral Integrity in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Treatment	-2.60***	-1.23***	-3.77***	-2.00***	-2.14***	-0.66***	-1.44***	-0.43***
	(0.06)	(0.07)	(0.16)	(0.18)	(0.06)	(0.07)	(0.06)	(0.07)
Constant	4.19***	3.18***	5.55***	4.74***	3.93***	3.07***	3.74***	2.83***
	(0.05)	(0.06)	(0.15)	(0.16)	(0.05)	(0.06)	(0.05)	(0.06)
adj. R^2	0.497	0.128	0.253	0.063	0.393	0.044	0.202	0.020
N	1915	2059	1915	2059	1915	2059	1915	2059

Note: Unstandardized OLS regression estimates are shown. Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As can be seen from Table D1, the difference-in-means between treatment groups and the control groups are substantively large and statistically distinguishable from zero across all outcomes. The estimation sample includes only respondents that have provided answers to all outcomes measures. Answers in the categories "Don't know" and "Prefer not to answer" are coded as missing.

Figure 1 in the main text presents these differences with the exception of compliance. Compliance is used as an alternative conceptualization of legitimacy (e.g., Levi et al. 2009). As Table D1 shows, information revealing electoral malpractice causes respondents to feel markedly less obliged to comply with government decisions. The effect is particularly pronounced in Denmark. However, the effect in Mexico is still substantively and statistically significant.

E Table for Figure 2 and 3

Figures 2 and 3 in the main text show marginal effects of treatment and predicted means conditional on partisanship. Figure E1 below presents the model used to estimate these. The table presents unstandardized OLS regression estimates for the conditional treatment effect with robust standard errors.

Table E1: Effect of Electoral Malpractice on Citizen Attitudes across Partisanship in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Treatment	-2.13***	-0.80***	-0.83**	-0.84***	-1.83***	-0.36**	-1.47***	-0.20
	(0.20)	(0.11)	(0.28)	(0.24)	(0.19)	(0.11)	(0.19)	(0.11)
Moderate Out-Partisan	0.73***	0.55**	2.36***	2.15***	0.59**	0.79***	0.61**	0.77***
	(0.22)	(0.20)	(0.35)	(0.41)	(0.21)	(0.16)	(0.21)	(0.16)
Neutral	0.81***	0.88***	4.45***	3.46***	0.80***	0.83***	0.45*	0.77***
	(0.21)	(0.16)	(0.34)	(0.33)	(0.20)	(0.14)	(0.20)	(0.16)
Moderate Co-partisan	0.99***	1.56***	6.20***	4.94***	1.05***	1.25***	0.68***	0.79***
	(0.20)	(0.17)	(0.30)	(0.33)	(0.19)	(0.15)	(0.20)	(0.17)
Strong Co-partisan	0.91***	2.02***	7.02***	6.44***	1.12***	1.83***	0.71***	0.98***
	(0.22)	(0.16)	(0.36)	(0.36)	(0.21)	(0.15)	(0.21)	(0.18)
Treatment ×								
Moderate Out-Partisan	-0.61**	-0.45*	-2.18***	-1.21**	-0.42	-0.46*	-0.14	-0.61**
	(0.23)	(0.22)	(0.37)	(0.46)	(0.23)	(0.19)	(0.23)	(0.19)
Treatment ×								
Neutral	-0.61**	-0.48*	-3.52***	-1.54***	-0.48*	-0.27	-0.00	-0.36*
	(0.23)	(0.19)	(0.37)	(0.39)	(0.22)	(0.17)	(0.22)	(0.18)
Treatment ×								
Moderate Co-Partisan	-0.59**	-0.75***	-4.01***	-1.75***	-0.31	-0.40*	0.16	-0.22
	(0.22)	(0.20)	(0.35)	(0.41)	(0.21)	(0.18)	(0.22)	(0.20)
Treatment ×								
Strong Co-Partisan	-0.03	-0.80***	-3.20***	-1.78***	0.18	-0.69***	0.48	-0.18
	(0.25)	(0.21)	(0.46)	(0.47)	(0.25)	(0.19)	(0.25)	(0.21)
Constant	3.44***	2.39***	1.38***	2.10***	3.17***	2.35***	3.22***	2.32***
	(0.19)	(0.10)	(0.26)	(0.21)	(0.18)	(0.10)	(0.18)	(0.09)
adj. R^2	0.526	0.247	0.496	0.340	0.463	0.165	0.257	0.075
N	1880	1996	1880	1996	1880	1996	1880	1996

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

The first row in Table E1 shows that learning upon malpractice caused strong out-partisan (the reference group) to update their views on elections and government in both countries. Only the estimated difference-in-means among strong out-partisans on compliance in Mexico fails to reach statistical significance at conventional levels. The following rows shows the estimated difference-in-means among non-treated partisans relative to strong out-partisans. These coefficients show that people tend to view elections as more fair if won by their own party even if there was no mention of malpractice in their vignettes. For instance, strong co-partisans in Denmark rate election fairness 0.91 scale points higher compared to strong out-partisans. This difference is a remarkable 2.02 in Mexico.

Similarly, people tend to view co-party governments as more legitimate even in the absence of malpractice. That is, people perceive governments as more entitled to make bindings decisions when they agree with them politically. The interaction terms shows that treatment produced a slightly larger effect on perceived election fairness among moderate out-partisans compared to strong out-partisans in Denmark. The negative effect among the former was 0.61 larger compared to the effect among the latter. Similarly, Table E1 shows that the difference-in-means among strong co-partisans was not distinguishable from the difference-in-means among strong out-partisans in Denmark.

In addition to the hypotheses in the main text, I also hypothesize that partisanship colors how people view elections in the absence of electoral malpractice. This corresponds to H4 in the pre-registration. A vast literature has consistently found that supporters of the winning party perceive elections as more fair than supporters of the losing party (Anderson et al. 2005; Cantú and García-Ponce 2015; Sances and Stewart 2015; Levy 2020; Flesken and Hartl 2018). Theoretically, it has been argued that people rationalize their views on elections when their preferred party loses (Sinclair et al. 2018). To examine whether people perceive elections as more fair when they side with the winners, we can compare the average PEI and legitimacy scores across partisanship among respondents in the control group. As Table E1 shows, both Danes and Mexicans show signs of a "winner's effect"; elections are viewed as more fair and governments are more legitimate by co-partisans of the winning party. These differences become even clearer when the model is estimated using a linear interaction parameter (see SI Appendix F below).

In Figure 2 in the main text and in Table E1 above, neutrals seem to sanction electoral malpractice quite a lot. Recall that the measure of partisanship is a five-point like-dislike score (cf. Abramowitz and Webster 2016, 2018; Wagner 2020), that captures how respondents feel towards the party the read about in the experiment. Neutrals on this scale are those that score three on this scale. That is, neutrals are exactly that; neutral. They have neither positive nor negative feelings towards the party they read about. As both Mexico and Denmark are multi-party democracies, such attitudes are entirely normal. Some respondents may feel closer to certain minor parties that were not included in this experiment, and have no particular feelings towards the ones mentioned in this experiment. However, it is also likely that the neutral group consists of a large part of respondents who do not pay particular interest in politics. Nevertheless, the harsh sanctioning from these respondents are quite interesting. With regards to perceived election fairness (PEI), neutrals have baseline level of 4.25 (Denmark) and 3.27 (Mexico). Upon learning about malpractice from a party they do not hold particular negative or positive affection for, they reduce such perception by 2.74 (Denmark) and 1.28 (Mexico) on average. In both countries, these adjustments are similar to those of moderate co- and out-partisans. Regarding support, neutrals hold - as expected - pre-treatment support levels around the midpoint of the scale. In Denmark, neutrals' baseline support level is 5.82, whereas the

baseline in Mexico is 5.55. Danish neutrals adjust their support by -4.36 upon learning about malpractice, whereas Mexican neutrals adjust by -2.38. In Table M3 below, show how large these drops are in relation to baseline levels. This table shows that the reaction among neutrals is quite similar to reactions among moderates on both sides.

F Reproducing Figure 2 with linear interaction

Table F1 is similar to Table E1, but estimating the interaction as a linear interaction parameter.

Table F1: Effect of Electoral Malpractice on Citizen Attitudes across Partisanship in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Treatment	-2.54***	-0.64***	-0.98**	-0.55	-2.20***	-0.25	-1.79***	-0.29*
	(0.17)	(0.14)	(0.30)	(0.29)	(0.17)	(0.13)	(0.17)	(0.13)
Co-Partisan	0.20***	0.50***	1.81***	1.62***	0.27***	0.43***	0.14**	0.25***
	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.03)	(0.04)	(0.04)
Treatment × Co-partisan	-0.01	-0.21***	-0.88***	-0.50***	0.04	-0.15***	0.13**	-0.05
	(0.05)	(0.05)	(0.09)	(0.10)	(0.05)	(0.04)	(0.05)	(0.05)
Constant	3.57***	1.88***	0.05	0.60*	3.12***	1.96***	3.33***	2.20***
	(0.16)	(0.12)	(0.27)	(0.26)	(0.15)	(0.11)	(0.15)	(0.11)
adj. R^2	0.521	0.246	0.483	0.340	0.458	0.166	0.252	0.073
N	1880	1996	1880	1996	1880	1996	1880	1996

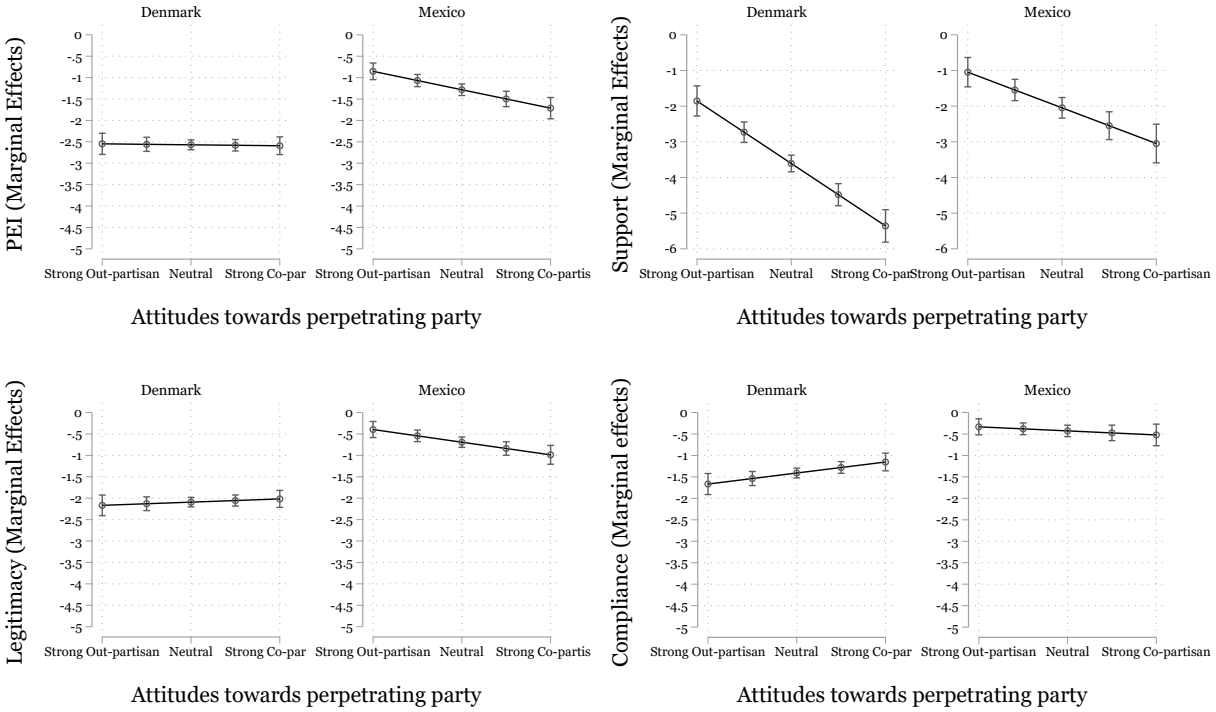
Note: Unstandardized OLS regression estimates are shown. Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As Table F1 clearly shows, the findings in the main text are not dependent on parametrical assumptions regarding the interaction term. As noted above, Table E1 clearly demonstrates the winner-loser gap in perceptions of electoral integrity and government legitimacy. The second parameter in all models shows the average difference in outcomes for a one-unit increase in partisanship (scale 1-5), for respondents in the control condition. These show that election winners view elections as more fair ($\hat{\beta}$ = .20 and .50 in Denmark and Mexico, respectively), and view governments as more legitimate ($\hat{\beta}$ = .27 and .43 in Denmark and Mexico, respectively). While a .20 increase in PEI appears quite modest, the gap between strong out-partisans and strong co-partisans is around 21.9 percentage points (p < .001). Similarly in Mexico, supporters of the winning party perceive the election to be significantly fairer (Δ 1 $\approx$ 12.5 percentage points gap in PEI, p < .001; Δ min-max $\approx$ 50.3 percentage points in PEI, p < .001).

Interestingly, receiving information revealing electoral malpractice does not produce a statistically significant difference-in-means among strong out-partisans in Mexico on support ($\hat{\beta} = -0.55$; se = 0.29) and perceived legitimacy ($\hat{\beta} = -0.25$; se = 0.13). As shown in Figure 3 in the main text, these respondents hold very negative views on government in the baseline, creating a floor effect.

Figure F1 shows the marginal effects of treatment conditional on partisanship as predicted by a linear interaction model (Table F1).

Figure F1: Effects of Electoral Manipulation on Citizen Attitudes Conditioned on Partisanship



Note: The figure displays mean values conditional on party affiliation and treatment assignment. Point estimates are depicted with a 95% confidence interval

Figure F1 is very similar to Figure 2 in the main text, validating the assumption of linearity in the interaction term (Hainmueller et al. 2019). The findings in Figure F1 are, therefore, not driven by model-specification. As shown, partisanship does not moderate reactions to malpractice in Denmark, although there seems to be a slight moderation with regards to compliance. Perhaps people tend to comply with a co-party government even though it meddled with elections to win office. In Mexico, Figure F1 shows that partisans react more strongly to malpractice by their party than to otherwise identical malpractice by opposition parties.

G Robustness to Non-Collapsed Treatment Groups

The experiments randomly assigned each respondent to one of three treatment conditions or a placebo condition. The treatment groups read vignettes about either vote-buying, voter pressure, or ballot-box stuffing. The control group read a vignette about the number of televised debates. Because people may not respond similarly to different *types* of malpractice (Mares and Visconti 2019; Collier and Vicente 2012; Harvey and Mukherjee 2018), I randomize three different types, allowing me to examine the consequences of malpractice more broadly. All analyses in the main paper collapsed treatment groups to estimate the average treatment effect of receiving information revealing electoral malpractice on people's attitudes toward government. Table G1 and G2 below replicates the findings without collapsing treatment groups. For simplicity, Table G2 estimates the linear interaction parameter, corresponding to Table F1 above.

Table G1: Effect of Electoral Malpractice on Perceptions of Electoral Integrity in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Vote-buying	-2.67***	-1.24***	-3.78***	-1.97***	-2.16***	-0.55***	-1.47***	-0.44***
	(0.07)	(0.09)	(0.19)	(0.21)	(0.07)	(0.08)	(0.07)	(0.08)
Voter coercion	-2.46***	-1.08***	-3.89***	-1.81***	-2.07***	-0.64***	-1.37***	-0.49***
	(0.08)	(0.09)	(0.19)	(0.22)	(0.07)	(0.08)	(0.08)	(0.08)
Ballot stuffing	-2.67***	-1.38***	-3.64***	-2.22***	-2.20***	-0.79***	-1.50***	-0.38***
	(0.07)	(0.09)	(0.19)	(0.21)	(0.07)	(0.08)	(0.08)	(0.08)
Constant	4.19***	3.18***	5.55***	4.74***	3.93***	3.07***	3.74***	2.83***
	(0.05)	(0.06)	(0.15)	(0.16)	(0.05)	(0.06)	(0.05)	(0.06)
adj. R^2	0.499	0.132	0.253	0.064	0.394	0.047	0.202	0.020
N	1915	2059	1915	2059	1915	2059	1915	2059

Note: Unstandardized OLS regression estimates are shown. Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As Table G1 shows, collapsing treatment groups do not influence the results. Indeed, the minor variation in effects between types of malpractice appears unsystematic. Treatment effects for all types of malpractices are negative and of similar size.

Importantly, Table G1 also demonstrates that respondents did not respond more or less harshly to certain types of malpractice depending on their plausibility. For instance, voter pressure is not equally plausible for all parties as the operationalization in the experiment concerned threatening voters with losing access to social programs.[1] Parties can only credible threaten to remove access to social benefits if they have (or are likely to have) government control. That is, of course, not equally likely for all parties. However, as Table G1 and E1 show, results for voter coercion are comparable to vote buying and ballot fraud. There is, therefore, reason to believe that the potential lack of plausibility is driven of certain party-malpractice combinations.

As Table G2 shows, the partisan moderation is similar across types of malpractice. The only difference

[1] I thank an anonymous reviewer for raising this point.

Table G2: Effect of Electoral Malpractice on Citizen Attitudes across Partisanship in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Vote-buying	-2.77***	-0.64***	-1.49***	-0.62	-2.34***	-0.21	-1.85***	-0.24
	(0.19)	(0.16)	(0.35)	(0.34)	(0.18)	(0.16)	(0.20)	(0.16)
Voter Coercion	-2.32***	-0.68***	-0.52	-0.58	-1.98***	-0.22	-1.71***	-0.34*
	(0.21)	(0.17)	(0.36)	(0.35)	(0.20)	(0.16)	(0.20)	(0.16)
Ballot stuffing	-2.51***	-0.59***	-0.96**	-0.44	-2.30***	-0.30	-1.82***	-0.27
	(0.20)	(0.16)	(0.37)	(0.35)	(0.19)	(0.16)	(0.21)	(0.16)
Co-Partisan	0.20***	0.50***	1.81***	1.62***	0.27***	0.43***	0.14**	0.25***
	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.03)	(0.04)	(0.04)
Vote-buying ×								
Co-partisan	0.04	-0.21***	-0.72***	-0.48***	0.07	-0.12*	0.14*	-0.07
	(0.06)	(0.06)	(0.12)	(0.12)	(0.06)	(0.05)	(0.06)	(0.06)
Voter coercion ×								
Co-partisan	-0.04	-0.14*	-1.08***	-0.42**	-0.02	-0.15**	0.13*	-0.05
	(0.06)	(0.06)	(0.12)	(0.13)	(0.06)	(0.05)	(0.06)	(0.06)
Ballot stuffing ×								
Co-partisan	-0.04	-0.29***	-0.81***	-0.60***	0.06	-0.17**	0.12	-0.03
	(0.06)	(0.06)	(0.13)	(0.13)	(0.06)	(0.05)	(0.06)	(0.06)
Constant	3.57***	1.88***	0.05	0.60*	3.12***	1.96***	3.33***	2.20***
	(0.16)	(0.12)	(0.27)	(0.26)	(0.15)	(0.11)	(0.15)	(0.11)
adj. R^2	0.524	0.252	0.486	0.340	0.459	0.168	0.252	0.072
N	1880	1996	1880	1996	1880	1996	1880	1996

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

is that the interaction term for ballot-stuffing and co-partisanship is insignificant. The coefficient estimate is, however, in the same direction as those for vote-buying and voter coercion and similar in size.

H Robustness to Alternative Partisanship Measures

Table H1 below replicates Table F1 and Table E1 using a different measure of partisanship. This measure of partisanship groups respondents into three categories (opponents, co-party leaners, and co-party identifiers) using three pre-treatment survey items (see SI Appendix A). I match respondents' party choice and self-reported identification as supporters of that party to the party mentioned in their vignettes.

Table H1: Effect of Electoral Malpractice on Citizen Attitudes across Partisanship in Denmark and Mexico

	PEI		Support		Legitimacy		Compliance	
	DEN	MEX	DEN	MEX	DEN	MEX	DEN	MEX
Treatment	-2.61***	-1.11***	-3.43***	-1.69***	-2.17***	-0.54***	-1.51***	-0.54***
	(0.07)	(0.08)	(0.17)	(0.18)	(0.07)	(0.07)	(0.07)	(0.07)
Co-party leaner	0.31*	1.30***	2.89***	3.58***	0.34*	1.17***	0.07	1.17***
	(0.15)	(0.18)	(0.29)	(0.36)	(0.14)	(0.14)	(0.14)	(0.14)
Co-party identifier	0.38**	1.51***	3.79***	4.86***	0.52***	1.38***	0.20	1.38***
	(0.12)	(0.13)	(0.26)	(0.29)	(0.11)	(0.12)	(0.12)	(0.12)
Treatment × leaner	-0.01	-0.69**	-0.99*	-1.18*	0.24	-0.68***	0.45*	-0.68***
	(0.19)	(0.24)	(0.42)	(0.50)	(0.18)	(0.20)	(0.19)	(0.20)
Treatment × identifier	0.28	-0.33	-0.82*	-1.03*	0.44**	-0.41**	0.47**	-0.41**
	(0.17)	(0.19)	(0.39)	(0.40)	(0.16)	(0.16)	(0.17)	(0.16)
Constant	4.11***	2.89***	4.77***	3.86***	3.83***	2.80***	3.71***	2.80***
	(0.06)	(0.07)	(0.16)	(0.16)	(0.06)	(0.06)	(0.06)	(0.06)
adj. R^2	0.512	0.213	0.375	0.233	0.432	0.120	0.223	0.120
N	1915	2059	1915	2059	1915	2059	1915	2059

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As Table H1 shows, the results in the main text are robust to using this measure of partisanship. Only three minor differences are found. First, the interaction between identifiers and treatment on PEI is insignificant in Mexico. Tables F1 and E1 showed negative interactions between co-partisanship and treatment. Second, the interaction between treatment and identifiers on legitimacy is positive and statistically significant in Denmark. Using the other measure of co-partisanship revealed no interaction in Tables F1 and E1. Third, both leaners and identifiers respond more strongly to treatment with regards to compliance in Mexico. In Tables F1 and E1, there were no differences between co- and out-partisans with regards to compliance in Mexico.

I Country Differences

Table I1 below presents unstandardized OLS regression estimates for the conditional effects of treatment on country (0 = Denmark; 1 = Mexico). The second parameter ("Mexico") is the estimated difference-in-means between control group respondents in Denmark and Mexico. As can be seen, Mexican respondents display a lower baseline level of PEI, support, legitimacy, and compliance, which suggests that there is greater general political apathy in Mexico.

Table I1: Country Differences in Citizens' Responses to Electoral Malpractice

	PEI	Support	Legitimacy	Compliance
Treatment	-2.54***	-3.49***	-2.03***	-1.33***
	(0.06)	(0.15)	(0.06)	(0.06)
Mexico	-0.97***	-0.72***	-0.79***	-0.86***
	(0.08)	(0.18)	(0.08)	(0.08)
Treatment × Mexico	1.29***	1.56***	1.37***	0.90***
	(0.09)	(0.21)	(0.09)	(0.09)
Constant	4.15***	5.41***	3.86***	3.69***
	(0.06)	(0.13)	(0.05)	(0.05)
adj. R^2	0.306	0.136	0.220	0.104
N	4613	4614	4457	4453

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

Table I2: Country Differences in Citizens' Partisan Responses to Electoral Malpractice

	PEI	Support	Legitimacy	Compliance
Treatment	-2.44***	-0.84*	-2.10***	-1.74***
	(0.16)	(0.33)	(0.15)	(0.16)
Co-partisan	0.20***	1.82***	0.27***	0.14***
	(0.04)	(0.09)	(0.04)	(0.04)
Treatment × Co-partisan	-0.03	-0.85***	0.03	0.14**
	(0.05)	(0.10)	(0.05)	(0.05)
Mexico	-1.63***	0.69	-1.09***	-1.05***
	(0.17)	(0.36)	(0.17)	(0.17)
Treatment × Mexico	1.82***	0.31	1.85***	1.46***
	(0.20)	(0.41)	(0.19)	(0.20)
Treatment × Mexico × Co-partisan	-0.20**	0.34*	-0.18**	-0.19**
	(0.06)	(0.13)	(0.06)	(0.06)
Constant	3.53***	-0.07	3.05***	3.27***
	(0.14)	(0.29)	(0.13)	(0.14)
adj. R^2	0.370	0.391	0.312	0.158
N	4440	4463	4311	4311

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As Table I1 shows, the interaction term is positive and statistically significant; treatment produced smaller

effects on PEI, support, legitimacy, and compliance in Mexico compared to Denmark. Table I2 shows unstandardized OLS regression estimates for the three-way interaction between treatment, partisanship, and country. As can be seen, the parameter estimate for the threeway interaction (Treatment × Mexico × Co-partisan) is statistically significant across outcomes. Looking at support, for example, Table I2 shows that the slope for Treatment × Co-partisan is steeper in Denmark. However, on the three other outcomes, the slope coefficients for the conditional effects of treatment and co-partisanship are steeper in Mexico. These differences are shown in Figure F1 above.

J Robustness to Attrition

Out of the total samples in both countries (N = 2,526 Denmark; N = 2,528 Mexico), 611 respondents (Denmark) (24 %) provided non-usable answers (i.e. don't know/prefer not to answer) to one of the main outcomes. In Mexico, 469 respondents (18.5 %) dropped out. An additional 35 Danes and 63 Mexicans dropped out because they did not answer questions regarding partisanship. Table J1 below shows that missingness due to dropout or answering "don't know" is not related to treatment assignment in Mexico. Among Danish respondents, those receiving vote-buying or voter coercion vignettes were approximately 7.2 percentage points and 6.4 percentage points less likely to drop out than respondents in the control condition, respectively. Danes receiving ballot stuffing were as likely to cause missingness as the control group.

Table J1: Logit model explaining a respondent's propensity to answer "don't know/prefer not to answer" on a key variable

	Denmark	Mexico
Vote-buying	-0.47***	0.00
	(0.13)	(0.14)
Voter Coercion	-0.40**	-0.05
	(0.13)	(0.14)
Ballot stuffing	-0.05	-0.26
	(0.13)	(0.15)
Constant	-0.92***	-1.41***
	(0.09)	(0.10)
N	2526	2528

Note: Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

While the slightly higher propensity to dropout among Danes in vote-buying and voter coercion groups is not ideal and could potentially bias the estimations in Tables D1 and E1, the substantial relationship between them is very weak. Indeed, the statistically significant relationship between treatment assignment and/or covariates and attrition is not, by itself, evidence of biases (Gerber and Green 2012). The predicted probability of dropping out among the control group in Denmark is 0.28 (95 % CI [0.25-0.32]). In comparison the predicted probabilities of dropping for vote-buying and voter coercion are 0.19 (95 % CI [0.18-0.25])and 0.21 (95 % CI [0.19-0.25]), respectively. Given the substantively large treatment effects, it is very doubtful that the relationship between missingness and vote-buying/voter coercion drives the results in the main texts. Below, I nevertheless take two approaches to examine whether the non-ignorable relation between treatment assignment and attrition in Denmark has substantially biased the findings.

The analyses in the main text are complete-case-analyses (CCA), which drops respondents that have at least partially missing data. Data that are not not missing-completely-at-random (MCAR) (see Rubin 1976) might produce biased parameter estimates (King et al. 2001). As Table J1 shows, treatment assignment and

attrition is correlated because respondents who received information that this election campaign had had more televised debates between the candidates for prime minister were more likely to provide a non-usable (missing) answer on at least one of the four outcomes relative to those reading about vote-buying or voter coercion. To examine whether the results were robust to potential attrition bias, I ran multiple imputations (MI). One of the key benefits of MI is that it can be implemented even when data are missing-not-at-random (MNAR) (Sullivan et al. 2018). That is, even when missingness is non-random given observed covariates, one can use MI to meaningfully predict values of missing variables.

In the first stage, I simulated multiple values for each missing observation using a linear imputation model with treatment condition, partisanship, economic performance evaluations, political interest, and the other outcomes as regressors. One of the advantages of MI is its' ability to use data on different outcomes and their associations to more precisely predict values of missing variables (Sullivan et al., 2018, p. 2623). To reduce the risks of outliers in imputed data set and to acknowledge variation and uncertainty in imputations, I created 100 imputed data sets ($m = 100$) for each outcome independently.[2] Doing so accounts for variance within each estimate as well as across the imputed data sets. In the second stage, I then re-ran the analysis from Table D1 and Figure 1 in the main text. Table J1 presents unstandardized OLS regression estimates on imputed data.

Table J2: Differences-in-Means using Imputed Data (MI) among Danish Respondents

	PEI	Support	Legitimacy	Compliance
Treatment	-2.50***	-3.49***	-2.02***	-1.32***
	(0.05)	(0.13)	(.05)	(.06)
	[-2.61 - -2.39]	[-3.76 - -3.22]	[-2.13 - -1.91]	[-1.44 - -1.20]
Constant	4.13***	5.42***	3.85***	3.67***
	(.04)	(.12)	(.04)	(.05)
N	2313	2285	2262	2240
Imputations	100	100	100	100

Note: Unstandardized OLS regression estimates are shown. Standard errors are in parentheses and 95 % CI in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

Table J1 shows that the main findings are not biased by non-random missingness. Regarding PEI, for instance, the parameter estimate using 100 imputed data sets is -2.50 (se = 0.05; 95 % CI [-2.61 - -2.39]; p < 0.001). In comparison, the parameter estimate in Table D1 and Figure 1 in the main text is -2.60 yielding an estimated bias of .10, which relative to scale is about 2 %. The parameter estimate across computed data sets are most compatible with a difference-in-means of -2.50 but are also reasonably compatible with any difference-in-means within the range of -2.39 and -2.61. This suggests two things. First, the parameter estimate presented in Figure 1 in the main text is not severely challenged by non-random missingness bias.

[2] A rule of thumb is that one should create at least as many imputed data sets at the percentage of missing data (Sullivan et al. 2018).

Second, even the lowest ATE compatible with the data is substantively large.

The same is true for political support. Recall that the estimated ATE in Figure 1 in the main text is -3.77 (se = 0.16; p < 0.001). Across imputed data sets, the parameter estimate most compatible with the data is -3.49 (se = 0.13; 95 % CI [-3.76 - -3.22]; p < 0.001). The estimated bias compounds to approximately 2.5 %. Again, any ATE estimate within the range of reasonable data compatibility is substantively and statistically significant. The last two columns of Table J2 show that the findings regarding legitimacy and compliance are also robust to non-random missingness. One thing to note is that CCA estimates for differences-in-means (like the ones in Table D1 and Figure 1 in the main text) all fall left of the parameter estimate using multiple imputations. As this section has shown, however, these biases are minor and do not change the conclusions of the main study.

Another way to gauge whether the findings are robust to biases from attrition is to place bounds on the treatment effect estimate (e.g., Gerber and Green, 2012, chap. 7.4). Doing so allows me to estimate the smallest and largest ATEs that would be found if the missing values had extremely high or low values. It, therefore, estimates the worst- and best-case scenarios (Sullivan et al. 2018) and enables me to ascertain whether any of the findings in the main text would be rendered inconclusive in either one of these scenarios. So, while Table J1 above gives reason to believe that the findings are relatively robust to attrition bias, Table J3 below presents unstandardized OLS regression estimates for the worst case (Lowest) and best case (Highest) scenarios. Placing extreme value bounds is particularly suitable for "bracketing the true ATE" (Gerber and Green, 2012, p. 227) when there is a narrow range of feasible outcomes. To bound the ATEs, I replaced missing values with either the highest or lowest possible scores on the four outcomes. In the worst-case scenario for PEI, for example, all respondents in the control group with missing values on PEI were coded as 1 (i.e. lowest possible score), while all missing values for treated respondents were coded as 5 (the highest possible score). The best-case analysis reversed this coding.

Table J3: Boundend ATE of Electoral Malpractice on Attitudes among Danish Respondents

	PEI		Support		Legitimacy		Compliance	
	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
Treatment	-1.68***	-2.74***	-1.91***	-4.33***	-1.23***	-2.31***	-0.54***	-1.73***
	(0.07)	(0.05)	(0.16)	(0.15)	(0.07)	(0.05)	(0.06)	(0.05)
Constant	3.59***	4.30***	4.66***	6.05***	3.41***	4.04***	3.30***	3.88***
	(0.06)	(0.04)	(0.14)	(0.13)	(0.06)	(0.04)	(0.05)	(0.04)
adj. R^2	0.195	0.546	0.051	0.305	0.114	0.445	0.024	0.271
N	2526	2526	2526	2526	2526	2526	2526	2526

Note: Unstandardized OLS regression estimates are shown. Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

As Table J3 shows, placing extreme value bounds on the ATE estimates does not challenge the conclusions

in the main paper. While the parameter estimates naturally fluctuates, the bounds indicate that the ATE in all cases would be in the same direction as reported in the main text, would be substantively large, and statistically significant at the 1 %-level. For example, in the scenario where all the missing values on legitimacy for control group respondents were extremely low and all missing values for the treatment group were extremely high, the difference-in-means is still -1.23 ($p < 0.001$), which is both substantively and statistically large.

K Preregistration Hypotheses

This Appendix section presents an anonymized version of the preregistration. Please note that the preregistration contains additional hypotheses (H7-H9) that are part of a different project.

Data Collection. Have any data been collected for this study already?
No, no data have been collected for this study yet.

Hypotheses

- H1: Electoral manipulation reduces perceived electoral integrity.
- H2: Electoral manipulation reduces political support for the incumbent government.
- H3: Electoral manipulation reduces perceived government legitimacy.
- H4: Individuals report higher levels of perceived electoral integrity, political support, and perceived government legitimacy when they share party affiliation with the winning party.
- H5: Individuals prefer an in-party government that has engaged in electoral manipulation to an out-party government that has not engaged in electoral manipulation.
- H6: The effects of electoral manipulation on perceptions of electoral integrity, political support, and government legitimacy are moderated by partisanship so that individuals judge the same electoral offense less harshly when they share party affiliation with the winning party.

Analyses. The study will run a series of OLS regressions to test the hypotheses. To test H1, H2, and H3, the study will perform an OLS to regress outcomes (perceptions of electoral integrity, political support, and legitimacy, respectively) on treatment conditions. This provides the average treatment effects on the primary outcomes. To test for heterogeneous effects of winner-loser status and partisanship, the study will run OLS estimations that regress outcomes (perceptions of electoral integrity, political support, and legitimacy) on treatment interacted by partisanship. Respondents' partisanship is measured pre-treatment. The study measures partisan affiliations in two ways.

- First, respondents' attitudes towards the governing parties - that are also mentioned in treatment vignettes - are measured on a five-point Likert-scale. Respondents are asked: "We would like to understand your attitudes toward the major political parties. Please rate the following parties on a scale from 1 to 5, where 1 means that you dislike the party a great deal, and 5 means you like the party a great deal."

- Second, respondents are asked which party or which party's candidate they would vote for if an election were held tomorrow. Subsequently, respondents are asked whether they identify as supporters of that party. A variable is then constructed which measures partisans, leaners, and non-partisans.

Sample Size. The study will collect nationally representative samples of 2500 respondents in each country. Below are power calculations for perceptions of electoral integrity (scale 1-5) and political support (scale 0-10)

Power analysis with regards to perceptions of electoral integrity (assuming a standard deviation of 1):

- The sample size of n = 2500 (comparing 1875 treated to 625 in the control group) allows detection of effect sizes on perceptions of electoral integrity (1-5) of 0.129 scale points (approximately 3.2 percentage points) with power 0.8.
- For effects between individual treatment conditions and the control group (comparison n = 1250), the lowest detectable effect size with a power of 0.8 is 0.158 (approximately 3.9 percentage points).

Power analysis (assuming a standard deviation of 1.5):

- The sample size of n = 2500 allows detection of effect sizes on perceptions of electoral integrity (1-5) of 0.19 (approximately 4.8 percentage points) with a power of 0.8.
- For comparisons between individual treatment conditions and the control conditions (comparison n = 1250), the lowest detectable effect size is 0.23 (approximately 5.75 percentage points) with power of 0.8.

Power analysis with regards to political support (assuming a standard deviation of 3.3):

- The sample size of n = 2500 (comparing 1875 treated to 625 in the control group) allows detection of effect sizes on political support (0-10) of 0.427 scale points (approximately 4.2 percentage points) with power 0.8.
- For effects between individual treatment conditions and the control group (comparison n = 1250), the lowest detectable effect size with a power of 0.8 is 0.52 (approximately 5.2 percentage points).

Outliers and Exclusions. The survey includes a simple, pre-treatment, instructional attention check, which aims to gauge whether respondents are taking the time to read the questions. This check is not scripted as a screener (i.e. respondents are not screened out of the survey), but if a large part of the sample fails to pass this check, the study will run analyses on both a split-sample of attentive respondents and the entire sample.

The attention check is as follows.

> "When a big news story breaks, people often go online to get the latest information about what is going on.
>
> In this question, we want to know if you are reading the questions. To show that you have read this much, select El Financiero as your answer below. No matter what you prefer, just choose this option.
>
> 1 Twitter
>
> 2 Facebook
>
> 3 El Debate
>
> 4 El Universal
>
> 5 Diario.mx
>
> 6 Proceso
>
> 7 Sopitas.com
>
> 8 El Financiero
>
> 9 Milenio
>
> 10 Excelsior
>
> 11 Don't know
>
> 11 Prefer not to answer"

Secondly, the survey includes a post-treatment attention check that asks respondents to recall which party won the election that they have just read about. This is meant to assess whether respondents took the time to read the experimental vignettes. While it is not the intention at the outset, the study has the possibility of examining complier average treatment effects in addition to average treatment effects if a strikingly large number of respondents fail to identify the winning party. Note, however, that the main analyses will be carried out on the entire sample regardless of attention or compliance.

L Comparing Marginal Means based on Figure 3

Table L1: Marginal Support Mean Comparison across Treatment and Partisanship (Denmark)

Comparison Groups			Difference	SE	95 % CI
Control × Mod. out-party	vs.	Control × Strong out-party	2.36***	(.347)	[1.67 - 3.04]
Control × Neutral	vs.	Control × Strong out-party	4.44***	(.340)	[3.78 - 5.12]
Control × Mod. co-party	vs.	Control × Strong out-party	6.20***	(.302)	[5.61 - 6.79]
Control × Strong co-party	vs.	Control × Strong out-party	7.02***	(.356)	[6.32 - 7.72]
Treated × Strong out-party	vs.	Control × Strong out-party	-0.83**	(.275)	[-1.37 - -0.29]
Treated × Mod. out-party	vs.	Control × Strong out-party	-0.65**	(.269)	[-1.18 - -0.12]
Treated × Neutral	vs.	Control × Strong out-party	0.09	(.279)	[-0.45 - 0.64]
Treated × Mod. co-party	vs.	Control × Strong out-party	1.35***	(.297)	[0.77 - 1.94]
Treated × Strong co-party	vs.	Control × Strong out-party	2.98***	(.373)	[2.25 - 3.72]
Control × Neutral	vs.	Control × Mod. out-party	2.08***	(.323)	[1.45 - 2.72]
Control × Mod. co-party	vs.	Control × Mod. out-party	3.84***	(.282)	[3.28 - 4.39]
Control × Strong co-party	vs.	Control × Mod. out-party	4.66***	(.340)	[3.99 - 5.32]
Treated × Strong out-party	vs.	Control × Mod. out-party	-3.19***	(.254)	[-3.69 - -2.69]
Treated × Mod. out-party	vs.	Control × Mod. out-party	-3.01***	(.247)	[-3.50 - -2.53]
Treated × Neutral	vs.	Control × Mod. out-party	-2.27***	(.259)	[-2.77 - -1.76]
Treated × Mod. co-party	vs.	Control × Mod. out-party	-1.00***	(.278)	[-1.54 - -0.45]
Treated × Strong co-party	vs.	Control × Mod. out-party	0.62†	(.357)	[-0.75 - 1.32]
Control × Mod. co-party	vs.	Control × Neutral	1.75***	(.273)	[1.21 - 2.29]
Control × Strong co-party	vs.	Control × Neutral	2.57***	(.332)	[1.92 - 3.22]
Treated × Strong out-party	vs.	Control × Neutral	-5.28***	(.244)	[-5.76 - -4.80]
Treated × Mod. out-party	vs.	Control × Neutral	-5.10***	(.236)	[-5.56 - -4.63]
Treated × Neutral	vs.	Control × Neutral	-4.35***	(.248)	[-4.84 - -3.87]
Treated × Mod. co-party	vs.	Control × Neutral	-3.08***	(.268)	[-3.61 - -2.56]
Treated × Strong co-party	vs.	Control × Neutral	-1.46***	(.350)	[-2.14 - -0.77]
Control × Strong co-party	vs.	Control × Mod. co-party	0.81**	(.293)	[0.24 - 1.39]
Treated × Strong out-party	vs.	Control × Mod. co-party	-7.03***	(.187)	[-7.40 - -6.66]
Treated × Mod. out-party	vs.	Control × Mod. co-party	-6.85***	(.178)	[-7.20 - -6.50]
Treated × Neutral	vs.	Control × Mod. co-party	-6.11***	(.193)	[-6.49 - -5.73]
Treated × Mod. co-party	vs.	Control × Mod. co-party	-4.84***	(.218)	[-5.27 - -4.41]
Treated × Strong co-party	vs.	Control × Mod. co-party	-3.21***	(.313)	[-3.83 - -2.60]
Treated × Strong out-party	vs.	Control × Strong co-party	-7.85***	(.266)	[-8.37 - -7.33]
Treated × Mod. out-party	vs.	Control × Strong co-party	-7.67***	(.259)	[-8.18 - -7.16]
Treated × Neutral	vs.	Control × Strong co-party	-6.93***	(.270)	[-7.46 - -6.40]
Treated × Mod. co-party	vs.	Control × Strong co-party	-5.66***	(.288)	[-6.22 - -5.09]
Treated × Strong out-party	vs.	Control × Strong co-party	-4.03***	(.366)	[-4.75 - -3.31]
Treated × Mod. out-party	vs.	Treated × Strong out-party	0.17	(.129)	[-0.07 - 0.43]
Treated × Neutral	vs.	Treated × Strong out-party	0.92***	(.149)	[0.63 - 1.21]
Treated × Mod. co-party	vs.	Treated × Strong out-party	2.19***	(.180)	[1.83 - 2.54]
Treated × Strong co-party	vs.	Treated × Strong out-party	3.82***	(.288)	[3.25 - 4.38]
Treated × Neutral	vs.	Treated × Mod. out-party	0.74***	(.137)	[0.47 - 1.01]
Treated × Mod. co-party	vs.	Treated × Mod. out-party	2.01***	(.170)	[1.68 - 2.34]
Treated × Strong co-party	vs.	Treated × Mod. out-party	3.64***	(.282)	[3.08 - 4.19]
Treated × Mod. co-party	vs.	Treated × Neutral	1.26***	(.186)	[0.90 - 1.63]
Treated × Strong co-party	vs.	Treated × Neutral	2.89***	(.292)	[2.32 - 3.47]
Treated × Strong co-party	vs.	Treated × Mod. co-party	1.62***	(.309)	[1.02 - 2.23]

Note: Marginal mean comparisons based on unstandardized OLS regression estimates. Standard errors in parentheses and 95 % CI in brackets. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

Table L2: Marginal Support Mean Comparison across Treatment and Partisanship (Mexico)

Comparison Groups			Difference	SE	95 % CI
Control × Mod. out-party	vs.	Control × Strong out-party	2.15***	(.412)	[1.34 - 2.95]
Control × Neutral	vs.	Control × Strong out-party	3.45***	(.331)	[2.80 - 4.10]
Control × Mod. co-party	vs.	Control × Strong out-party	4.94***	(.328)	[4.29 - 5.58]
Control × Strong co-party	vs.	Control × Strong out-party	6.44***	(.362)	[5.73 - 7.15]
Treated × Strong out-party	vs.	Control × Strong out-party	-0.84***	(.235)	[-1.30 - -0.38]
Treated × Mod. out-party	vs.	Control × Strong out-party	0.09	(.275)	[-0.44 - 0.63]
Treated × Neutral	vs.	Control × Strong out-party	1.08***	(.274)	[0.53 - 1.61]
Treated × Mod. co-party	vs.	Control × Strong out-party	2.35***	(.305)	[1.74 - 2.94]
Treated × Strong co-party	vs.	Control × Strong out-party	3.81***	(.350)	[3.13 - 4.50]
Control × Neutral	vs.	Control × Mod. out-party	1.30**	(.435)	[0.45 - 2.15]
Control × Mod. co-party	vs.	Control × Mod. out-party	2.79***	(.432)	[1.94 - 3.63]
Control × Strong co-party	vs.	Control × Mod. out-party	4.29***	(.458)	[3.39 - 5.19]
Treated × Strong out-party	vs.	Control × Mod. out-party	-2.99***	(.366)	[-3.71 - -2.27]
Treated × Mod. out-party	vs.	Control × Mod. out-party	-2.05***	(.393)	[-2.82 - -1.28]
Treated × Neutral	vs.	Control × Mod. out-party	-1.07**	(.393)	[-1.84 - -0.30]
Treated × Mod. co-party	vs.	Control × Mod. out-party	0.19	(.414)	[-0.61 - 1.00]
Treated × Strong co-party	vs.	Control × Mod. out-party	1.66***	(.449)	[0.78 - 2.55]
Control × Mod. co-party	vs.	Control × Neutral	1.48***	(.356)	[0.78 - 2.18]
Control × Strong co-party	vs.	Control × Neutral	2.98***	(.388)	[2.22 - 3.74]
Treated × Strong out-party	vs.	Control × Neutral	-4.30***	(.273)	[-4.83 - -3.77]
Treated × Mod. out-party	vs.	Control × Neutral	-3.36***	(.308)	[-3.96 - -2.75]
Treated × Neutral	vs.	Control × Neutral	-2.38***	(.308)	[-2.98 - -1.77]
Treated × Mod. co-party	vs.	Control × Neutral	-1.11**	(.335)	[-1.76 - -0.45]
Treated × Strong co-party	vs.	Control × Neutral	0.36	(.377)	[-0.37 - 1.10]
Control × Strong co-party	vs.	Control × Mod. co-party	1.50***	(.385)	[0.74 - 2.25]
Treated × Strong out-party	vs.	Control × Mod. co-party	-5.78***	(.268)	[-6.31 - -5.26]
Treated × Mod. out-party	vs.	Control × Mod. co-party	-4.84***	(.304)	[-5.44 - -4.24]
Treated × Neutral	vs.	Control × Mod. co-party	-3.86***	(.304)	[-4.46 - -3.27]
Treated × Mod. co-party	vs.	Control × Mod. co-party	-2.59***	(.331)	[-3.24 - -1.94]
Treated × Strong co-party	vs.	Control × Mod. co-party	-1.12**	(.373)	[-1.85 - -0.38]
Treated × Strong out-party	vs.	Control × Strong co-party	-7.28***	(.309)	[-7.89 - -6.67]
Treated × Mod. out-party	vs.	Control × Strong co-party	-6.34***	(.341)	[-7.01 - -5.67]
Treated × Neutral	vs.	Control × Strong co-party	-5.36***	(.340)	[-6.04 - -4.69]
Treated × Mod. co-party	vs.	Control × Strong co-party	-4.1***	(.365)	[-4.81 - -3.37]
Treated × Strong out-party	vs.	Control × Strong co-party	-2.62***	(.404)	[-3.41 - -1.83]
Treated × Mod. out-party	vs.	Treated × Strong out-party	0.93***	(.200)	[0.54 - 1.33]
Treated × Neutral	vs.	Treated × Strong out-party	1.92***	(.200)	[1.52 - 2.31]
Treated × Mod. co-party	vs.	Treated × Strong out-party	3.18***	(.239)	[2.71 - 3.66]
Treated × Strong co-party	vs.	Treated × Strong out-party	4.66***	(.295)	[4.08 - 5.24]
Treated × Neutral	vs.	Treated × Mod. out-party	0.97***	(.246)	[0.49 - 1.46]
Treated × Mod. co-party	vs.	Treated × Mod. out-party	2.25***	(.279)	[1.70 - 2.79]
Treated × Strong co-party	vs.	Treated × Mod. out-party	3.72***	(.328)	[3.07 - 4.37]
Treated × Mod. co-party	vs.	Treated × Neutral	1.27***	(.278)	[0.72 - 1.81]
Treated × Strong co-party	vs.	Treated × Neutral	2.74***	(.327)	[2.10 - 3.39]
Treated × Strong co-party	vs.	Treated × Mod. co-party	1.47***	(.353)	[0.77 - 2.17]

Note: Marginal mean comparisons based on unstandardized OLS regression estimates. Standard errors in parentheses and 95 % CI in brackets. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

Tables L1 and L2 compare the marginal mean across partisanship and treatment conditions for Danish and Mexican respondents, respectively. For example, the first row in Table L1 compares a non-treated moderate out-partisan to a non-treated strong out-partisan in Denmark. As the table shows, moderate out-partisans are more supportive of the elected government in the vignettes than strong out-partisans. Tables L1 and L2 clearly show that across all partisan groups, voters prefer a "clean" government to a government that has rigged elections. That is, people have clear preferences for governments that win fair elections.

For the purposes of this study, the most important comparisons are between treated co-partisans and non-treated out-partisans. These comparisons can establish the level of support that people *would have answered* had they been given a scenario in which a out-party government wins an election fair, which can then be compared to their levels of support they express for a co-party government that won using malpractice. A similar approach to establish baselines or priors is used in Arias et al. (2018). Recall that the measure of government support is scaled from 0-10. Hence, Table L1 shows for Danish respondents that treated strong co-partisans express 29.8% ($p < .001$) and 6% ($p < .1$) more support for a co-party government that they know has cheated in the election than they would have expressed for a "clean" out-party government they feel strongly and moderately negative towards, respectively.

Table L2 shows a similar picture for Mexico. Strong co-partisans in the treatment group display 38.1% ($p < .001$) and 16.6% ($p < .001$) more support for a co-party government that they know has cheated in the election than they would have expressed for a "clean" out-party government they feel strongly and moderately negative towards, respectively. Indeed, treated strong co-partisans are as supportive of a co-party government as they would be towards a "clean" government that they do not hold strong attitudes towards.

M Probing for Floor Effects

In the main text, the control group is used to estimate what respondents in the treatment groups *would have answered* had they not been treated. As such, the main text uses the control group to establish baseline levels of support (cf. Arias et al. 2018; Gerber et al. 2010). As shown in the main text, the treatment produces substantively large differences between groups. Moreover, the findings in the main text clearly demonstrates that people do not apply a partisan double standard when evaluating the use of malpractice in elections. Indeed, the results seems to suggest that people react more strongly to co-party malpractice, which is puzzling. Among Danish respondents, the effect of learning about malpractice on government support seems to be greater among co-partisans. Among Mexican respondents, the effects seem to be greater among co-partisans on all outcomes.

One concern with such finding is the potential for floor effects. That is, the larger effect among co-partisans could be driven by a mechanical feature of the measurement. Because partisans are, obviously, more supportive of a co-party government *ex ante* (i.e. prior to receiving information about malpractice), their pre-treatment level of government support will naturally be substantively higher. Consequently, co-partisans will have more room to adjust their levels of support downwards, simply because they start from a higher baseline. Out-partisans, in contrast, have lower pre-treatment scores on all measured outcomes, and so it is possible that these respondents cannot adjust their scores anymore. Consequently, it is possible that the larger marginal effect among co-partisans is "just" an artifact of the measurement strategy.

Following Reuter and Szakonyi (2021), one can imagine an extreme scenario in which the revelation of malpractice causes all respondents to express the lowest possible level of support for the government (i.e. a score of 0 on the support scale). Because out-partisans will have lower pre-treatment scores, they will only exhibit a marginal treatment effect of 1.38 and 3.74 for strong and moderate out-partisans (in Denmark), respectively. In contrast, strong co-partisans will exhibit a marginal effect of 8.40. But these differences may purely be an artifact of the measurement, rather than an indication of genuine differences in responses to information revealing electoral interference. Tables M1 and M2 below display distributions of the outcome variables among control group respondents. These distributions provide a first glance of the potential for floor effects.

Table M1 display means and standard deviations for the three outcome variables among control group respondents across partisanship. These means represent the baseline levels of election fairness judgments, government support, and perceived legitimacy for the different partisan groups. As shown in Table M1 there is substantial pre-treatment (baseline) variation in government support across partisanship as expected. People feel markedly more supportive of a co-party government. The mean baseline level of government

Table M1: Distribution of Outcome Variables for Control Group (Denmark)

Group	PEI		Support		Legitimacy	
	Mean	SD	Mean	SD	Mean	SD
Strong Out-partisans	3.44	1.48	1.38	2.05	3.17	1.45
Moderate Out-partisans	4.17	1.10	3.74	2.40	3.76	1.12
Neutrals	4.25	0.97	5.82	2.16	3.97	.90
Moderate Co-partisans	4.43	0.74	7.59	1.72	4.22	0.69
Strong Co-partisans	4.35	0.90	8.40	1.95	4.29	0.87

Note: Descriptive statistics across partisan groups in the control condition. Denmark (N = 1880).

support among strong out-partisans is just 1.38 (SD = 2.05), indicating extremely negative affection towards an out-party government. In contrast, the baseline level of government support among strong co-partisans is 8.40 (SD = 1.95). The difference in pre-treatment means, in other words, is 7.02 ($p < 0.001$) (Table L1 and L2 above shows significance tests for all support mean comparisons). Perhaps unsurprisingly, the differences on PEI and legitimacy are much smaller. Regardless, the distribution of support among different partisan groups in the control conditions sheds light on the potential for floor effects. In particular, these distributions show that co-partisans have much more room to update their levels of support compared to out-partisans who cannot adjust their support levels as much.

Table M2 displays these distributions among control group respondents in the Mexican sample. Table M2 largely paints the same picture as with the Danish sample. There are, unsurprisingly, very large differences in baseline levels of government support between co- and out-partisans.

Table M2: Distribution of Outcome Variables for Control Group (Mexico)

Group	PEI		Support		Legitimacy	
	Mean	SD	Mean	SD	Mean	SD
Strong Out-partisans	2.39	1.38	2.09	2.92	2.34	1.33
Moderate Out-partisans	2.94	1.32	4.24	2.68	3.14	0.98
Neutrals	3.27	1.26	5.55	2.5	3.17	1.02
Moderate Co-partisans	3.95	1.2	7.03	2.26	3.59	1.07
Strong Co-partisans	4.41	1.00	8.53	2.41	4.17	0.91

Note: Descriptive statistics across partisan groups in the control condition. Mexico (N = 1996).

In contrast to the Danish sample as shown in Table M1, however, is that there are larger discrepancies among Mexican co- and out-partisan in the control condition on PEI and legitimacy. The difference-in-means between strong out- and strong co-partisans on baseline levels of perceived electoral integrity is 2.02. Similarly, the difference in baseline levels of perceived legitimacy is also bigger among Mexican participants compared to Danish participants. Once more, Table M2 indicates that the somewhat larger effects among co-partisans are likely driven by a mechanical feature of the measurement. Moreover, the distributions among

Mexican respondents in the control group also suggests that the somewhat larger effects among co-partisans could be driven by a floor effect.

Another way of gauging whether the larger marginal effects among co-partisans (as shown in Figure 2 in the main text) is comparing the percent change within-group (cf. Reuter and Szakonyi 2021). This allows a comparison between the relative changes across partisanship. Table M3 shows the conditional treatment effects expressed as percent drops. These drops take differences in baseline levels into account. Conditional effects are shown in Figure 2 in the main text (based on Table E1 above) and baseline levels are shown in Table M1 and M2 above.

Table M3: Conditional Treatment Effects as Percent Drops

Group	PEI		Support		Legitimacy	
	Denmark	Mexico	Denmark	Mexico	Denmark	Mexico
Strong Out-partisans	-61.9%	-33.4%	-60.1%	-40.1%	-57.4%	-15.3%
Moderate Out-partisans	-65.7%	-42.1%	-80.4%	-48.3%	-59.6%	-26.1%
Neutrals	-64.4%	-39.1%	-74.7%	-42.8%	-58.1%	-19.5%
Moderate Co-partisans	-61.4%	-39.2%	-63.8%	-36.8%	-50.4%	-20.8%
Strong Co-partisans	-49.4%	-36.2%	-47.9%	-30.7%	-38.2%	-25.2%

Note: Marginal effects expressed at percent drops within partisan groups. Mexico (N = 1996) and Denmark (N = 1880). Calculated as $\frac{difference}{baseline} \times 100$

Table M3 shows suggest that the somewhat smaller conditional effects among out-partisans as displayed in Figure 2 in the main text could quite possibly be driven by floor effects. Recall that among Danish respondents, the treatment effect on government support appeared larger among co-partisans (Figure 2 in main text). However, when taking into account that partisans, naturally, have higher levels of pre-treatment support for a co-party government, it becomes clear that the drop in support among out-partisans relative to baseline levels is a large effect. Learning about out-party malpractice caused strong out-partisans to adjust their level of support downwards with approximately -0.83 (SE = 0.28), which is small change compared the other partisan groups. However, when taking these respondents' baseline levels of government into account, this difference correspond to a 60.1% drop. In contrast the drop in support among strong partisans who learn about co-party malpractice is -4.03 (SE = .33). In absolute terms, this drop is far greater than the drop among out-partisans. However, this "only" correspond to a 47.9% percent drop when considering the relatively higher level of pre-treatment government support.

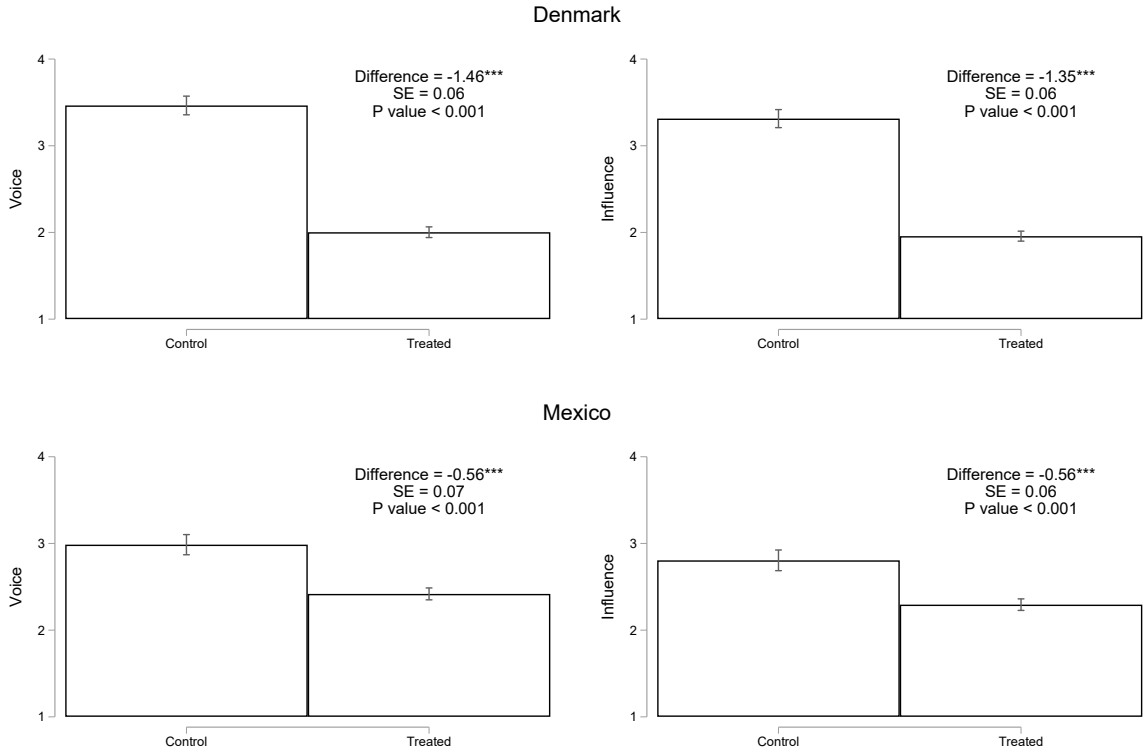
Table M3 paints a similar picture regarding Mexican respondents. Figure 2 also showed that treatment produced greater drops in government support among Mexican co-partisans. However, when taking into account that co-partisans have more room to adjust their support, these difference disappear similar to the Danish respondents. However, in contrast to the Danish respondents, Figure 2 in the main text also shows that Mexican co-partisan react more strongly regarding perceived electoral integrity and government

legitimacy. These differences in effects could also be driven by floor effects. The relative (percent) drops on these outcomes in Table M3, however, paints a more nuanced picture. On one hand, the percent legitimacy drop among strong co-partisans is 25.2%, whereas the same drop among strong out-partisan is 15.3%. Although partisans tend to view a co-party government as more legitimate *ex ante* compared to an out-party, the relative drop in legitimacy remains greater among these respondents when taking baseline levels into account. On the other hand, the drop among moderate out-partisans correspond to 26.1%. Regarding PEI, Table M3 suggests that the percent drop is greater among moderate co- and out-partisans and neutrals.

N Did Respondents Believe Treatment?

The experiment in this study is a direct/applied experiment, which means that respondents were given "factual/hypothetical" information. The experiment then compares attitudes towards government between respondents who received such information to the control group who did not receive such information. In contrast to indirect/theoretical experiments, the aim in direct experiment is not to induce a change in some latent variable (e.g., sentiments or feelings). Rather, the treatment and the independent variable is one and the same (Mutz 2021). In such designs, manipulation check have less value (Mutz 2011, 2021). Indeed, it has been argued that including manipulation checks can even amplify, zero, or interact with treatment (Hauser et al. 2018). In other words, one can influence the "true" treatment effect of receiving information by including a manipulation check, probing whether participants believe the information. On the other hand, not having a manipulation necessarily implies that it is more difficult to ascertain how seriously respondents took the treatment.

Figure N1: Manipulation Check. Respondents believed the election was less fair



Note: The figure displays mean values conditional on treatment assignment based on unstandardized OLS regression estimates. Depicted with a 95% confidence interval. N (Denmark) = 1915 and N (Mexico) = 2059.

As an alternative to a factual manipulation check (e.g., Kane and Barabas 2019), one can examine

whether respondents in the treatment group actually believed that the hypothetical election (in which a party engaged in malpractice) was less fair than control group respondents, by investigating further implications. For instance, one can examine respondents' feelings of voice and influence. These measures capture the extent to which people feel that the hypothetical election they read about gave ordinary people a change to express their views and influence politics. All else being equal, respondents should feel that elections gave them less opportunity to express their views and to influence politics, if they believed that these elections were rigged. When ballot-boxes are stuffed with fake ballots, for example, ordinary voters' voice and influence rapidly diminish. The experiment measured both outcomes. Answers range on a 5-point Likert-type scale (higher values indicates greater voice and influence). Figure N1 presents the unconditional difference-in-means from unstandardizesd OLS regression estimates with robust standard errors.

Figure N1 shows that both Danish and Mexican respondents significantly lowered their perceived voice opportunities and influence. These findings suggests that treated respondents did in fact believe the information they were given about the election, which taken together with the findings regarding election fairness judgments (Figure 1 in the main text) mitigates potential concerns about treatment take-up.

O Were All Party-Malpractice Combinations Plausible?

Respondents assigned to treatment read about either vote-buying, voter pressure, or ballot-box stuffing (see Table 1 in the main text). While vote-buying and ballot-stuffing can be perpetrated by the incumbent government and opposition parties alike, voter pressure (as conceptualized in this experiment) is a more feasible manipulation tactic for incumbent governments (or at least for parties who are likely to enter into government). Hence, a concern is that respondents did not perceive all treatments as equally plausible across combinations of parties and manipulation tactics. For example, if respondents believe that it is unlikely that the Liberals are going to enter into government in Denmark, or that PAN are going to win the Presidency in Mexico, they may respond differently to learning that these parties tried to pressure voters into voting for them by threatening to take away their access to social benefits. The empirical implication of this is that the estimated effects of voter pressure differ depending on which party perpetrates it.

To examine this further, Table O1 and O2 below presents split-sample analyses, presenting treatment effects for election fairness judgments with-in parties. That is, the tables below split the sample into smaller samples, comparing only respondents who read about the same party.

Table O1: Split Sample Analyses of the Effects on PEI (Denmark)

	Only Social Democrats	Only Liberals
Vote-Buying	-2.55***	-2.79***
	(0.10)	(0.09)
Voter Intimidation	-2.31***	-2.61***
	(0.11)	(0.10)
Ballot Stuffing	-2.49***	-2.84***
	(0.10)	(0.10)
Constant	4.11***	4.26***
	(0.07)	(0.07)
adj. R^2	0.453	0.546
N	952	963

Note: Unstandardized OLS regression estimates. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

Table O1 shows that all treatments produced negative effects on perceived election fairness for both the Social Democratic Party and the Liberals. Moreover, the estimated effect sizes are remarkably similar. Table O2 shows that Mexicans also responded similarly to the different party-manipulation combinations. Voter intimidation (threathening to take social benefits away from voters if they did not vote for a given party) produced similar effects on perceived election fairness for all parties in the experiment. The findings demonstrate that the results are not affected by potential differences in plausibility of party-manipulation type combinations.

Table O2: Split Sample Analyses of the Effects on PEI (Mexico)

	Only Morena	Only PRI	Only PAN
Vote-Buying	-1.29***	-1.09***	-1.33***
	(0.15)	(0.15)	(0.15)
Voter Intimidation	-1.03***	-1.06***	-1.18***
	(0.16)	(0.15)	(0.16)
Ballot Stuffing	-1.47***	-1.34***	-1.35***
	(0.15)	(0.14)	(0.15)
Constant	3.42***	2.96***	3.17***
	(0.11)	(0.11)	(0.11)
adj. R^2	0.128	0.131	0.142
N	690	695	674

Note: Unstandardized OLS regression estimates. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-sided tests).

References

Abramowitz, A. I. and S. Webster (2016). The rise of negative partisanship and the nationalization of U.S. elections in the 21st century. *Electoral Studies 41*, 12–22.

Abramowitz, A. I. and S. W. Webster (2018). Negative Partisanship: Why Americans Dislike Parties But Behave Like Rabid Partisans. *Political Psychology 39*(S1), 119–135.

Anderson, C. J., A. Blais, S. Bowler, T. Donovan, and O. Listhaug (2005). *Losers' Consent: Elections and Democratic Legitimacy*. Oxford: Oxford University Press.

Anduiza, E., A. Gallego, and J. Munoz (2013). Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption. *Comparative Political Studies 46*(12), 1664–1692.

APSA (2020). Principles and Guidance for Human Subjects Research. *Technical Report*, DOI: https://connect.apsanet.org/hsr/principles–and–guidance/.

Arias, E., H. Larreguy, J. Marshall, and P. Querubín (2018). Priors Rule: When do Malfesance Revelations Help or Hurt Incumbent Parties? *NBER Working Paper No. 24888 retrieved from: http://www.nber.org/papers/w24888*, 1–50.

Cantú, F. and O. García-Ponce (2015). Partisan losers' effects: Perceptions of electoral integrity in Mexico. *Electoral Studies 39*, 1–14.

Collier, P. and P. C. Vicente (2012). Violence, bribery, and fraud: the political economy of elections in sub-saharan africa. *Public Choice 153*(1-2), 117–147.

Daxecker, U. and H. Fjelde (2020). Electoral Violence, Partisanship and Perceptions of Election Quality: A Survey Experiment in West Bengal, India. *Unpublished Manuscript*.

Druckman, J. N. and K. R. Nelson (2003). Framing and Deliberation: How Citizens' Conversations Limit Elite Influence. *American Journal of Political Science 47*(4), 729–745.

Flesken, A. and J. Hartl (2018). Party support, values, and perceptions of electoral integrity. *Political Psychology 39*(3), 707–724.

Gaines, B. J., J. H. Kuklinski, and P. J. Quirk (2007). The Logic of the Survey Experiment Reexamined. *Political Analysis 15*(1), 1–20.

Gerber, A. S. and D. P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton & Company.

Gerber, A. S., D. P. Green, E. H. Kaplan, and H. L. Kern (2010). Baseline, placebo, and treatment: Efficient estimation for three-group experiments. *Political Analysis 18*(3), 297–315.

Graham, M. H. and M. W. Svolik (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review 114*(2), 392–409.

Gutiérrez-Romero, R. and A. Lebas (2020). Does electoral violence affect vote choice and willingness to vote? Conjoint analysis of a vignette experiment. *Journal of Peace Research*, 002234331989267.

Hainmueller, J., J. Mummolo, and Y. Xu (2019). How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice. *Political Analysis 27*(2), 163–192.

Harvey, C. J. and P. Mukherjee (2018). Methods of election manipulation and the likelihood of post-election protest. *Government and Opposition*, 1–23.

Hauser, D. J., P. C. Ellsworth, and R. Gonzalez (2018). Are Manipulation Checks Necessary? *Frontiers in Psychology 9*(998), 1–10.

Kane, J. V. and J. Barabas (2019). No harm in checking: Using factual manipulation checks to assess attentiveness in experiments. *American Journal of Political Science 63*(1), 234–249.

King, G., J. Honaker, A. Joseph, and K. Scheve (2001). Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review 95*(1), 49–69.

Levi, M., A. Sacks, and T. Tyler (2009). Conceptualizing legitimacy, measuring legitimating beliefs. *American Behavioral Scientist 53*(3), 354–375.

Levy, M. (2020). Winning cures everything? Beliefs about voter fraud, voter confidence, and the 2016 election. *Electoral Studies*, in press.

Luskin, R. C., J. S. Fishkin, and R. Jowell (2002). Considered Opinions: Deliberative Polling in Britain. *British Journal of Political Science 32*(3), 455–487.

Mares, I. and G. Visconti (2019). Voting for the lesser evil: evidence from a conjoint experiment in romania. *Political Science Research and Methods*, 1–14.

Mutz, D. C. (2011). *Population-Based Survey Experiments*. New Jersey: Princeton University Press.

Mutz, D. C. (2021). Improving Experimental Treatments in Political Science. In J. Druckman and D. P. Green (Eds.), *Advances in Experimental Political Science*, pp. 219–238. Cambridge University Press.

Mutz, D. C. and B. Reeves (2005). The New Videomalaise: Effects of Televised Incivility on Political Trust. *American Political Science Review 99*(1), 1–15.

Reuter, O. J. and D. Szakonyi (2021). Electoral Manipulation and Regime Support: Survey Evidence from Russia. *World Politics 73*(2), 275–314.

Rubin, D. B. (1976). Inference and missing data. *Biometrika 63*(3), 581–592.

Sances, M. W. and C. Stewart (2015). Partisanship and confidence in the vote count: Evidence from u.s. national elections since 2000. *Electoral Studies 40*, 176–188.

Sinclair, B., S. S. Smith, and P. D. Tucker (2018). "It's Largely a Rigged System": Voter Confidence and the Winner Effect in 2016. *Political Research Quarterly 71*(4), 854–868.

Sniderman, P. M. (2018). Some advances in the design of survey experiments. *Annual Review of Political Science 21*(1), 259–275.

Sullivan, T. R., I. R. White, A. B. Salter, P. Ryan, and K. J. Lee (2018). Should multiple imputation be the method of choice for handling missing data in randomized trials? *Statistical Methods in Medical Research 27*(9), 2610–2626.

Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy 30*(3), 20–32.

Szakonyi, D. (2021). Candidate Filtering: The Strategic Use of Electoral Manipulations in Russia. *British Journal of Political Science*, 1–22.

Teele, D. L. (2021). Virtual consent: The bronze standard for experimental ethics. In J. Druckman and D. P. Green (Eds.), *Advances in Experimental Political Science*, pp. 130–145. Cambridge University Press.

Tomz, M. and J. L. P. Weeks (2020). Public opinion and foreign electoral intervention. *American Political Science Review*, 1–18.

Tyler, T. R. (1997). The psychology of legitimacy: A relational perspective on voluntary deference to authorities. *Personality and Social Psychology Review 1*(4), 323–345.

Tyler, T. R. (2006). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology 57*(1), 375–400.

Wagner, M. (2020). Affective polarization in multiparty systems. *Electoral Studies*.