Are citizens vigilant of changing the balance of power?

Testing evaluations of checks and balances reforms

Honorata Mazepus

ORCID: 0000-0002-5005-9399

Department of Political Sciences, University of Amsterdam

Nieuwe Achtergracht 166, Postbus 15578

1001 NB Amsterdam, Netherlands

h.mazepus@uva.nl

Abstract: Democratic checks and balances are crucial accountability mechanisms that prevent the abuse of power and protect citizens. Thus, citizens committed to democratic principles should be vigilant of reforms changing the balance of power. Using a mixed-design survey experiment with Polish citizens, this article investigates whether citizens are more vigilant of reform proposals that alter democratic checks and balances than other reforms. It also tests the effects of party cues on the intentions assigned to reformers proposing to alter democratic rules. The results show that (1) the in-party supporters evaluated the intentions to reform checks and balances as benign, whereas the out-party supporters were more vigilant and saw them as malicious; (2) participants were less vigilant of an administrative reform and evaluated it as less ill-intended; (3) these effects held across supporters of different political parties – both incumbent and opposition. Overall, citizens are rather vigilant of changing democratic checks and balances.

Keywords: party cues; checks and balances; epistemic vigilance; motivated reasoning, intentions; reforms; survey experiment; democratic erosion; Poland

Ethics approval: This research was reviewed and approved by the Ethics Board of the Faculty of Governance and Global Affairs at Leiden University, approval number 2021-011-ISGA-Mazepus. The consent form used in the study is available in the Appendix Online (AO1).

Introduction

The processes of undermining democratic checks and balances have been reported across new and old democracies (Lührmann et al., 2020). Elected political actors with authoritarian tendencies use reforms to get rid of limitations on their power: they expand the executive competences, reform courts, change electoral rules, and alter laws on political parties, protests, and the media. While the declared levels of citizens' support for democracy is stable and relatively high (Anderson et al., 2021; Voeten, 2016) and most citizens in Europe have a rather coherent democracy belief system (Hernández, 2019), many recent studies questioned whether citizens can be the guardians of liberal-democratic institutions. Research showed that citizens endorse reforms eroding the balance of power (Şaşmaz et al., 2022), that their understanding of democracy is not stable (Bryan, 2023; Krishnarajan, 2023), and that they trade off democratic rules for partisanship (Littvay et al., 2024; Simonovits et al., 2022). But is this an accurate representation of citizens' relation to democratic institutions?

So far, most studies in democratic contexts have focused on the effects of democratic transgressions by politicians on voting preferences of citizens. Candidate choice experiments showed that both partisanship and receiving preferred policies decrease the punishment of democracy-undermining behavior when it comes to voting (Frederiksen, 2024; Graham & Svolik, 2020; Saikkonen & Christensen, 2023). Vote-choice, however, does not tell the full story about the evaluations of transgressions (Aarslew, 2023; Jacob, 2024) and cannot be equated with support for altering the institutions of liberal democracy. Building on the rich body of findings, this study takes the research agenda further and proposes to look beyond the candidate-choice and to incorporate epistemic vigilance theory to better understand the volatility of citizens' commitment to democratic principles. This article integrates the theory of epistemic vigilance (Sperber et al., 2010) from cognitive

psychology with partisan motivated reasoning theory (Taber & Lodge, 2006) to test what drives the evaluations of reforms potentially undermining democratic institutions. Epistemic vigilance is a set of cognitive mechanisms that helps us avoid deception and harm. Epistemic vigilance and partisan motivated reasoning theories make mostly similar predictions about the political behavior, but they provide different explanations for it: the former proposes that in-group favoritism is a consequence of (reasonable) vigilance, while the latter frames it more in terms of bias. The article aims to integrate these theories and provide a complementary explanation of citizens' support for democracy by focusing on the vigilance toward politicians' *intentions* (Braley et al., 2023).

Specifically, the article investigates *the vigilance* of reforms altering the balance of power. It focuses on two triggers of epistemic vigilance, namely *who* proposes such reforms and *what* reforms are being evaluated. First, it probes the partisanship trigger by focusing on the *perceived intentions* attributed to politicians proposing institutional reforms. The tested proposition is that citizens assign different intentions (benign or malicious) to the reformers of institutions depending on their coalitional (party) allegiance. Second, it tests the effect of *the stakes* involved in the proposed reforms and the effect of party cues on evaluations of different types of reform proposals: reforms undermining different types of checks and balances (high stakes reforms), and a procedural administrative reform (an example of a low stakes reform.)

Empirically, the article presents the results of a pre-registered[1] original survey experiment with a quota-based sample of Polish citizens (N = 1040). In a mixed design, each participant was randomly assigned to a party cue (in-party, no party, or out-party) and evaluated four scenarios about

---

[1] See pre-registration at: https://osf.io/tc2k8/?view_only=60c0d64fe613482287657149df6e7e33

proposed reforms in random order. Participants indicated perceived intentions (benign vs. malicious) of the reformers for checks and balances reforms (changes to the appointments of judges, removal of the presidential term limit, and new rules for protest registration) and a reform changing the deadline for income-tax declaration. The experiment uses a realistic scenario that provides a valid justification of a reform that potentially undermines democracy.

The analysis showed that party cues had consistent effects on how participants evaluated the intentions of the reformers: if the reform was proposed by the out-party, participants assigned more malicious intentions, if by the in-party, participants assigned more benign intentions. These party-cue effects were present also in the case of the reform with lower stakes. However, the study found a clear difference in how citizens evaluated reforms of checks and balances and a rather inconsequential administrative reform.

This study offers new insights about the causal relation between partisanship driven vigilance, intentions assigned to the reformers, and support for different types of reform proposals to further refine the theory of democratic attitudes of citizens. Overall, citizens seem much more vigilant of the reforms of checks and balances and evaluate such reforms more negatively. Importantly, even controlling for the specific party support, altering checks and balances is not seen as desirable. This indicates that epistemic vigilance is triggered when politicians initiate reforms endangering democratic rules, and that overall, citizens support these reforms less than low-stakes reforms. At the same time, motivated reasoning leads to different evaluations of the intentions of the reformers and different levels of support for the reforms depending on the party allegiance.

Epistemic vigilance

Epistemic vigilance is a set of cognitive mechanisms that helps us distinguish harmful from beneficial information and reduces our susceptibility to persuasion (Mercier, 2017). The theory of epistemic vigilance proposes that our minds are equipped with such mechanisms that are activated by the relevant information from the environment (Sperber et al., 2010). These mechanisms make us more sensitive to and weary of particular information, and in the context of politics, they should make us watchful of the behavior and *intentions* of politicians who propose important reforms. Epistemic vigilance "(unlike distrust) is not the opposite of trust, it is the opposite of blind trust" (Sperber et al., 2010). More specifically, epistemic vigilance can be best seen as a mechanism that leads to mistrust (Jennings et al., 2021), skepticism (Norris, 2022), and evaluative trust (De Blok et al., 2022), and therefore acts as a shield against gullibility (Mercier, 2017).

The mechanisms of epistemic vigilance should be activated when citizens interact in their socio-political environment and consider political proposals. As such, epistemic vigilance can be a good departure point for explaining citizen evaluations of reform proposals undermining democratic checks and balances, by zooming in on the *perceived intentions* of the reformers: harmful or benign. This study focuses on two triggers of epistemic vigilance in the context of checks and balances reforms: party cues and the relevance (stakes) of the proposed reforms.

Reforms undermining democratic checks and balances and epistemic vigilance

In most scenarios of democratic erosion, authoritarian political actors come to power in free and fair elections and gradually undermine liberal-democratic institutions (Lührmann & Lindberg, 2019). Those actors target different democratic checks and balances – institutions and mechanisms that help to control those in power and keep them accountable (Bermeo, 2016).

Following Lührmann, Marquart & Mechkova (2020), we can differentiate between three types of checks and balances. Vertical ones pertain mainly to electoral processes that make representation of citizens possible and should allow voting poorly performing politicians out of office. Horizontal ones are the constraints imposed by different state institutions on each other (parliament, government, president, courts, term-limits, and specialized agencies). Diagonal ones are mechanisms allowing citizens to monitor and criticize the government, and mobilize opposition in-between elections (e.g. media, protests, social movements, and civil society organization). The democratic erosion within newer and older democracies affects especially the horizontal and diagonal checks and balances (Boese et al., 2022), which will be the focus of the experiment.

Overall, evidence shows that citizens attribute high importance to parliamentary and judicial oversight, media freedom, and protest rights (see, e.g. Mazepus & Toshkov, 2022). Therefore, in principle, it should be difficult to convince citizens that a reform of checks and balances is desirable. Epistemic vigilance theory helps us understand under what conditions individuals could accept the arguments in favor of such reforms. First, reforms of checks and balances are not among the main issues of political campaigns in democracies. Rarely are they at the center of the political battle and if so, then never presented with arguments in support of destroying democracy. In fact, even authoritarian leaders use democratic arguments to justify their transgressions (see Matovski, 2021). This means that the proposals might sound like they are in line with democratic rules and hence align with pre-existing beliefs of citizens (Mercier, 2017).

However, epistemic vigilance facilitates evaluation of other aspects of information as well. The pre-existing assumptions about the source of information is likely to play a role, especially in the context of politics. The information about the reasons for reforms proposed by the incumbents could be interpreted differently because of the assumptions about the intentions of the "allies" and

"opponents" – those who have our interests at stake and those who do not. Our prior beliefs about the actor shape the evaluation of the actor's proposals and the justifications they provide for them. This mechanism constitutes a defense system against being misled or abused, especially important in complex information environments. In the domain of politics, partisanship can act as a cue of allegiance and activate such vigilance towards received information. Therefore, even when individuals value democratic rules and support specific checks and balances, their attitudes towards reforms might fluctuate depending on who proposes the reforms and whether they perceive them as a threat to their interest.

Assuming that in general citizens are supportive of democratic rules, proposals of reforms affecting the balance of power should activate epistemic vigilance and spark two main questions in the minds of citizens: "who to trust?" and "what to believe?" (Mercier, 2020). Thus, the relevant preexisting beliefs when citizens face information about reforming political institutions are likely to be the estimation of the intentions of the actor proposing them (who to trust), and the relevance of the issue and arguments about the potential effects of the reform (what to believe).

Who to trust: estimating the intentions of the actor proposing reforms

Since epistemic vigilance is a set of mechanisms for avoidance of harm and deception, deciding who constitutes a trustworthy source of information involves a judgment about who is likely to have our interest at stake (or, in other words, whose interests align with ours; Mercier 2020, p. 92) and unlikely to harm us (Levi & Stoker, 2000). In the context of politics, harm can come in different forms, for example undesirable policies (e.g., allowing more migrants to enter our country when we oppose it) or exclusion (e.g., not receiving benefits or outright discrimination). While it is highly unlikely that in the current political environment any proposed reform of checks and balances would be presented as undermining democracy, citizens can read the intentions behind such reforms

7

and their potential consequences in various ways. Therefore, the mechanisms of epistemic vigilance should be triggered to pay attention to *who* is proposing the reform and evaluate whether the reformer can be expected to benefit or harm the citizen. In other words, can one trust that the reformer's intentions are good. Because citizens and politicians rarely interact directly with each other, political identity markers are an important cue for who to trust.

Attachments to social identity groups like political parties or ethnic groups were found to drive citizens' evaluations of information and their perceptions of objective facts (e.g. Leeper & Slothuus, 2014). Political information, often very complex and difficult to grasp, rarely comes without partisan or group identity cues. Most of information includes signals of partisanship such as names of politicians, their function, and/or their party affiliation. It has been shown that our mental alliance-detection system is geared towards cues such as party affiliation (Pietraszewski et al., 2015), which help us navigate complex political choices. Following from this, people are likely to evaluate policy and reform proposals differently depending on whether they are proposed by a party they identify themselves with (see as having their interest in mind) or a party they oppose (see as not aligned with their interest). This gap in evaluations depending on the political affiliation is similar to the winner/loser gap in satisfaction with democracy, endorsement of elections, and assigning importance to the liberal aspects of democracy including checks and balances (Anderson et al., 2005; Hansen et al., 2019; Mazepus & Toshkov, 2022).

Different evaluations of policy proposals can result from political orientation or policy preferences that simply align with a supported party (Krishnarajan, 2023), but they can also be driven by motivated reasoning involving an identity cue (Bayes et al., 2020; see also Petersen et al., 2013) – wherein the information about the identity of the actor responsible for a policy determines our attitude towards the policy. This means that the perceived reasons for introducing a policy or reform

8

could depend on who proposes it, activating either motivated reasoning or motivated skepticism (Taber & Lodge, 2006). However, what in the motivated reasoning logic would be seen as 'my group bias' could also be seen as (justified) vigilance of *intentions* of others: people in our team have benign intentions (they have good reasons to pursue a particular action), while the intention of those on the opposing team are malicious (Sloman & Rabb, 2019).

As a consequence, if an individual sees their interest as aligned with the political party that puts a proposal forward, they will assign more benign intentions to their reform proposals and trust democratic reasons for such reforms: e.g., ensuring independence of judges, strengthening democracy and the rule of law. If an individual does not see the alignment of interest, they will assign more malicious intentions to their reform proposals: undermining democracy, trying to get rid of the constraints on power. As reforms of checks and balances are potentially more harmful to the out-party supporters, they should be more vigilant of these proposals, more sensitive to the intentions of the reformers, and more likely to attribute malicious motivations to them. Thus, following the motivated skepticism logic, we can formulate a general expectation that a cue of out-group party makes citizens more vigilant of the intentions behind the reforms of checks and balances. More specifically, we can formulate hypotheses about the specific perceived intentions of the reformers:

H1a: A cue of in-group party increases the perception of <u>benign intentions</u> behind a reform altering the system of checks and balances (by comparison with no party cue and out-group party cue).

H1b: A cue of out-group party increases the perception of <u>malicious intentions</u> behind a reform altering the system of checks and balances (by comparison with no party cue and in-group party cue).

What to believe: the stakes of the proposed reform

Epistemic vigilance comes with processing and social costs and, to be efficient, it should be activated for information that is relevant to the receiver and not for any kind of information (Sperber et al., 2010). Epistemic vigilance should be activated especially in high-stake (relevant) situations, i.e., when the consequences of accepting (wrong) information are serious (Sperber et al., 2010). This is consistent with the role of personal relevance played in the assessment of information in the Elaboration Likelihood Model: relevance (stakes) increases if the consequences for our life are significant (Petty & Cacioppo, 1986, p. 145).

Different reforms can have different consequences and therefore are of different relevance. Avoiding domination and exploitation has always been one of the main concerns of individuals when delegating power to others (Boehm, 1999). Democratic checks and balances are crucial accountability mechanisms that prevent the abuse of power and assure equal treatment in front of the law and the rights that allow contestation of incumbents. Thus, reforms changing the balance of power are high stake reforms that can harm citizens, and therefore they should be especially vigilant of them. Moreover, high stakes in the context of political reforms can mean differential effects of reforms on individuals that align themselves with the incumbents and on individuals who align themselves with the opposition[2]. Just like proposed policies on issues such as immigration, health care, or taxes have differential effects on citizens with different interests, reforms altering checks and balances differentially affect citizens.

---

[2] This is true especially in the short run. Long-term, dismantling of democratic checks and balances is harmful to most citizens.

First of all, when citizens are in the opposition to the government, they are more vulnerable to the potential abuses of power and therefore information that the checks and balances will be altered should be more relevant to them. These citizens have a reason to worry that such changes will weaken controls on the executives, which can have negative consequences for them personally. This aligns with the evidence that checks and balances might be especially important to those on the losing side of elections (Mazepus & Toshkov, 2022). Hence, we can expect that the stakes of altering the system of checks and balances are higher (the relevance is stronger) for those who are not supporting the current government and expect them to have lower support for the reform:

H2: *A cue of an out-group party proposing the reform of checks and balances increases <u>negative evaluation</u> of this reform (by comparison to a cue of an in-group and to a no party cue)*[3].

Furthermore, epistemic vigilance mechanisms should be activated more strongly when the reforms are relevant in general. In the context of politics, this would pertain to the reforms with potentially differential effects on citizens and to important issues (e.g., migration, welfare system, or political system). As a consequence, not everything (not every political decision) should trigger vigilance and not everything proposed by the out-group should be automatically categorized as malicious, but only particular proposals that could harm the position of a citizen. This is consistent with the available evidence about no effects of party cues on 'consensus issues' (Slothuus & De Vreese, 2010).

---

[3] The numbers of hypotheses have been adjusted by comparison to the pre-registration to align with the structure of the paper. 'Signal' was replaced with 'cue' in the phrasing in this article and 'coalition' with 'party' as these are more accurate expressions. H3a was rephrased without changing the intention of the hypothesis (see AO12).

Moreover, citizens should be in general more weary of reforms that could undermine democracy than less relevant reforms, irrespective of who is proposing such a reform.

Therefore, this study tests a general expectation that citizens respond differently to the reforms altering checks and balances than to administrative (non-political) reforms and, specifically, that citizens are more divided about the intentions behind the reforms of checks and balances than behind the administrative one.

More precisely, H3 proposes an interaction between the party cue and specific reform, namely that

H3: The effect of party-cue on assigned intentions is stronger for the reforms with higher stakes (checks and balances) than reforms with lower stakes (administrative reform).[4]

Methods: survey experiment with mixed design

To test the causal relations specified in the hypotheses, this study employed a novel survey experiment. Using experimental design avoids the problem of relying on self-report when researching democratic values of citizens. This seems especially relevant when researching checks and balances and the rules of law (cf. Kaftan, 2024). The study uses original data from Poland, a relevant case to study attitudes of citizens with experience of democratization and democratic erosion. Poland used to be considered a model-example of transition from communist authoritarianism to liberal democracy in the region, but has experienced democratic backsliding since the Law and Justice party won elections and formed the government in 2015 (Karv, 2022, p. 61). The government supported by the parliamentary majority and the president from the same party

---

[4] See AO12 for the list of hypotheses.

has continuously undermined the authority of the Constitutional Tribunal, independence of judges, women's rights, separation between the church and the state, and the quality and objectivity of the public media. The main attacks over the two terms in government were launched at the horizontal and diagonal checks and balances.

Consequently, altering checks and balances is a salient issue in Poland, where many arguments for and against changing institutions are presented by politicians and the media. The presence of such reforms in public debate makes it more likely that Polish participants will be able to imagine the scenarios of reforms (by comparison to participants from contexts where such reforms are never debated or introduced) and contributes to the ecological validity of the study. In addition, although in multi-party systems political polarization tends to be lower than in two-party systems such as the USA and the UK, the party preferences of Polish citizens are relatively highly polarized (Bettarelli et al., 2023). Therefore, Poland offers the most-likely case to detect the my-party bias in a multi-party context.

At the same time, the salience of the reforms might make it more difficult to convince participants about which party (in-party or out-party) is proposing reforms of checks and balances and might make participants in the control condition associate the reforms with the party who has been responsible for democratic backsliding in the last years. To account for this problem, two measures were taken. First, the scenarios present arguments for reforms that are not obviously anti-democratic (elite rhetoric is rarely openly anti-democratic) making it more plausible that any party proposes them. For example, 'getting rid of corrupt judges' as the aim of judicial reform could be imagined as a proposal by the opposition party Civic Platform to reverse the reforms introduced by the Law and Justice party. Second, the study used an elaborate comprehension check (discussed below) to make sure participants pictured the right political party responsible for the reforms. Importantly, the study

used scenarios with realistic representation of the reforms where the governing party offers an argument in favor of the reform, but the most direct consequence of the reform meant to undermine democratic checks and balances is also presented. In addition, the first sentence of the vignette makes sure that there are no doubts about the validity of the elections and that all participants imagine a government with electoral legitimacy as a baseline.

The experiment used a mixed design: it combined a between- and within-subject manipulation (see Appendix Online AO2 for the survey flow). Participants of the study were randomly assigned to one of three *party cue* conditions: control (no party cue), 'in-party' (cue that the party they support governs), or 'out-party' (cue that the party they oppose governs). Each participant evaluated four *reform* vignettes: two of horizontal checks and balances (reform of the appointment of judges and reform of the presidential term), one of diagonal check and balance (reform of protest registration rules), and one administrative-type reform (tax statement deadline change). The vignettes about different reforms were presented to participants in a random order.

An example vignette – about the reform of registration of protests – reads as follows:

> Imagine that **one of the political parties in your country/a political party you support and feel very close to/ a political party you oppose and do not feel close to at all** wins the majority at a fully competitive and fair election. This means that the party gets to form the government and implement various policies.

One of the ministers proposes **to introduce new law for registration of protests**. As a result, it will be more difficult to organize a protest. The minister informs that this reform is needed to assure safety of citizens' gatherings. The government supports this plan.[5]

After reading each vignette, participants were asked to answer three questions assessing the reception of the manipulation and the comprehension of the vignette text[6].

The survey captured vigilance indirectly by measuring the perceived intentions behind the reform proposals. The measures of perceived intentions are new and were developed for the purpose of this study. After reading each reform vignette, participants assessed the intentions of the reformers by answering 'What is the most likely motivation of this government behind introducing such a reform?'. They answered two questions measuring *benign intentions*, one more specific and one more general: 'The government proposes this reform to assure that citizens can protest in a safe way' and 'The government proposes this reform to strengthen democracy' and two questions measuring *malicious intentions* "The government proposes this reform to have fewer limitations on their power, so the government can do whatever it wants' and "The government proposes this reform to undermine democracy'. These questions were adjusted to match the reforms described in each vignette (for the full list of vignettes and questions, see AO3 & AO4) and measured on the scale

---

[5] The vignette is based on the experimental study in [blinded for review]. The current version of the vignettes about the reform of appointment of judges and the change of the tax statement deadline was pre-tested in a pre-registered pilot study with N = 206 students at [blinded for review] (preregistration available here: https://osf.io/5ydzr/?view_only=1df5871eb60542039a7e59ae4776bc44).

[6] The results of a soft launch of the data collection revealed that a large percentage of participants are not reading the vignette text, are unable to answer the comprehension check questions correctly, and therefore the manipulation is not received by a large proportion of participants. Because of this an alternative strategy than the pre-registered one was chosen to increase the attention, comprehension and thus reception of the manipulation. For full discussion of the used strategy, see the appendix AO6.

from 1- 'Very unlikely' to 7- 'Very likely'. To measure the *support for the reforms*, participants answered

two questions capturing the normative and personal support respectively in response to each

vignette: 'To what extent would you find such a reform justified? (from 1- 'Not at all justified' to 7-

'Fully justified')' and 'To what extent would you agree with such a reform? (from 1- 'Fully disagree'

to 7- 'Fully agree')'. Most of the answers to these pairs of questions were strongly correlated

(between r = .61 and r = .93). The scale for the support of the reform was highly consistent and

reliable (Cronbach's alpha between 0.93 and 0.96). The scale reliability for the new measures of

intentions was acceptable to high (six out of eight scales had ah Cronbach's alpha above 0.80, see

AO4). Following the pre-registration, I created indexes averaging the scores for each of the pairs of

answers (see AO4 & AO7) to more reliably measure the dependent variables.

Data

The data for this study was collected in Poland in November-December 2021 by the survey

company *Respondi* (currently *Bilendi*). The presented study was included in a questionnaire together

with another study, which followed after the experiment and the main measures. Participants

completed the survey online in Polish. The quotas on gender, age, and education were applied to

reflect the Polish urban population. The total sample of N = 1040 was collected[78]. The sample was

52% female, with average age of M = 44 (SD = 14.3), good distribution of participants across the

---

[7] I conducted power analysis for ANOVA Repeated measures using G*power. I converted effect size f to eta squared with the formula $\eta^2 = f^2 /(1 + f^2)$. The study had power = .90 to detect between-subject and within-subject effects of $\eta^2 = 0.01$ or larger. The study had insufficient power to detect interaction effects smaller than $\eta^2 = 0.02$.

[8] Due to a programming mistake by the survey company, I have collected additional 350 responses. The additional data show very similar patterns to the data presented in the main text. For the discussion of the mistake, see appendix AO12. For the report of the additional sample analysis, see appendix AO14.

age groups, and minimum of 18 and maximum of 69 years old. The implication of this sampling strategy is that the sample does not represent the oldest citizens of Poland who are the most difficult to access online and constitute around 15% of the Polish population. In terms of political party support, participants were distributed well across the major political parties (25% voted for Law and Justice and 26% for the Civic Platform in the last elections), though by comparison with the population, voters of Law and Justice are underrepresented. The sample overrepresents citizens with higher education. For further description of the sample and measures of ecological validity, see appendix AO5.

After participants completed another study (on unrelated topic), they answered questions measuring their attitudes towards democratic institutions. Overall, and in line with previous studies, the participants expressed high level of support for different elements of democracy (see Table 1).

Three of the items measuring attitudes toward democracy are related to the checks and balances manipulated in the experimental scenarios: judicial independence, right to protest, and presidential term limit. As the mean scores are very high, we can assume that participants in our sample value these rules and should, in principle, be vigilant of their violations. However, the voters of different parties value checks and balances to a different extent: the voters of Law and Justice value all checks and balances the least of all voters and significantly less than the voters of Civic Platform (Table 1)

Table 1. Importance of different checks and balances measured with the question 'Please indicate to what extent you agree or disagree with the following statements from 1- 'Fully disagree' to 7- 'Fully agree". N = 1040. The total sample is N = 1040 (PiS = 260, PO = 273, Other = 269, No vote = 238).

Checks & balances	Question	M	SD
Courts' fairness	It is important that courts treat everyone the same	6.25	1.39

	PiS (Law and Justice)	6.09	1.44
	PO (Civic Platform)	6.48	1.20
	Other	6.26	1.40
	No vote	6.14	1.49
Free & fair elections	Citizens should have the right to pick their political representatives in free and fair elections	**6.32**	**1.27**
	PiS (Law and Justice)	6.20	1.28
	PO (Civic Platform)	6.56	1.04
	Other	6.28	1.34
	No vote	6.23	1.40
Freedom of speech	It is NOT important that opposition parties are free to criticize the government	**2.71**	**1.90**
	PiS (Law and Justice)	3.44	1.91
	PO (Civic Platform)	2.08	1.74
	Other	2.47	1.79
	No vote	2.89	1.90
Judicial independence	Judges should be independent from elected politicians	**6.19**	**1.40**
	PiS (Law and Justice)	5.75	1.61
	PO (Civic Platform)	6.56	1.04
	Other	6.32	1.28
	No vote	6.10	1.53
Media freedom	It is important that the media are free to criticize the government	**5.87**	**1.52**
	PiS (Law and Justice)	5.31	1.67
	PO (Civic Platform)	6.32	1.23
	Other	6.09	1.37
	No vote	5.70	1.61

Judicial oversight	It is NOT important that the courts are able to stop the government acting beyond its authority	**2.56**	**1.89**
	PiS (Law and Justice)	3.35	1.92
	PO (Civic Platform)	2.05	1.81
	Other	2.33	1.82
	No vote	2.54	1.76
Democracy overall	Democracy has its flaws, but it is the best system of governance invented so far	**5.45**	**1.65**
	PiS (Law and Justice)	5.38	1.52
	PO (Civic Platform)	6.07	1.39
	Other	5.40	1.71
	No vote	4.89	1.77
Right to protest	Citizens have to be able to express their dissatisfaction with political developments through protests	**6.02**	**1.42**
	PiS (Law and Justice)	5.55	1.53
	PO (Civic Platform)	6.44	1.14
	Other	6.18	1.25
	No vote	5.84	1.57
Presidential term limit	It is important that the post of the president is not held by one person for more than two terms	**5.42**	**1.84**
	PiS (Law and Justice)	4.86	1.95
	PO (Civic Platform)	6.01	1.57
	Other	5.45	1.83
	No vote	5.32	1.81

Analysis

Between 66% and 72% of participants passed all three manipulation and comprehension checks across the experimental conditions (see AO6), therefore we can be confident that around two thirds

of participants received the manipulation as intended (see Mutz & Pemantle, 2015). The analyses

presented here include the full sample (N = 1040)[9]. To assess whether participants could imagine the

proposed scenario of reforms in the in-party condition, I used an ecological validity check question

(see AO5). For the voters of Law and Justice and Confederation it was the easiest to imagine that

their party proposes a reform altering the appointment and dismissal of judges. For the voters of

The Left and Civic Platforms it was the most difficult to imagine. However, the results indicate that

even for The Left and Civic Platform voters these reform proposals were within the scope of

imaginable (somewhere around the mid-point of the 7-point scale).

As specified in the pre-registration, to test the hypotheses, I performed three mixed-design

ANOVAs with 'party cue' as between-subject factor and 'reform' as within-subject factor for the

three following outcome variables: support for the reforms, benign intentions attribution, and

malicious intentions attribution.

Support for the reforms

The first analysis focused on the effects of party cue and reform type on support for such reforms.

The results of the mixed design ANOVA showed that there was a statistically significant effect of

the party cue on the reform support, $F(2, 1037) = 17.03$, $p < .001$, $\eta^2 = .03$ (AO10). Figure 1

illustrates the results and shows the distribution of data and the main effect of party cue. Pairwise

comparisons (see AO10) confirmed these patterns of data and showed that participants who

received the 'in-party' cue were more supportive of the reforms than participants who received the

'out-party' cue ($M_{diff} = 0.65$, $p < .001$). Furthermore, participants who received the 'out-party' cue

---

[9] All data and code needed to replicate this analysis will be available upon publication.

differed in their evaluation of the reform from those who were in 'no party' condition only when evaluating the reform of appointment of judges (see Table 15A in AO10 and marginal means plot in AO11). There was no significant difference between 'no party' and 'out-party' conditions in their average evaluations of all the other reforms. Participants in the 'in-party' condition are different from those in the other two conditions and expressed significantly higher approval of the reforms of presidential term, protests, and tax statement deadline.

Thus hypothesis (H2) that a cue of out-party increases negative evaluation of the reform of checks and balances (by comparison to a cue of in-group and no party cue) is supported only by the results for the evaluation of the reform of the appointment of judges. In contrast to H2, the results for the evaluation of all the other reforms suggest that it is the cue of an in-party that drives the support for the reforms up.

It is important to note that on average the support for checks and balances reforms is below the mid-point (4) on the 7-point scale. The only reform that received a neutral to positive average evaluation is the reform of the tax statement deadline: it received a significantly higher level of support than any of the checks and balances reform. Moreover, the distribution of data clearly shows a larger concentration of negative scores for the evaluations of the checks and balances reforms than for the evaluation of the tax statement deadline reform (Figure 1).

Figure 1. Violin plots show the support for reforms across conditions of party-cue (in-party, no party, and out-party). The white dot shows the median, the dark bar shows the interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density. Higher values represent higher support.
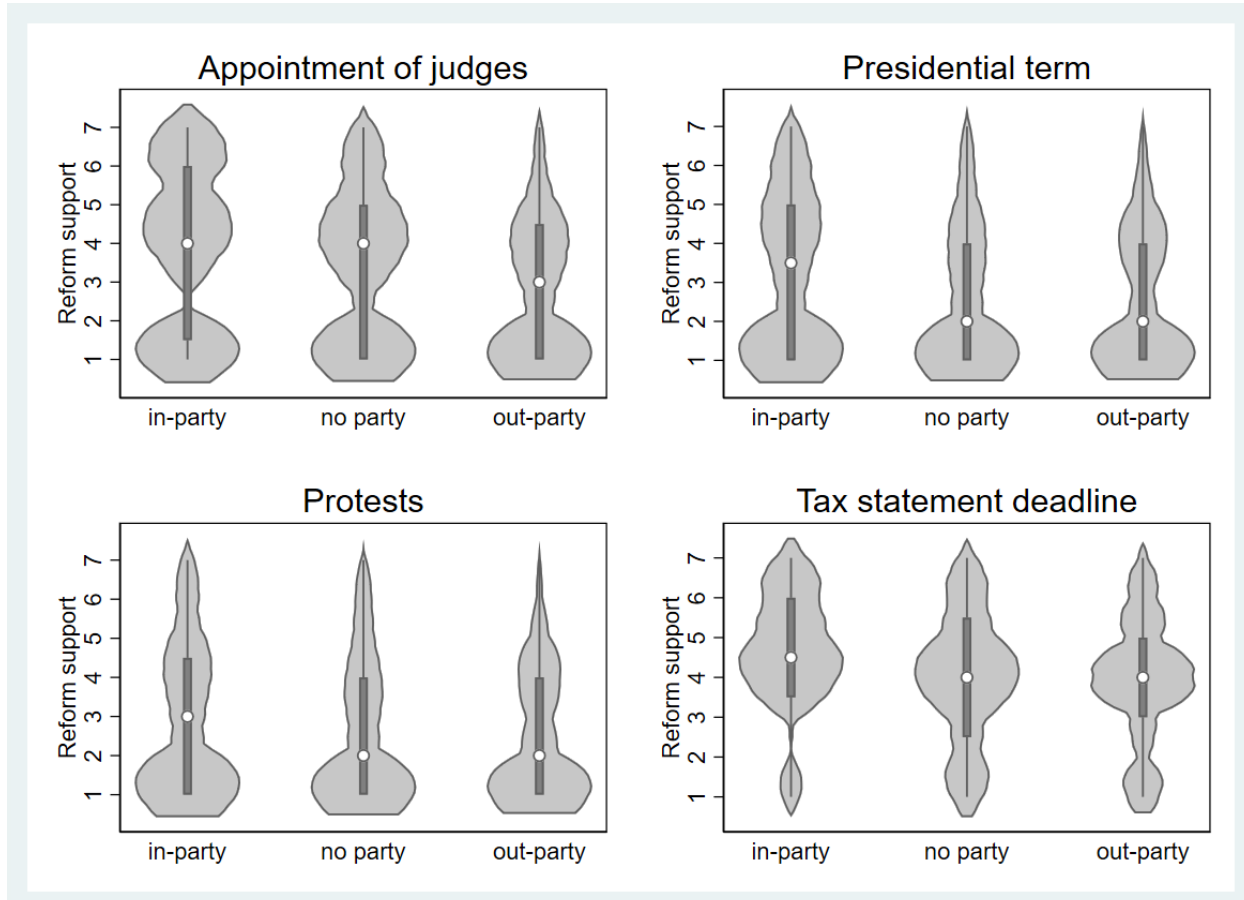
Tests confirm these observed patterns. There was a significant within-effect[10] of the reform type on the support for the reform, $F(2.830, 2934.43) = 166.79$ $p < .001$, $\eta p^2 = .14$. The reform of the tax statement deadline was evaluated significantly more positively than any of the checks and balances reforms by participants (from $M_{diff} = -.612$ by comparison with the reform of appointment of judges to $M_{diff} = -1.26$ with the reform of protest). The pairwise comparisons showed that there are also significant differences in support between the different checks and balances reforms: the

---

[10] I present the results of the Huynh-Feldt estimate for the within-subject and interaction tests to correct for the violation of the sphericity assumption (tested with the Mauchly test of sphericity). The Greenhouse-Geisser estimate is greater than .75 for all three dependent variables, therefore a less conservative correction can be used (see Field, 2013, p. 548). All test are presented in AO10.

reforms of protest registration (M_{diff} = -652, p < .001) and of the presidential term limit (M_{diff} = -534, p < .001) received significantly less support than the reform of appointment of judges (see A09).

In addition, the analysis also showed a significant interaction between party cue and reform type ($F(5.659, 2934.43) = 3.42$, p = .03, ηp^2 = .01). In the case of the reform of the tax statement deadline the difference between those in the 'in-party' and 'out-party' party was smaller than in the case of the other reforms (M_{diff} = .43, p = .002). For comparison, the difference between those in the 'in-party' and 'out-party' condition in support for the reform of the appointment of judges was M_{diff} = .93, p < .001, for the reform of the presidential term M_{diff} = .67, p < .001, and for the reform of protests M_{diff} = .55, p < .001.

Overall, these results confirm the general expectation that citizens respond differently to the reforms altering checks and balances than to administrative (non-political) reforms. The within-tests show that there is a difference in support for the checks and balances reforms by comparison with the support for the administrative reform. Moreover, while the party cues have the same patterns of effects, the difference between the 'in-party' and 'out-party' is smaller when evaluations pertain to the low-stakes reform.

Benign intentions attribution

Turning to the analysis of perceived intentions, the results of mixed-design ANOVA showed that there was a statistically significant effect of the party cue on the level of benign intentions assigned to the reformers, $F(2, 1037) = 12.34$, p < .001, ηp^2 = .02. The graphs in Figure 2 show that participants who received the 'in-party' cue assigned consistently more benign intentions to the reformers than participants in the condition without a party cue and participants who received an

23

'out-party' cue. Pairwise comparisons revealed that across the reforms the difference between the 'in-party' condition and the rest is larger (and always significant) than between the 'out-party' condition and the neutral 'no party' condition (not significant in 3 out of 4 cases). Only in the case of the reform of the appointment of judges the difference between the opponents and those in the neutral condition meet the standard significance level with M_{diff} = .30, p = .05. These results are in line with the hypothesis (H1a) that a cue of in-group party increases the perception of benign intentions behind a reform altering the system of checks and balances (by comparison with 'no-party' cue and 'out-party' cue).

There was also a significant effect of the reform type on the level of benign motivations assigned to the reformers, $F(2.727, 2827.61)$ = 190.51, p < .001, ηp^2 = .16. The results show that participants on average assigned more benign intentions to the reformers when evaluating the reform of the tax statement deadline (M = 4.33, SE = .05) than when evaluating checks and balances reforms.

Overall, participants were rather skeptical of reforms of checks and balances and on average did not agree that the reformers introduce them to improve the functioning of democracy and to enhance the balance of power: the level of assigned benign intentions is below the mid-point (4) of the 7-point scale. The intentions behind the reform of the protest registration (diagonal check on the power) were evaluated as least benign (M = 2.91, SE = .06) and participants attributed significantly less benign intentions to the protest reform than to the appointment of judges (M_{diff} = .46, p < .001) and the reform of the presidential term limit (M_{diff} = .44, p < .001), whereas the difference between the benign intentions assigned to the reforms of the appointment of judges and of the presidential term was not significant (see Table 18A, AO10).

Figure 2. Violin plots show the attribution of benign intentions across conditions of party-cue (in-

party, no party, and out-party). The white dot shows the median, the dark bar shows the interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density. Higher values indicate attribution of more benign intentions.
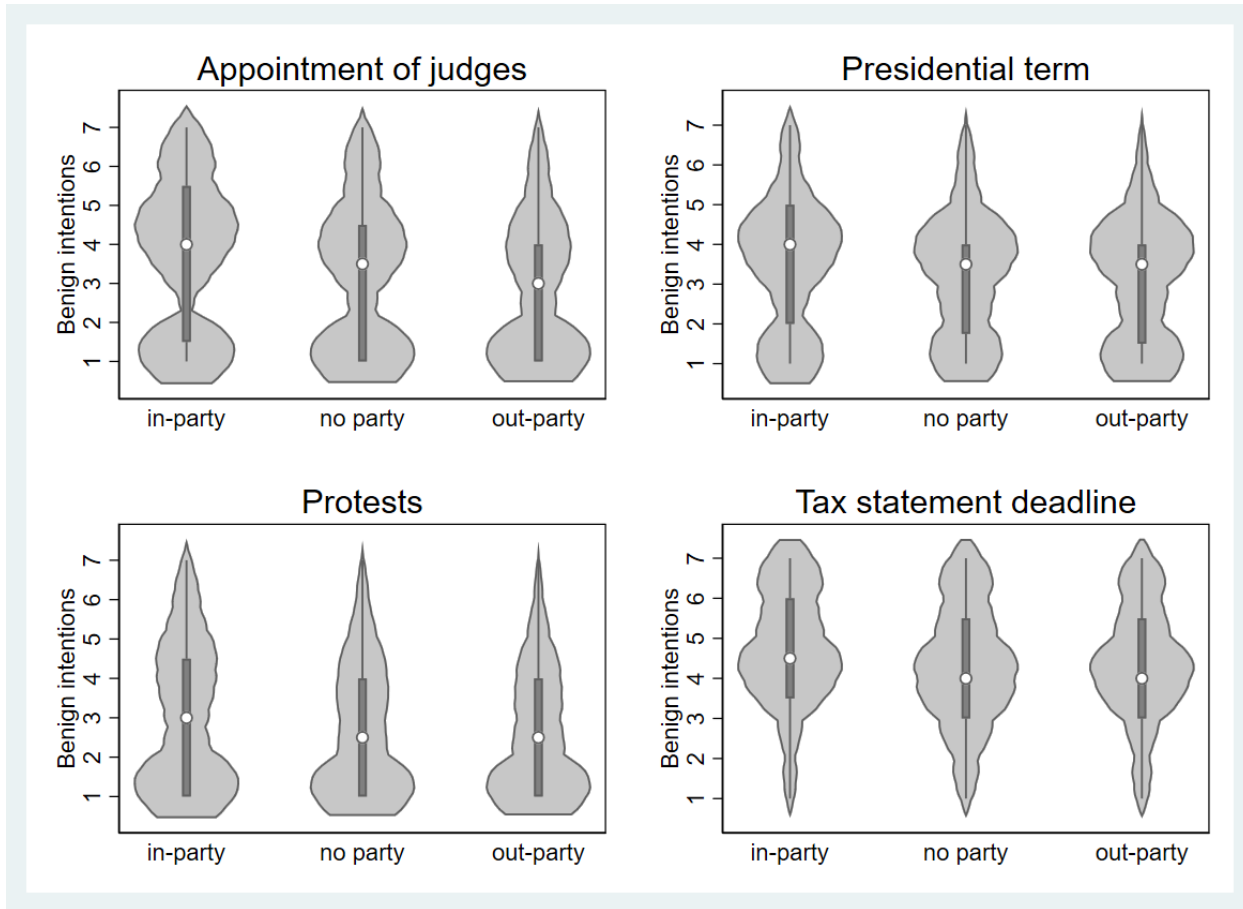
The difference between the participants in the out-party condition and the in-party condition when evaluating the reform of the tax statement deadline was the smallest across all reforms ($M_{diff} = -.30$, $p = .022$). By comparison, this difference was the largest for the evaluation of intentions attributed to the reform of the appointment of judges ($M_{diff} = -.73$, $p < .001$). However, the interaction between the party cue and the reform type was not significant ($F(5.453, 2827.61) = 1.56$, $p = .163$, $\eta p^2 = .003$). Hence, we find no evidence for H3 that the effect of party-cue on assigned intentions is stronger for the reforms with higher stakes (checks and balances) than reforms with lower stakes

(administrative reform).

Malicious intentions attribution

The results of the mixed design ANOVA showed that there was also a statistically significant effect of the party cue on the level of malicious intentions assigned to the reformers, $F(2, 1037) = 12.60$, $p < .001$, $\eta p^2 = .02$. The graphs in Figure 3 show that participants who received the 'out-party' cue assigned consistently more malicious intentions to the reformers than participants in the condition without a party cue and participants who received an 'in-party' cue. Pairwise comparisons revealed that for the reform of the appointment of judges and registration of protests, the difference between the 'out-party' condition and the rest of participants is larger than between the participants in the 'in-party' and 'no party' conditions . Participants in the 'out-party' condition attributed significantly more malicious intentions to the reform of the appointment of judges than participants in the 'no-party' condition ($M_{diff} = .59$, $p < .001$) and participants in the 'in-party' condition ($M_{diff} = .69$, $p < .001$). Similarly, participants in the 'out-party' condition attributed significantly more malicious intentions to the reform of protest registration than participants in the 'no-party' condition ($M_{diff} = .44$, $= .002$) and participants in the 'in-party' condition ($M_{diff} = .43$, $p = .002$) and 'no-party' condition (this difference is not statistically significant). For these reforms, the differences between participants in the 'in-party and 'no-party' conditions were not statistically significant (see Table 25A in AO10). Only in the case of the reform of the presidential term the difference between participants in the 'in-party' and those in the neutral condition meets the standard significance level with $M_{diff} = -.29$, $p = .035$ (i.e., supporters of the party assigned significantly less malicious intentions to the reformers than those who did not know which party is proposing the reform). These results are largely in line with the hypothesis (H1b) that a cue of an out-group party increases the perception of malicious intentions behind a reform altering the system of checks and balances (by

26

comparison with 'no party' condition and 'in-party' cue).

The analysis shows that here too there was a significant effect of the reform type on the level of malicious motivations assigned to the reformers, $F(26.65, 2747.23)= 355.651$, $p < .001$, $\eta p^2 = .26$. Participants on average assigned less malicious intentions to the reformers when evaluating the reform of the tax statement deadline than when evaluating checks and balances reforms (see Table 23A in AO10). Participants were rather skeptical of such reforms and on average leaned towards evaluating the motivations of the reformers as aimed to undermine the functioning of democracy and skewing the balance of power: the level of assigned malicious intentions is above the mid-point (4) of the 7-point scale. Again, mirroring the results of the benign intentions analysis, participants attributed most malicious intentions to the protest registration reform ($M = 5.06$, $SE = .06$). Participants were least suspicious of the reform of the tax statement deadline and assigned much lower level of malicious intentions to the reformers proposing it ($M = 3.18$, $SE = .06$).

The interaction between the party cue and reform was significant ($F (5.30, 2747.23) = 2.41$, $p = .03$). The data distributions show that the difference between the in-party and out-party party conditions was the smallest for the evaluation of intentions behind the tax statement deadline reform ($M_{diff} = -.27$, $p = .049$) by comparison with the largest difference for evaluation of the appointment of judges reform ($M_{diff} = -.69$, $p < .001$). When evaluating the tax statement deadline reform, there were no significant differences between the participants in the neutral condition and participants in the 'out-party' condition, and no differences between the participants in the neutral condition and participants in the 'in-party' condition (see Table 25A in AO10).

Figure 3. Violin plots show the attribution of malicious intentions across conditions of party-cue (in-party, no party, and out-party). The white dot shows the median, the dark bar shows the

interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density. Higher values indicate attribution of more malicious intentions.
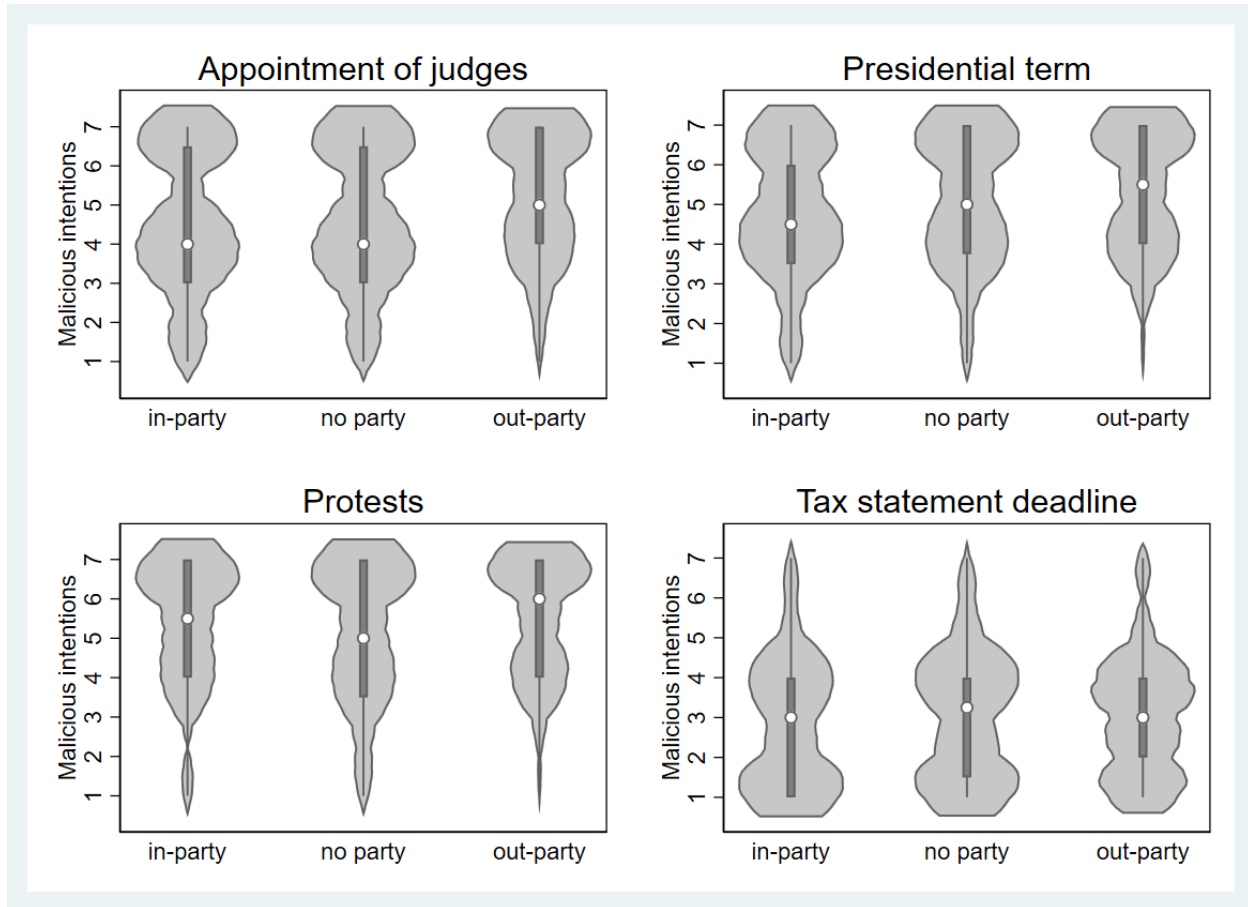
Overall, looking at the results of the perceived intentions analysis, the evidence for the hypothesis H3 (the effect of party-cue on assigned intentions is stronger for the reforms with higher stakes than reforms with lower stakes) is inconclusive. The results show that the scores of participants are concentrated much more at the extremes of the scales for attribution of benign (lowest scores) and malicious (highest scores) when they evaluate checks and balances reforms than when they evaluate the less consequential reform of the tax statement deadline. Furthermore, in all cases the difference between in-party and out-party cue in assigning intentions was larger when participants evaluated checks and balances reforms than the reform of tax statement deadline. However, this interaction

was significant for the malicious intentions, but not for the benign intentions. Importantly, these effects are small, while the sample size did not allow for detection of such small effects.

Robustness check

To test for potential differential treatment effects across different party voters, I conducted factorial ANOVAs for each checks and balances reform (appointment of judges, protest registration, and presidential term limit) and for all three dependent variables (support, benign intentions, and malicious intentions). The analyses showed that the party vote (Law & Justice, Civic Coalition, Other, No vote) was significantly associated with each of these dependent variables meaning that the voters of different parties differ in their level of support for reforming checks and balances. This is in line with the descriptive patterns of importance assigned to checks and balances presented in Table 1. In 'no party condition', the voters of the Civic Coalition expressed the lowest support for each of the reforms, while the voters of the Law and Justice expressed the highest support. This finding can be interpreted in two ways: the voters of PiS are the most supportive of such reforms or the 'no party condition' was seen as reflective of the status quo in Poland at the time, namely the government of PiS. Looking at the averages across all three conditions, the voters of PiS are the most supportive of all reforms and assigned least malicious intentions to the reformers (see AO8). This result too might point to the fact that the no-party condition has been interpreted in line with the *status quo* (i.e. that the PiS government was responsible for the reforms) and/or reflect that the PiS voters care less about protecting democratic checks and balances than the voters of Civic Platform. Importantly, in none of the models was the interaction between the condition (no party, in-party, out-party) and party vote significant, showing that the treatment affected all party voters similarly (see AO13).

Additional analyses with the restricted sample (N = 601) of participants who responded

correctly to all checks questions and were not prompted to read the vignettes again show robustness of the experimental results to the sample specifications (see AO15).

Discussion and conclusion

What do citizens make of proposals to alter democratic rules? Are they vigilant of changing the balance of power and weakening accountability mechanisms? The results of this study offer three main contributions.

First of all, party cues do affect how citizens evaluate the proposed reforms and how they evaluate the *intentions* of the reformers. The experiment showed clear evidence of 'my party bias': those participants who receive a cue that a government they support proposes a reform attributed more benign intentions to the reformers than others. By contrast, citizens who received a cue that the reform is proposed by a government they oppose were much more vigilant of such proposals and assigned more malicious authoritarian intentions to the government. This is in line with evidence that the supporters of incumbents might be more willing to trade off democracy for preferred policy and economic benefits (Graham & Svolik, 2020; Singer, 2018). When evaluating the same reform, citizens rely on pre-existing beliefs about the reformer (whom to trust). This evidence aligns with both theories, epistemic vigilance (being weary of who proposes the reform) and motivated reasoning (following the party cues) and contributes to the explanation why in societies with high level of pronounced support for liberal democratic principles, citizens' commitment to democracy seems volatile and contextual.

Are checks and balances reforms special? First, checks and balances reforms were supported significantly less than the administrative reform. Also the intentions of the reformers were seen as more malicious and less benign when they proposed reforms undermining democratic rules. This

aligns with the idea of epistemic vigilance: citizens are more weary of reforms with high stakes. However, the party-cues influenced not only reforms of checks and balances, but also the low-stakes administrative reform. So although a rather inconsequential change of a tax statement deadline could be considered a 'consensus issue' (Slothuus & De Vreese, 2010), citizens' evaluations were still affected by who proposes it: when the reform is proposed by 'my party', the intentions behind the reform are perceived as more benign and more support for such reform is expressed. The results show that motivated reasoning was active no matter the stakes of the reforms. However, although the evidence for the interaction between party cues and assigned intentions was inconclusive, the difference between in-party and out-party supporters in their approval of the reforms was much smaller in the case of the tax statement deadline than all the checks and balances. Moreover, irrespective of the party issuing the proposal, the administrative reform was not seen as ill-intended, by contrast with the checks and balances reforms. Both these patterns show that epistemic vigilance in a way dampens the effect of motivated reasoning in the case of less relevant reforms and co-triggers partisanship in the case of the reforms changing the balance of power.

Theoretically, this means that our shared psychology equips us with mechanisms to avoid deception and harm and, as a consequence, be skeptical of the political proposals by an out-group. This leads to a difference in the evaluations of the intentions of the politicians whom we believe have our interest in mind (in-party) vs. those who do not (out-party). It also leads to more vigilance of actions with high stakes. While the theories of epistemic vigilance and motivated reasoning strive to explain the same outcomes, they often do not speak to each other: the former proposes that in-group favoritism is a consequence of (reasonable) vigilance, while the latter frames it more in terms of bias. This study showed that when it comes to politics, we likely observe an interplay between them. The next step in integrating epistemic vigilance with motivated reasoning is to theorize a boundary

31

between a healthy dose of skepticism of the out-party (epistemic vigilance) and the blind-following of the in-party line (motivated reasoning). On the flip side, we would benefit from learning what are the limits of motivated reasoning and whether epistemic vigilance can help us find them.

Future research can address several limitations of this study. Although the experiment uses three different reform scenarios, the results do not necessarily generalize to different checks and balances and different violations that could lead to backsliding. In addition, the results show some variability in the evaluations of different types of reforms and somewhat different levels of intentions assigned to them. For example, participants consistently evaluated the proposed reform of protest registration in most negative terms. It could be that the differences in evaluations of the reforms of checks and balances reflect familiarity with particular reforms (e.g., Polish participants were more likely to be familiar with the debates about the courts reforms). Another potential explanation is that the protest registration change represents a reform with the highest stakes in the eyes of citizens (especially the opposition voters). Future studies could expand the scope of tested reforms to other checks and balances, measure the importance of different types of reforms and their evaluations directly, and further refine the experimental instrument, including the measures of perceived intentions of the reformers. Alternative designs are needed to test the epistemic vigilance activation in the context of politics more directly. In addition, future studies could theorize and test different triggers – beyond party cues and stakes/relevance – of epistemic vigilance in the context of politics.

Furthermore, the presented study applies an experimental design to be able to detect the causal link between party cues and intentions assigned to the reformers. Despite the strengths of such design, experimental effects do not necessarily directly translate to the effects on attitudes and behavior outside of the experimental setting. In addition, the effects found in the experiment cannot be translated directly to the population, because the used sample is not fully representative. At the same

time, if the evaluations of the reforms and the reformers were affected by these rather weak party cues, it is not unreasonable to expect stronger effects among citizens participating in real-life political struggles. Importantly, because of insufficient power, the test of the interaction between party cues and reform type is inconclusive, and the robustness of the found effects would need to be tested with a larger sample.

Finally, this study used data from Poland to test general theories about party-cues that garnered largest support from research conducted in polarized systems such as the US (Gidengil et al., 2022). It expanded evidence that in the systems with multiple parties the party cues can affect citizens' attitudes and their evaluations of institutional reforms. Further studies could test if these effects are present in multiparty systems with different levels of democracy and lower polarization.

The results imply that democracy can become vulnerable when the ruling majority is not too worried about the dismantling of institutions that hold the political elites accountable, i.e., when the majority turns into passive bystanders (Saikkonen & Christensen, 2023). In the analyzed case of Poland at the time of data collection, the incumbent government represented the voters who were on average more supportive of altering democratic accountability mechanisms (the voters of Law and Justice), which might have contributed even further to the democratic backsliding (see Wunsch et al., 2025). This study tells us that if a government that has a parliamentary majority starts shifting the political playing field and abusing their power, the majority of citizens who are on the side of this government will be less vigilant of potentially long-term negative effects of such changes. By contrast, the political opponents will be the most vigilant and suspicious of the intentions behind altering checks and balances and therefore more likely to try to prevent the erosion of democratic principles. Importantly, although the supporters of different parties have different levels of approval for reforms undermining checks and balances, the results show that, overall, the majority of citizens

are vigilant of changing the balance of power and are not enthusiastic about reforms changing liberal democratic rules.

References

Aarslew, L. F. (2023). Why Don't Partisans Sanction Electoral Malpractice? *British Journal of Political Science*, *53*(2), 407–423. https://doi.org/10.1017/S0007123422000126

Anderson, C. J., Blais, A., Bowler, S., Donovan, T., & Listhaug, O. (2005). *Losers' Consent: Elections and Democratic Legitimacy.* https://academic.oup.com/book/32905?itm_content=Oxford_Academic_Books_0&itm_campaign =Oxford_Academic_Books&itm_source=trendmd-widget&itm_medium=sidebar

Anderson, C. J., Bol, D., & Ananda, A. (2021). Humanity's attitudes about democracy and political leaders: Patterns and trends. *Public Opinion Quarterly*, *85*(4), 957–986.

Bayes, R., Druckman, J. N., Goods, A., & Molden, D. C. (2020). When and How Different Motives Can Drive Motivated Political Reasoning. *Political Psychology*, *41*(5), 1031–1052. https://doi.org/10.1111/pops.12663

Bermeo, N. (2016). On Democratic Backsliding. *Journal of Democracy*, *27*(1), 5–19.

Bettarelli, L., Reiljan, A., & Van Haute, E. (2023). A regional perspective to the study of affective polarization. *European Journal of Political Research*, *62*(2), 645–659. https://doi.org/10.1111/1475-6765.12548

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press.

Boese, V. A., Lundstedt, M., Morrison, K., Sato, Y., & Lindberg, S. I. (2022). State of the world 2021: Autocratization changing its nature? *Democratization*, *29*(6), 983–1013. https://doi.org/10.1080/13510347.2022.2069751

Braley, A., Lenz, G. S., Adjodah, D., Rahnama, H., & Pentland, A. (2023). Why voters who value democracy participate in democratic backsliding. *Nature Human Behaviour*, *7*(8), 1282–1293. https://doi.org/10.1038/s41562-023-01594-w

Bryan, J. D. (2023). What Kind of Democracy Do We All Support? How Partisan Interest Impacts a Citizen's Conceptualization of Democracy. *Comparative Political Studies*, *56*(10), 1597–1627. https://doi.org/10.1177/00104140231152784

De Blok, L., Van Der Meer, T., & Van Der Brug, W. (2022). Policy area satisfaction, perceptions of responsibility, and political trust: A novel application of the REWB model to testing evaluation-

based political trust. *Journal of Elections, Public Opinion and Parties*, *32*(1), 129–150. https://doi.org/10.1080/17457289.2020.1780433

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. SAGE.

Frederiksen, K. V. S. (2024). Do Partisanship and Policy Agreement Make Citizens Tolerate Undemocratic Behavior? *The Journal of Politics*. https://doi.org/10.1086/726938

Gidengil, E., Stolle, D., & Bergeron-Boutin, O. (2022). The partisan nature of support for democratic backsliding: A comparative perspective. *European Journal of Political Research*, *61*(4), 901–929. https://doi.org/10.1111/1475-6765.12502

Graham, M. H., & Svolik, M. W. (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. https://doi.org/10.1017/S0003055420000052

Hansen, S. W., Klemmensen, R., & Serritzlew, S. (2019). Losers lose more than winners win: Asymmetrical effects of winning and losing in elections. *European Journal of Political Research*, *58*(4), 1172–1190. https://doi.org/10.1111/1475-6765.12329

Hernández, E. (2019). Democracy belief systems in Europe: Cognitive availability and attitudinal constraint. *European Political Science Review*, *11*(4), 485–502. https://doi.org/10.1017/S1755773919000286

Jacob, M. S. (2024). Citizens as a democratic safeguard? The sequence of sanctioning elite attacks on democracy. *American Journal of Political Science*, ajps.12847. https://doi.org/10.1111/ajps.12847

Jennings, W., Stoker, G., Valgarðsson, V., Devine, D., & Gaskell, J. (2021). How trust, mistrust and distrust shape the governance of the COVID-19 crisis. *Journal of European Public Policy*, *28*(8), 1174–1196. https://doi.org/10.1080/13501763.2021.1942151

Kaftan, L. (2024). Lip service to liberal democracy in Western Europe? *European Political Science Review*, 1–18. https://doi.org/10.1017/S1755773924000079

Karv, T. (2022). Does the democratic performance really matter for regime support? Evidence from the post-communist Member States of the European Union. *East European Politics*, *38*(1), 61–82. https://doi.org/10.1080/21599165.2021.1878150

Krishnarajan, S. (2023). Rationalizing Democracy: The Perceptual Bias and (Un)Democratic Behavior. *American Political Science Review*, *117*(2), 474–496. https://doi.org/10.1017/S0003055422000806

Leeper, T. J., & Slothuus, R. (2014). Political Parties, Motivated Reasoning, and Public Opinion Formation. *Political Psychology*, *35*(S1), 129–156. https://doi.org/10.1111/pops.12164

Levi, M., & Stoker, L. (2000). Political Trust and Trustworthiness. *Annual Review of Political Science*, *3*(1), 475–507. https://doi.org/10.1146/annurev.polisci.3.1.475

Littvay, L., McCoy, J. L., & Simonovits, G. (2024). It's Not Just Trump: Americans of Both Parties Support Liberal Democratic Norm Violations More Under Their Own President. *Public Opinion Quarterly*, nfae042. https://doi.org/10.1093/poq/nfae042

Lührmann, A., & Lindberg, S. I. (2019). A third wave of autocratization is here: What is new about it? *Democratization*, *26*(7), 1095–1113. https://doi.org/10.1080/13510347.2019.1582029

Lührmann, A., Marquardt, K. L., & Mechkova, V. (2020). Constraining governments: New indices of vertical, horizontal, and diagonal accountability. *American Political Science Review*, *114*(3), 811–820.

Matovski, A. (2021). *Popular dictatorships: Crises, mass opinion, and the rise of electoral authoritarianism*. Cambridge University Press. https://books.google.com/books?hl=en&lr=&id=P2lHEAAAQBAJ&oi=fnd&pg=PR8&dq=matovsky+popular+dictatorships&ots=URurB1ZofT&sig=I3hPAIsqg8J_b0JNNOP4s1bM9FY

Mazepus, H., & Toshkov, D. (2022). Standing up for Democracy? Explaining Citizens' Support for Democratic Checks and Balances. *Comparative Political Studies*, *55*(8), 1271–1297. https://doi.org/10.1177/00104140211060285

Mercier, H. (2017). How Gullible are We? A Review of the Evidence from Psychology and Social Science. *Review of General Psychology*, *21*(2), 103–122. https://doi.org/10.1037/gpr0000111

Mercier, H. (2020). *Not Born Yesterday: The Science of Who We Trust and What We Believe*. Princeton University Press. https://doi.org/10.1515/9780691198842

Mutz, D. C., & Pemantle, R. (2015). Standards for Experimental Research: Encouraging a Better Understanding of Experimental Methods. *Journal of Experimental Political Science*, *2*(2), 192–215. https://doi.org/10.1017/XPS.2015.4

Norris, P. (2022). *In praise of skepticism: Trust but verify*. Oxford University Press. https://books.google.com/books?hl=en&lr=&id=J-t5EAAAQBAJ&oi=fnd&pg=PP1&dq=in+praise+of+skepticism&ots=0UpRKBSo4M&sig=RDtPDSNDxcCbz-XcxSGP2r5kKM4

Petersen, M. B., Skov, M., Serritzlew, S., & Ramsøy, T. (2013). Motivated Reasoning and Political Parties: Evidence for Increased Processing in the Face of Party Cues. *Political Behavior*, *35*(4), 831–854. https://doi.org/10.1007/s11109-012-9213-1

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123–205). Academic Press. https://doi.org/10.1016/S0065-2601(08)60214-2

Pietraszewski, D., Curry, O. S., Petersen, M. B., Cosmides, L., & Tooby, J. (2015). Constituents of political cognition: Race, party politics, and the alliance detection system. *Cognition*, *140*, 24–39.

Saikkonen, I. A.-L., & Christensen, H. S. (2023). Guardians of Democracy or Passive Bystanders? A Conjoint Experiment on Elite Transgressions of Democratic Norms. *Political Research Quarterly*, *76*(1), 127–142. https://doi.org/10.1177/10659129211073592

Şaşmaz, A., Yagci, A. H., & Ziblatt, D. (2022). How Voters Respond to Presidential Assaults on Checks and Balances: Evidence from a Survey Experiment in Turkey. *Comparative Political Studies*, *55*(11), 1947–1980. https://doi.org/10.1177/00104140211066216

Simonovits, G., McCoy, J., & Littvay, L. (2022). Democratic Hypocrisy and Out-Group Threat: Explaining Citizen Support for Democratic Erosion. *The Journal of Politics*, *84*(3), 1806–1811. https://doi.org/10.1086/719009

Singer, M. (2018). Delegating Away Democracy: How Good Representation and Policy Successes Can Undermine Democratic Legitimacy. *Comparative Political Studies*, *51*(13), 1754–1788. https://doi.org/10.1177/0010414018784054

Sloman, S. A., & Rabb, N. (2019). Thought as a determinant of political opinion. *Cognition*, *188*, 1–7. https://doi.org/10.1016/j.cognition.2019.02.014

Slothuus, R., & De Vreese, C. H. (2010). Political Parties, Motivated Reasoning, and Issue Framing Effects. *The Journal of Politics*, *72*(3), 630–645. https://doi.org/10.1017/S002238161000006X

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Taber, C. S., & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Voeten, E. (2016). Are people really turning away from democracy? *SSRN 2882878*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2882878

Wunsch, N., Jacob, M. S., & Derksen, L. (2025). The demand side of democratic backsliding: How divergent understandings of democracy shape political choice. *British Journal of Political Science*, *55*, e39.

Appendix Online

Are citizens vigilant of changing the balance of power? Testing evaluations of checks and balances reforms

Contents

AO1. Ethics approval

This study was approved by the [blinded for review] on 8 September 2021 under the reference number [blinded for review].

The researchers obtained informed consent from participants using the following form translated to Polish:

Welcome to this study conducted by [blinded for review] and [blinded for review]. Please read the information below before you continue.

Aim of the study: The study aims to understand political behavior and preferences of citizens. We are interested in your true views and preferences.

Duration: The survey takes around 25 minutes.

Compensation: Participants will be compensated in line with the rate agreed on with the survey company Respondi.

Potential risks: There are no expected risks or side-effects of participation.

Anonymity: Your information remains confidential. We do not ask for your name or any other identifying information. Your answers cannot be traced back to you.

Data and publication: The collected data is stored safely. In publications, there will be no possibility to link individual answers with identifying information about respondents. Following the standards of open science, the anonymized data will be submitted to data repositories for research purposes.

Consent to participate

I am at least 18 years of age and participate voluntarily in the study. I was informed about the nature of the study and any possible side effects.

1: I want to participate in this survey

2: I do not want to participate in this survey

AO2. Survey flow

```
                    ┌─────────────────────────────────────────┐
                    │   Survey introduction/Informed consent   │
                    └─────────────────────────────────────────┘
                                      │
                                      ▼
              ┌──────────────────────────────────────────────────┐
              │ Pre-experimental questions: demographics, sound check │
              └──────────────────────────────────────────────────┘
                                      │
                                      ▼
              ┌──────────────────────────────────────────────────┐
              │   Random assignment to an experimental condition   │
              └──────────────────────────────────────────────────┘
                  │                    │                    │
                  ▼                    ▼                    ▼
          ┌──────────────┐     ┌──────────────┐     ┌──────────────┐
          │  No-party    │     │  In-party    │     │  Out-party   │
          │  (n = 352)   │     │  (n = 345)   │     │  (n = 343)   │
          └──────────────┘     └──────────────┘     └──────────────┘
                                      │
                                      ▼
              ┌──────────────────────────────────────────────────┐
              │ Presentation of four reform vignettes (random order) │
              └──────────────────────────────────────────────────┘
```

Appointment of judges | Term limit | Protest | Tax statement deadline

Manipulation & comprehension checks

If failed: redirection to re-read and answer again

DV measures: reform support, benign intentions, malicious intentions

Unrelated study

Democratic principles questions

End of survey

AO3. Vignettes

Reform of appointment of judges:

Imagine that **one of the political parties** in your country/**a political party you support and feel very close to/ a political party you oppose and do not feel close to at all** wins the majority at a fully competitive and fair election. This means that the party gets to form the government and implement various policies.

One of the ministers proposes to implement **a reform that makes changes to the way judges are appointed and dismissed.** As a result, some of the current judges will be dismissed from their posts. The minister claims that this reform is needed to get rid of the corrupt judges that were working for the previous government. The government supports this plan.

Reform of presidential term-limit:

Imagine that **one of the political parties** in your country/**a political party you support and feel very close to/ a political party you oppose and do not feel close to at all** wins the majority at a fully competitive and fair election. This means that the party gets to form the government and implement various policies.

One of the ministers proposes **to get rid of the two-term limit for presidency**. As a result, the same person will be allowed to be the president for more than two 5-year terms. The minister claims that this reform is needed to assure the continuity of the state and democracy. The government supports this plan.

Reform of protest registration:

Imagine that **one of the political parties** in your country/**a political party you support and feel very close to/ a political party you oppose and do not feel close to at all** wins the majority at a fully competitive and fair election. This means that the party gets to form the government and implement various policies.

One of the ministers proposes to introduce **new law for registration of protests.** As a result, it will be more difficult to organize a protest. The minister informs that this reform is needed to assure safety of citizens' gatherings. The government supports this plan.

For the tax reform:

Imagine that **one of the political parties** in your country/**a political party you support and feel very close to/ a political party you oppose and do not feel close to at all** wins the majority at a fully competitive and fair election. This means that the party gets to form the government and implement various policies.

4

One of the ministers proposes to introduce **a new deadline for submitting tax statements.** This means that the deadline for submitting tax statements will fall on the 1st of June rather than 1st of May. The minister claims that this reform is needed to improve the efficiency of the tax reporting system. The government supports this plan.

AO4. Measures of dependent variables

The wording of dependent variable questions measuring perceived government motivations is different for the vignettes about each of the checks and balances reforms and for the administrative reform. The questions, however, aim to measure the same underlying idea. The questions measuring the evaluation of each reform are the same across the different reform vignettes.

Reform of appointment of judges:

In your opinion, what is the most likely motivation of this government behind introducing such a reform?
Please indicate your answer on the scale from 1- 'Very unlikely' to 7- 'Very likely'.

The government proposes this reform to:
Benign motivations: assure independence of judges, so the judges can check the government
Benign motivations: strengthen democracy
$r = .74$, $p < .001$, Cronbach's alpha $= 0.85$

Malicious motivations: have fewer limitations on their power, so the government can do whatever it wants
Malicious motivations: undermine democracy
$r = .66$, $p < .001$, Cronbach's alpha $= 0.80$

Reform of presidential term-limit:
In your opinion, what is the most likely motivation of this government behind introducing such a reform?
Please indicate your answer on the scale from 1- 'Very unlikely' to 7- 'Very likely'.

The government proposes this reform to:
Benign motivations: assure that the president can check the government in line with the mechanisms of checks and balances of power
Benign motivations: strengthen democracy
$r = .64$, $p < .001$, Cronbach's alpha $= 0.63$

Malicious motivations: have fewer limitations on their power, so the government can do whatever it wants

Malicious motivations: undermine democracy
$r = .61$, $p < .001$, Cronbach's alpha $= 0.75$

Reform of protest registration:

In your opinion, what is the most likely motivation of this government behind introducing such a reform?
Please indicate your answer on the scale from 1- 'Very unlikely' to 7- 'Very likely'.

The government proposes this reform to:
Benign motivations: assure that citizens can protest in a safe way
Benign motivations: strengthen democracy
$r = .72$, $p < .001$, Cronbach's alpha $= 0.84$

Malicious motivations: have fewer limitations on their power, so the government can do whatever it wants
Malicious motivations: undermine democracy
$r = .63$ $p < 001$, Cronbach's alpha $= 0.77$

For the tax reform:

In your opinion, what is the most likely motivation of this government behind introducing such a reform?
Please indicate your answer on the scale from 1- 'Very unlikely' to 7- 'Very likely'.

The government proposes this reform to:
Benign motivations: make it easier for citizens to file taxes
Benign motivations: improve the functioning of the state administration
$r = .64$, $p < .001$, Cronbach's alpha $= 0.78$

Malicious motivations: make it more difficult for citizens to file their taxes
Malicious motivations: worsen the functioning of the state administration
$r = .71$, $p < .001$, Cronbach's alpha $= 0.83$

For all vignettes:

Reform support: To what extent would you find such a reform justified? Please indicate your answer on the scale from 1- 'Not at all' to 7- 'fully'
Reform support: To what extent would you agree with such a reform? Please indicate your answer on the scale from 1- 'Very negatively' to 7- 'Very positively'

Courts: $r = .93$, $p < .001$, Cronbach's alpha $= 0.96$
Presidential term: $r = .91$, $p < .001$, Cronbach's alpha $= 0.95$

Protests: r = .91, p < .001, Cronbach's alpha = 0.95
Tax statement deadline: r = .87, p < .001, Cronbach's alpha = 0.93

AO5. Sample description and ecological validity checks

Variable	Population	Sample
Gender		
Women	52.34	52
Men	47.66	48
Age		
18-29	15.68	19.71
30-39	19.15	21.71
40-49	18.98	20.87
50-59	14.71	17.02
60-69	16.33	20.67
> 69	15.14	0
*Higher education**	29.10	40.96
Vote in 2019		
Polish Coalition	8.55	6
Law and Justice	43.59	25
The Left	12.56	12
Confederation	6.81	8
Civic Platform	27.40	26
Did not vote	38.26	13

The sample included participants supporting different political parties: 25% of participants indicated that they casted their vote in the last parliamentary elections for Law and Justice, 26% for the Civic Coalition (with Civic Platform as one of the parties in this block), 12% for the Left (with Democratic Left Alliance, Spring, and Together), 8% for Confederation, 6% for Polish Coalition (with Polish People's Party and Kukiz 15), and 13% indicated that they did not vote. This shows that supporters of different main political parties were included in the sample and the sample was not politically skewed towards one side of the political spectrum. In terms of education, 5% of participants indicated an education level below upper secondary, 54% upper secondary, and 41% tertiary (bachelor or master) education.

To check for the ecological validity of the presented scenarios, participants answered the question 'Was it difficult to imagine that the reforms described in the scenarios are being considered in Poland?' for each of the reforms (getting rid of the two-term limit for presidency, changing the appointment and dismissal of judges, changing the registration of protests, changing the tax statement deadline) on the scale from 1- "Very difficult" to 7- "Very easy". The results show that on average it was not too difficult for participants to imagine the reforms presented in the scenarios. The mean answer to this question for each reform was above the mid-point of the scale. The reform changing the tax statement deadline was the easiest to imagine as being considered in Poland (M = 4.89, SD = 1.79), followed by the registration of protests (M = 4.68, SD = 1.20), the appointment of judges (M = 4.61, SD = 2.04), and the presidential term reform (M = 4.36, SD = 2.07).

Given the situation in Poland at the time of the study, it could be that it was easier to imagine that a particular party proposes reforms of checks and balances. To check if this was the case, participants were asked 'Is it difficult to imagine that your **favorite** party in Poland considers a reform changing the rules of the appointment and dismissal of judges?' and answered on the scale from 1- "Very difficult" to 7- "Very easy". Table 6A shows the mean responses to this question per party-vote for participants who were in the in-party condition.

Table 1A. Mean answers to the question 'Is it difficult to imagine that your favorite party in Poland considers a reform changing the rules of the appointment and dismissal of judges?' on the scale from 1- "very difficult" to 7- "very easy". The results are for participants in the in-party condition across parties they voted for.

Party vote	M	SD	N
Koalicja Polska (z PSL i Kukizem 15)	3.42	2.06	24
Prawo i Sprawiedliwość (PiS)	4.82	1.75	83
Lewica (z SLD, Wiosną i Razem)	3.14	1.82	44
Konfederacja (z KORWiN i Ruchem Narodowy	4.07	2.23	27
Koalicja Obywatelska (z PO, Nowoczesną, Inicjatywą Polską i Partią Zielonych)	3.36	2.07	85

AO6. Manipulation/comprehension checks

After each vignette and with the text of the vignette visible on the screen participants were asked to answer three questions (see below): two checking if the crucial manipulated information was picked up (which party introduces the reform and what type of reform is proposed) and one about information that was not manipulated to at least partially address the potential issue of second-guessing the design of the experiment (which political figure proposed the reform). If a participant answered one of these questions incorrectly, they were prompted with the text again and asked to read again and answer correctly to continue. All participants moved to the dependent variable measures after answering the manipulation check questions for the second time irrespective of their answers.

a) In the text above, we asked you to imagine a party winning the election. According to the scenario you read, which party won the election?
1- A party you support and feel very close to
2- A party you oppose and do not feel close to at all
3- The Green party
4- The Social Democratic Party
5- The information about which party won was not specified in the text

b) In the text above, we asked you to imagine a reform. According to the scenario, which reform was proposed?
1- Reform of the appointment and dismissal of judges
2- Reform of registration of protests
3- Reform of the presidential term limit
4- Reform of the tax statement deadline

5- Reform of the military

c) In the text above, we asked you to imagine that someone proposes a reform. According to the scenario you read, who proposed the reform?
1- The president
2- The Prime-minister
3- One of the members of parliament
4- One of the ministers
5- An opposition party leader

Table 2A. Frequencies for answers to the manipulation and comprehension check questions. The '1st try pass/fail' indicates the number of participants who got all three answers correct/who got one of the three answers incorrect at the first attempt. The '2nd try pass/fail' indicates the total number of participants who answered all three questions correctly/one of the three questions incorrectly for after the second try. 'Total pass' and 'total fail' indicate the final totals per each question.

Condition	1st try pass/fail	2nd try pass/fail	Reform	Question	total pass	total fail
no party						
N = 352						
	214/138	239/113	courts	party	249	103
	61%/39%	68%/32%		reform	316	36
				actor	307	45
	225/127	238/114	term limit	party	257	95
	64%/36%	68%/32%		reform	305	47
				actor	299	53
	198/154	234/118	protest	party	256	96
	56%/44%	66%/34%		reform	300	52
				actor	293	113
	227/125	250/102	tax deadline	party	275	77
	64%/36%	71%/29%		reform	306	46

			actor	305	47

in-party

N = 345

207/138	246/99	courts	party	269	76
60%/40%	71%/29%		reform	305	40
			actor	300	45
223/122	250/95	term limit	party	274	71
65%/35%	72%/28%		reform	305	40
			actor	295	50
218/127	246/99	protest	party	276	69
63%/37%	71%/29%		reform	289	62
			actor	304	41
215/130	242/103	tax deadline	party	264	81
62%/38%	70%/30%		reform	302	43
			actor	298	47

out-party

N = 343

211/132	242/101	courts	party	258	85
62%/38%	71%/29%		reform	297	46
			actor	294	49
216/127	238/105	term limit	party	255	88
63%/37%	69%/31%		reform	289	54
			actor	298	45
218/125	243/100	protest	party	266	77
64%/36%	71%/29%		reform	292	51
			actor	291	52

210/133	237/106	tax deadline	party	251	92
61%/39%	69%/31%		reform	296	47
			actor	294	49

AO7. Correlation tables of dependent variables

Table 3A. Evaluations of courts reform

	courts benign	courts malicious	courts reform support
courts benign	1	-.604***	.822***
courts malicious	-.604***	1	-.645***
courts reform support	.822***	-.645***	1

Note. *** p < .001, N = 1039

Table 4A. Evaluations of presidential term reform

	presidential term benign	presidential term malicious	presidential term support
presidential term benign	1	-.446***	.719***
presidential term malicious	-.446***	1	-.585***
presidential term support	.719***	-.585***	1

Note. *** p < .001, N = 1039

Table 5A. Evaluations of the reform of protest registration

	protests benign	protests malicious	protests support
protests benign	1	-.562***	.844***
protests malicious	-.562***	1	-.554***
protests support	.844***	-.554***	1

Note. *** p < .001, N = 1039

Table 6A. Evaluations of the reform of tax statement deadline

	tax deadline benign	tax deadline malicious	tax deadline support
tax deadline benign	1	-.382***	.749***
tax deadline malicious	-.382***	1	-.345***
tax deadline support	.749***	-.345***	1

Note. *** p < .001, N = 1039

AO8. Descriptive statistics for dependent variables per condition

Table 7A . Estimated means for reform support across party cues and reforms.

Party cue	Reform	Mean	SE	95% CI
no party	appointment of judges	3.57	.111	3.35, 3.79
(N = 352)	presidential term	2.78	.104	2.57, 2.98
	protests	2.73	.102	2.53, 2.93
	tax statement deadline	3.97	.096	3.78, 4.16
in-party	appointment of judges	3.89	.112	3.68, 4.11
(N = 345)	presidential term	3.36	.105	3.15, 3.56
	protests	3.15	.103	2.95, 3.35
	tax statement deadline	4.37	.097	4.18, 4.56
out-party	appointment of judges	2.97	.112	2.75, 3.19
(N = 343)	presidential term	2.69	.106	2.49, 2.90
	protests	2.60	.103	2.39, 2.80

tax statement deadline	3.93	.097	3.74, 4.12	

Table 8A. Estimated means for benign intentions across party cues and reforms.

Party cue	Reform	Mean	SE	95% CI
no party	appointment of judges	3.33	.106	3.12, 3.54
(N = 352)	presidential term	3.29	.092	3.11, 3.47
	protests	2.82	.097	2.63, 3.01
	tax statement deadline	4.22	.092	4.04, 4.40
in-party	appointment of judges	3.76	.107	3.55, 3.97
(N = 345)	presidential term	3.59	.093	3.40, 3.77
	protests	3.21	.098	3.02, 3.40
	tax statement deadline	4.54	.093	4.35, 4.72
out-party	appointment of judges	3.03	.108	2.82, 3.24
(N = 343)	presidential term	3.15	.094	2.97, 3.33
	protests	2.69	.098	2.50, 2.89
	tax statement deadline	4.22	.094	4.04, 4.41

Table 9A . Estimated means for malicious intentions across party cues and reforms.

Party cue	Reform	Mean	SE	95% CI
no party	appointment of judges	4.47	.103	4.27, 4.67
(N = 352)	presidential term	4.87	.096	4.68, 5.06
	protests	4.91	.099	4.71, 5.10

	tax statement deadline	3.18	.094	2.99, 3.36
in-party	appointment of judges	4.37	.104	4.16, 4.57
(N = 345)	presidential term	4.58	.097	4.39, 4.77
	protests	4.92	.100	4.73, 5.12
	tax statement deadline	3.04	.095	2.85, 3.23
out-party	appointment of judges	5.06	.105	4.85, 5.26
(N = 342)	presidential term	5.16	.098	4.97, 5.35
	protests	5.35	.100	5.16, 5.55
	tax statement deadline	3.31	.095	3.12, 3.49

Table 10A . Observed means and standard deviations per reform for support, benign intentions, and malicious intentions. N = 1040.

	Mean	SD
Support		
Appointment of judges	**3.48**	**2.11**
PiS (Law and Justice)	4.87	1.78
PO (Civic Platform)	2.39	1.89
Other	3.32	2.03
No vote	3.38	1.93
Protest	**2.82**	**1.93**
PiS (Law and Justice)	4.08	1.97
PO (Civic Platform)	1.93	1.53
Other	2.48	1.80
No vote	2.87	1.68
Presidential term limit	**2.94**	**1.97**
PiS (Law and Justice)	3.93	2.05
PO (Civic Platform)	2.20	1.78
Other	2.77	1.89
No vote	2.96	1.74
Tax statement deadline	**4.09**	**1.81**
PiS (Law and Justice)	4.83	1.73
PO (Civic Platform)	3.76	1.83
Other	3.99	1.76
No vote	3.77	1.73

Benign intentions		
Appointment of judges	**3.37**	**2.02**
PiS (Law and Justice)	4.61	1.74
PO (Civic Platform)	2.51	1.92
Other	3.09	1.91
No vote	3.33	1.88
Protest	**2.91**	**1.83**
PiS (Law and Justice)	4.12	1.77
PO (Civic Platform)	2.02	1.47
Other	2.56	1.79
No vote	2.97	1.70
Presidential term limit	**3.34**	**1.74**
PiS (Law and Justice)	4.18	1.75
PO (Civic Platform)	2.68	1.67
Other	3.14	1.60
No vote	3.42	1.57
Tax statement deadline	**4.33**	**1.74**
PiS (Law and Justice)	4.98	1.64
PO (Civic Platform)	4.07	1.82
Other	4.30	1.66
No vote	3.92	1.63
Malicious intentions		
Appointment of judges	**4.63**	**1.96**
PiS (Law and Justice)	3.66	1.83
PO (Civic Platform)	5.50	1.91
Other	4.84	1.81
No vote	4.46	1.83
Protest	**5.06**	**1.86**
PiS (Law and Justice)	4.15	1.86
PO (Civic Platform)	5.80	1.62
Other	5.36	1.72
No vote	4.85	1.83
Presidential term limit	**4.87**	**1.82**
PiS (Law and Justice)	4.05	1.86
PO (Civic Platform)	5.52	1.74
Other	5.09	1.69
No vote	4.76	1.68
Tax statement deadline	**3.17**	**1.177**
PiS (Law and Justice)	2.83	1.81
PO (Civic Platform)	3.29	1.83
Other	3.28	1.69
No vote	3.30	1.70

AO9. Plots of means and standard deviations for dependent variables

Figure 1A. Means and standard deviations for support of reforms across party-cue conditions (in-party, no party, out-party)

Appointment of judges

Protest

Presidential term limit
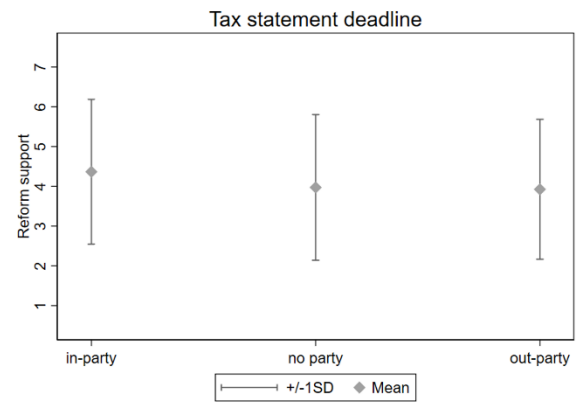

Tax statement deadline
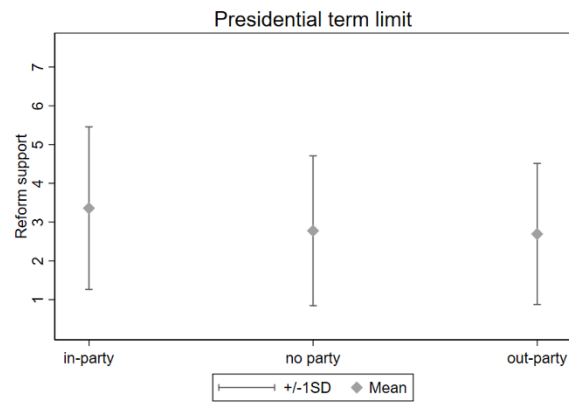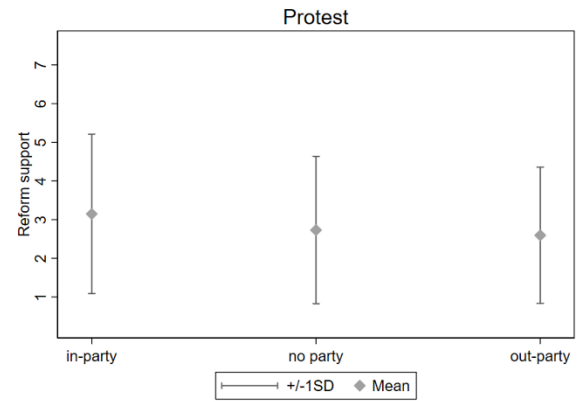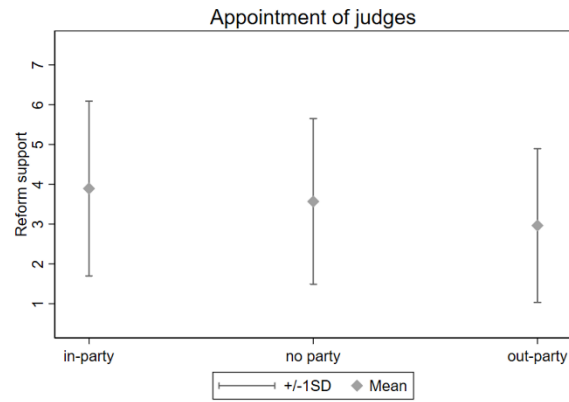
Figure 2A. Means and standard deviations for attribution of benign intentions to reforms across party-cue conditions (in-party, no party, out-party)
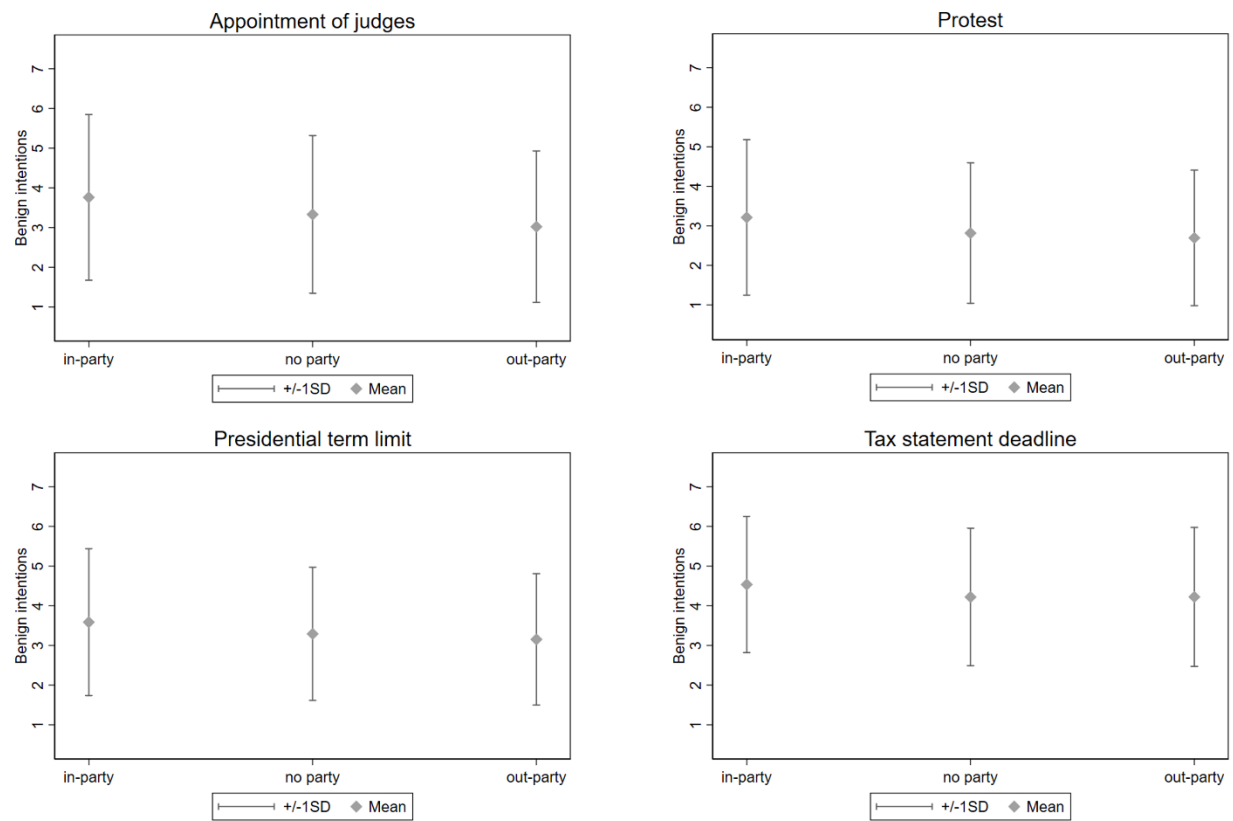
Figure 3A. Means and standard deviations for attribution of malicious intentions to reforms across party-cue conditions (in-party, no party, out-party)
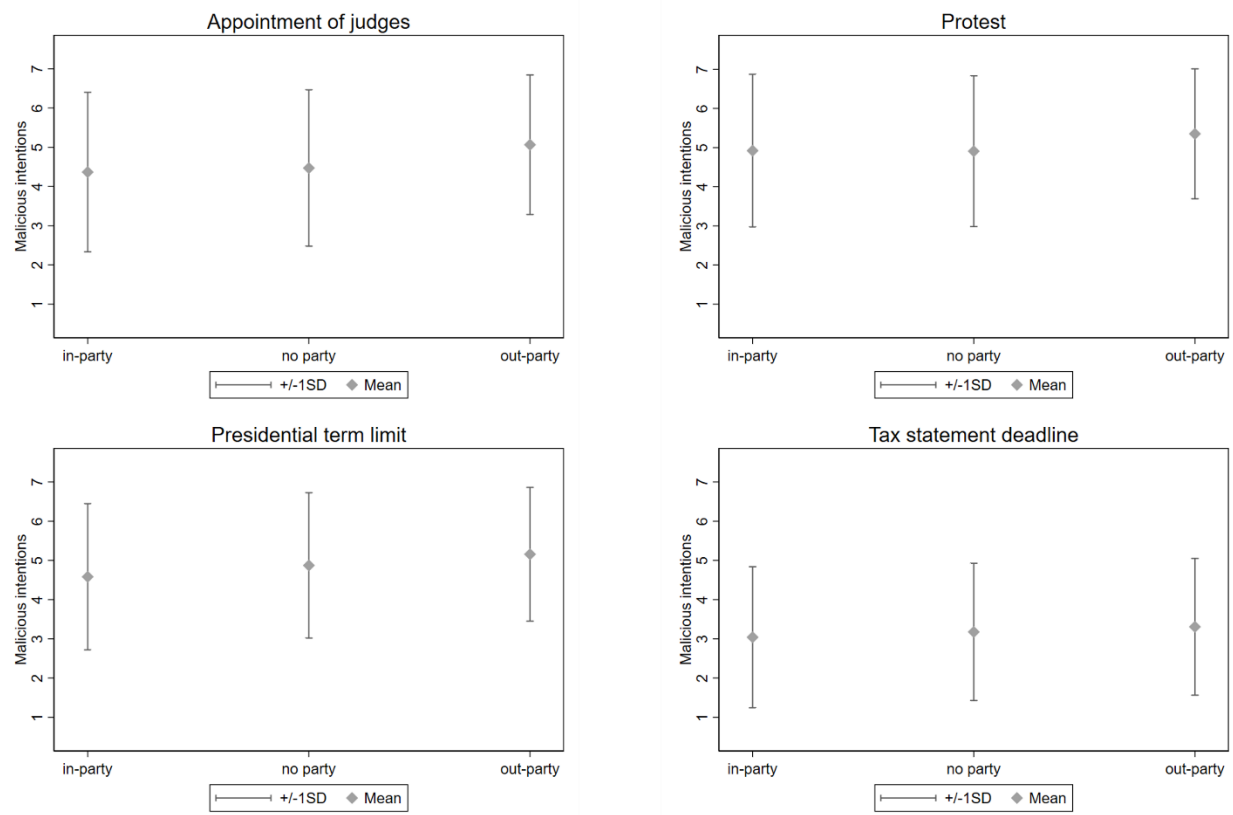
A10. Mixed-design ANOVAs results and pairwise comparisons

Reform support

Table 11A. Results of the mixed ANOVA for *reform support* (sphericity not assumed)

Factor	Test	df	F	p	Partial η^2
Reform (within effect)	Sphericity Assumed	3	166.79	<.001	0.139
	Greenhouse-Geisser	2.816	166.79	<.001	0.139
	Huynh-Feldt	2.830	166.79	<.001	0.139
	Lower-bound	1.000	166.79	<.001	0.139
Party cue (between effect)	Between-subject	2	17.03	<.001	0.032
Reform * party-cue (interaction effect)	Sphericity Assumed	6	3.420	.002	0.007
	Greenhouse-Geisser	5.632	3.420	.003	0.007
	Huynh-Feldt	5.659	3.420	.003	0.007

Lower-bound	2.000	3.420	.033	0.007
Error	1036			

Table 12A. Pairwise comparisons for *reform support* across party-cue conditions

Reference condition	Compared condition	M$_{diff}$	SE	p	95% CI
No party	In-party	-.430	.112	<.001	-0.650. -0.209
No party	Out-party	.218	.112	.053	-0.003. 0.438
In party	Out-party	.647	.113	<.001	0.426. 0.869

Table 13A. Pairwise comparisons for *reform support* across reforms

Reference condition	Compared condition	M$_{diff}$	SE	p	95% CI
Appointment of judges	Term limit	.534	.061	<.001	0.414. 0.654
	Protest	.652	.058	<.001	0.539. 0.765
	Tax statement deadline	-.612	.070	<.001	-0.749. -0.476
Term limit	Protest	.118	.054	<.030	0.011. 0.224
	Tax statement deadline	-1.146	.069	<.001	-1.282. -1.011
Protest	Tax statement deadline	-1.264	.067	<.001	-1.395. -1.134

Table 14A . Pairwise comparisons for the *reform support* across party-cue and reform conditions.

Party cue	Reform	M$_{diff}$	SE	p	95% CI
no party					
appointment of judges (ref.)	presidential term	.794	0.105	<.001	0.588, 1.000
	protests	.839	0.099	<.001	0.645, 1.034
	tax statement deadline	-.403	0.120	0.001	-0.638, -0.169
presidential term (ref.)	protest	.045	0.093	0.626	-0.138, 0.229
	tax statement deadline	-1.197	0.119	<.001	-1.431, -0.964
Protest (ref)	Tax statement deadline	-1.243	0.114	<.001	-1.467, -1.018

in-party					
appointment of judges (ref.)	presidential term	.536	0.106	<.001	0.328, 0.744
	protests	.746	0.100	<.001	0.550, 0.943
	tax statement deadline	-.472	0.121	<.001	-0.709, -0.236
presidential term (ref.)	protest	.210	0.094	0.026	0.025, 0.395
	tax statement deadline	-1.009	0.120	<.001	-1.244, -0.773
Protest (ref.)	Tax statement deadline	-1.219	0.116	<.001	-1.446, -0.992
out-party					
appointment of judges (ref.)	presidential term	.275	0.107	0.010	0.066, 0.484
	protests	.368	0.101	<.001	0.171, 0.566
	tax statement deadline	-.965	0.121	<.001	-1.203, -0.727
presidential term (ref.)	protest	.094	0.095	0.324	-0.092, 0.279
	tax statement deadline	-1.240	0.121	<.001	-1.476; -1.003
Protest (ref.)	Tax statement deadline	-1.333	0.116	<.001	-1.561, -1.106

Table 15A. Pairwise comparisons for the *reform suppor*t across reform and party-cue conditions.

Reform	Party-cue	M_{diff}	SE	p	95% CI
Appointment of judges					
No party (ref.)	In-party	-.325	0.157	0.039	-0.633, -0.016
	Out-party	.602	0.157	0.000	0.293, 0.911
In-party (ref)	Out-party	.926	0.158	0.000	0.616, 1.237
Presidential term					
No party (ref.)	In-party	-.582	0.148	0.000	-0.873, -0.292

	Out-party	0.083	0.148	0.578	-0.209, 0.374
In-party (ref)	Out-party	.665	0.149	0.000	0.372, 0.958

Protests

No party (ref.)	In-party	-.418	0.145	0.004	-0.702, -0.133
	Out-party	0.131	0.145	0.369	-0.154, 0.416
In-party (ref)	Out-party	.548	0.146	0.000	0.262, 0.835

Tax statement deadline

No party (ref.)	In-party	-.394	0.137	0.004	-0.662, -0.126
	Out-party	0.040	0.137	0.769	-0.228, 0.309
In-party (ref)	Out-party	.434	0.138	0.002	0.164, 0.704

Benign intentions

Table 16A. Results of the mixed ANOVA for *benign intentions* attribution (sphericity not assumed)

Factor	Test	df	F	p	Partial η^2
Reform	Sphericity Assumed	3	190.509	<.001	0.155
(within effect)	Greenhouse-Geisser	2.714	190.509	<.001	0.155
	Huynh-Feldt	2.727	190.509	<.001	0.155
	Lower-bound	1	190.509	<.001	0.155
Party cue	Between-subject	2	12.34	<.001	0.023
(between effect)					
Reform * party-cue	Sphericity Assumed	6	1.557	.156	0.003
(interaction effect)	Greenhouse-Geisser	5.427	1.557	.163	0.003
	Huynh-Feldt	5.453	1.557	.163	0.003
	Lower-bound	2.000	1.557	.211	0.003

Table 17A. Pairwise comparisons for *benign intentions attribution* across party-cue conditions

Reference condition	Compared condition	M_{diff}	SE	p	95% CI
No party	In-party	-.359	.103	.001	-0.561. -0.156
No party	Out-party	.142	.103	.171	-0.061. 0.345
In party	Out-party	.501	.104	<.001	0.297. 0.705

Table 18A. Pairwise comparisons for *benign intentions attribution* across reforms

Reference condition	Compared condition	M_{diff}	SE	p	95% CI
Appointment of judges	Term limit	.027	.058	.641	-0.087. 0.141
	Protest	.464	.054	<.001	0.358. 0.571
	Tax statement deadline	-0.955	.071	<.001	-1.093. -0.816
Term limit	Protest	.437	.051	<.001	0.337. 0.537
	Tax statement deadline	-.982	.065	<.001	-1.110. -0.854
Protest	Tax statement deadline	-1.419	.066	<.001	-1.549. -1.289

Table 19A. Pairwise comparisons for the *benign intentions attribution* across party-cue and reform conditions.

Party cue	Reform	M_{diff}	SE	p	95% CI
no party					
appointment of judges (ref.)	presidential term	0.038	0.100	0.700	-0.157, 0.234
	protests	.516	0.093	0.000	0.333, 0.698
	tax statement deadline	-.889	0.121	0.000	-1.127, -0.651
presidential term (ref.)	protest	.477	0.087	0.000	0.306, 0.649
	tax statement deadline	-.928	0.112	0.000	-1.147, -0.708
Protest (ref)	Tax statement deadline	-1.405	0.114	0.000	-1.629, -1.181
in-party					
appointment of judges (ref.)	presidential term	.174	0.101	0.084	-0.024, 0.371
	protests	.549	0.094	0.000	0.365, 0.0734
	tax statement deadline	-.774	0.123	0.000	-1.015, -0.533
presidential term (ref.)	protest	.375	0.088	0.000	0.202, 0.549
	tax statement deadline	-.948	0.113	0.000	-1.170, -0.726

Protest (ref.)	Tax statement deadline	-1.323	0.115	0.000	-1.550, -1.097

out-party

appointment of judges (ref.)	presidential term	-0.123	0.101	0.225	-0.321, 0.075
	protests	.329	0.095	0.001	0.144, 00.0514
	tax statement deadline	-1.205	0.123	0.000	-1.446, -0.963
presidential term (ref.)	protest	.452	0.089	0.000	0.278, 0.626
	tax statement deadline	-1.082	0.114	0.000	-1.305, -0.859
Protest (ref.)	Tax statement deadline	-1.534	0.116	0.000	-1.76, -1.306

Table 20A . Pairwise comparisons for the *benign intentions attribution* across reform and party-cue conditions.

Reform	Party-cue	M_{diff}	SE	p	95% CI
Appointment of judges					
No party (ref.)	In-party	-.430	0.151	0.005	-0.726, -0.133
	Out-party	.303	0.151	0.046	0.006, 0.600
In-party (ref)	Out-party	.733	0.152	0.000	0.434, 1.032
Presidential term					
No party (ref.)	In-party	-.294	0.131	0.025	-0.552, -0.037
	Out-party	0.142	0.131	0.280	-0.116, 0.400
In-party (ref)	Out-party	.436	0.132	0.001	0.177, 0.695
Protests					
No party (ref.)	In-party	-.396	0.138	0.004	-0.667, -0.125
	Out-party	0.117	0.138	0.400	-0.155, 0.388
In-party (ref)	Out-party	.513	0.139	0.000	0.240, 0.786

Tax statement deadline

No party (ref.)	In-party	-.315	0.131	0.017	-0.572, -0.057
	Out-party	-0.012	0.131	0.926	-0.270, 0.246
In-party (ref)	Out-party	.302	0.132	0.022	0.043, 0.561

Malicious intentions

Table 21A. Results of the mixed ANOVA for *malicious intentions* attribution (sphericity not assumed)

Factor	Test	df	F	p	Partial η^2
Reform	Sphericity Assumed	3	355.65	<.001	0.255
(within effect)	Greenhouse-Geisser	2.637	355.65	<.001	0.255
	Huynh-Feldt	2.649	355.65	<.001	0.255
	Lower-bound	1.000	355.65	<.001	0.255
Party cue	Between-subject	2	12.60	<.001	0.024
(between effect)					
Reform * party-cue	Sphericity Assumed	6	2.411	.025	0.005
(interaction effect)	Greenhouse-Geisser	5.274	2.411	.031	0.005
	Huynh-Feldt	5.298	2.411	.031	0.005
	Lower-bound	2.000	2.411	.090	0.005

Table 22A. Pairwise comparisons for *malicious intentions attribution* across party-cue conditions

Reference condition	Compared condition	M_{diff}	SE	p	95% CI
No party	In-party	-.359	.103	.001	-0.561. -0.156
No party	Out-party	.142	.103	.171	-0.061. 0.345
In party	Out-party	.501	.104	<.001	0.297. 0.705

Table 23A. Pairwise comparisons for *malicious intentions attribution* across reforms

Reference condition	Compared condition	M_{diff}	SE	p	95% CI
Appointment of judges	Term limit	.027	.058	.641	-0.087. 0.141
	Protest	.464	.054	<.001	0.358. 0.571

	Tax statement deadline	-0.955	.071	<.001	-1.093. -0.816
Term limit	Protest	.437	.051	<.001	0.337. 0.537
	Tax statement deadline	-.982	.065	<.001	-1.110. -0.854
Protest	Tax statement deadline	-1.419	.066	<.001	-1.549. -1.289

Table 24A. Pairwise comparisons for the *malicious intentions attribution* across party-cue and reform conditions.

Party cue	Reform	M_{diff}	SE	p	95% CI
no party					
appointment of judges (ref.)	presidential term	-.402	0.100	0.000	-0.599, -0.205
	protests	-.437	0.097	0.000	-0.628, -0.247
	tax statement deadline	1.294	0.122	0.000	1.055, 1.533
presidential term (ref.)	protest	-0.036	0.089	0.692	-0.211, 0.140
	tax statement deadline	1.696	0.122	0.000	1.456, 1.936
Protest (ref)	Tax statement deadline	1.732	0.127	0.000	1.483, 1.980
in-party					
appointment of judges (ref.)	presidential term	-.216	0.101	0.033	-0.415, -0.017
	protests	-.555	0.098	0.000	-0.748, -0.362
	tax statement deadline	1.326	0.123	0.000	1.085, 1.568
presidential term (ref.)	protest	-.339	0.090	0.000	-0.516, -0.162
	tax statement deadline	1.542	0.123	0.000	1.300, 1.784
Protest (ref.)	Tax statement deadline	1.881	0.128	0.000	1.630, 2.133
out-party					

appointment of judges (ref.)	presidential term	-0.101	0.102	0.322	-0.300, 0.099
	protests	-.288	0.099	0.004	-0.482, -0.094
	tax statement deadline	1.753	0.124	0.000	1.510, 1.996
presidential term (ref.)	protest	-.187	0.091	0.039	-0.365, -0.009
	tax statement deadline	1.854	0.124	0.000	1.610, 2.097
Protest (ref.)	Tax statement deadline	2.041	0.129	0.000	1.788, 2.293

Table 25A. Pairwise comparisons for the *malicious intentions attribution* across reform and party-cue conditions.

Reform	Party-cue	M_{diff}	SE	p	95% CI
Appointment of judges					
No party (ref.)	In-party	0.104	0.147	0.481	-0.185, 0.392
	Out-party	-.588	0.147	0.000	-0.877, -0.299
In-party (ref)	Out-party	-.692	0.148	0.000	-0.982, -0.401
Presidential term					
No party (ref.)	In-party	.290	0.137	0.035	0.020, 0.559
	Out-party	-.287	0.137	0.037	-0.557, -0.018
In-party (ref)	Out-party	-.577	0.138	0.000	-0.848, -0.306
Protests					
No party (ref.)	In-party	-0.014	0.140	0.920	-0.289, 0.261
	Out-party	-.439	0.141	0.002	-0.715, -0.163
In-party (ref)	Out-party	-.425	0.141	0.003	-0.702, -0.147
Tax statement deadline					
No party (ref.)	In-party	0.136	0.134	0.311	-0.127, 0.398

	Out-party	-0.129	0.134	0.334	-0.392, 0.134
In-party (ref)	Out-party	-.265	0.135	0.049	-0.529, -0.001

AO11. Marginal means plots based on the results of the mixed-design ANOVA

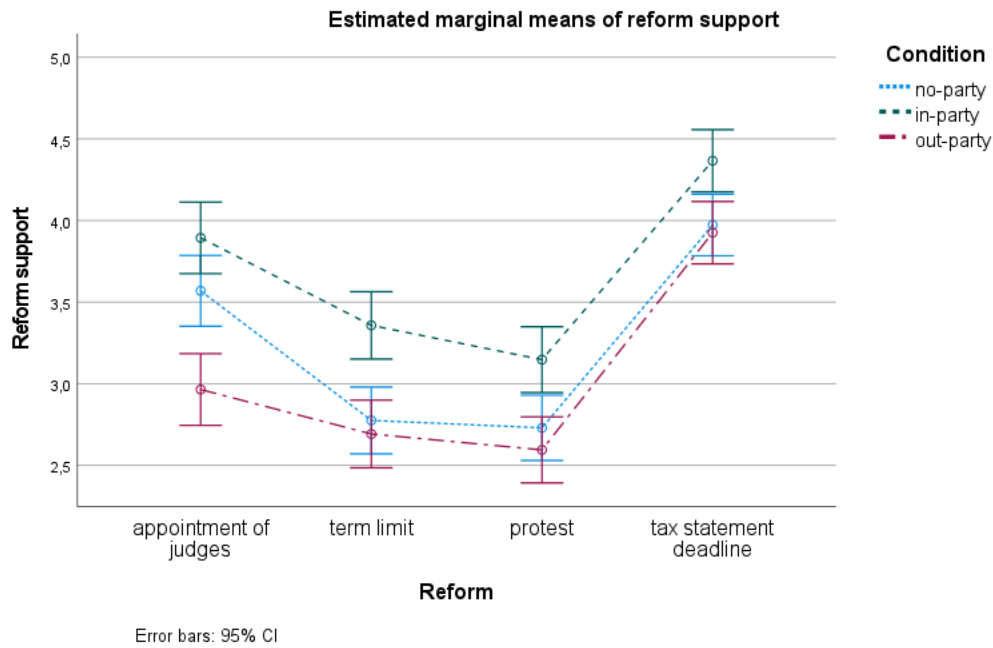**Figure 4A.** Estimated marginal means of reform support across reform conditions

Figure 5A. Estimated marginal means of benign intentions across reform conditions.
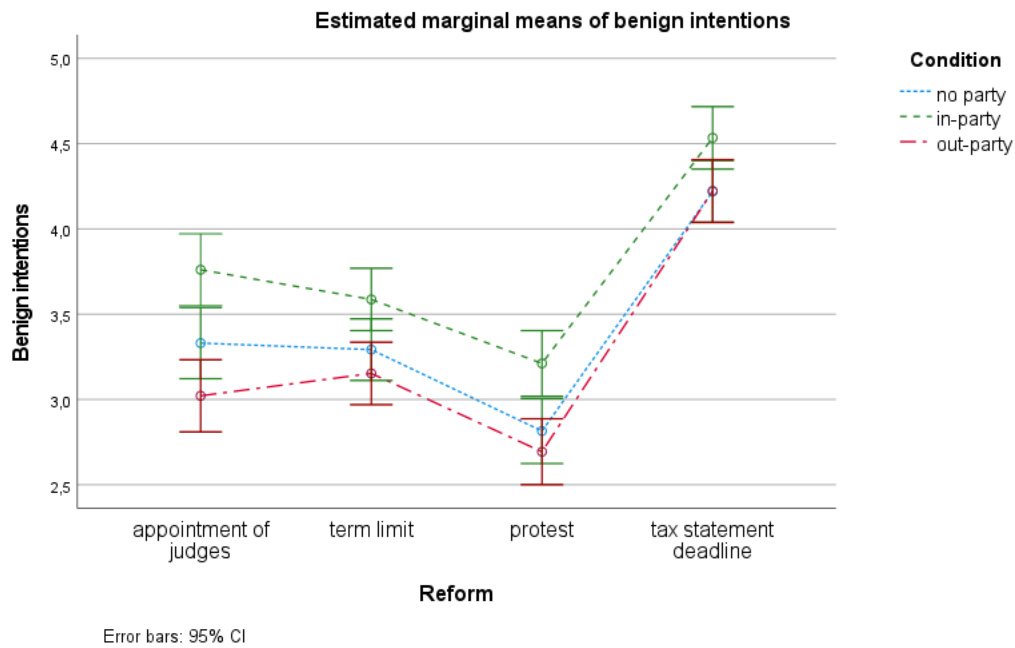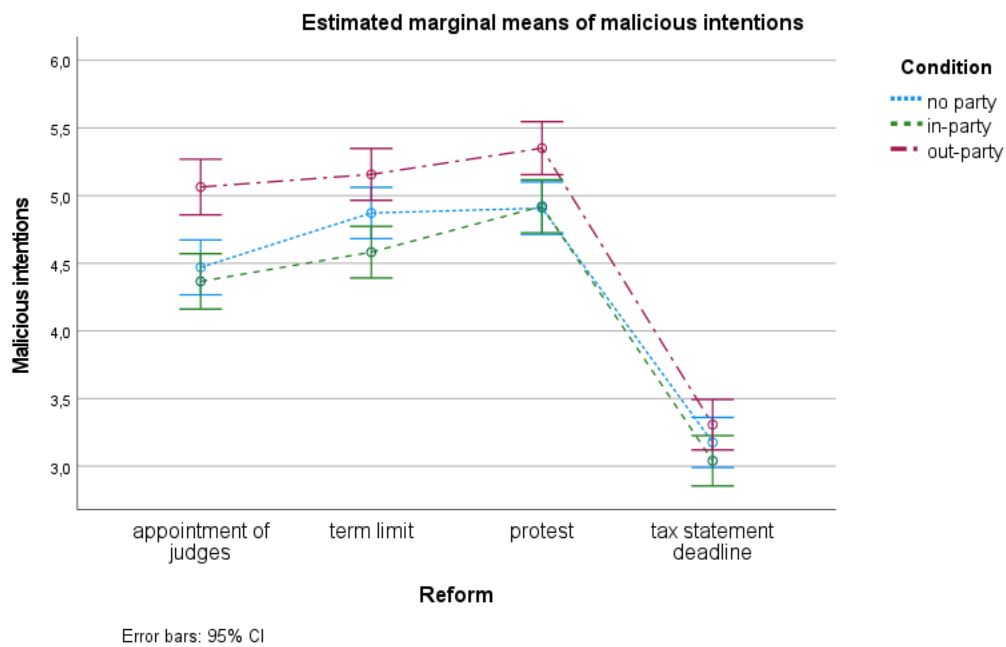

Figure 6A. Estimated marginal means of malicious intentions across reform conditions.

AO12. Deviations from pre-registration

The main study was pre-registered (available here: https://osf.io/tc2k8/?view_only=60c0d64fe613482287657149df6e7e33). Due to a low passing rate of the comprehension checks and to a mistake in programming by the survey company the study deviated in two main ways from the pre-registration:

- Participants were encouraged to read the text of the vignette again and to answer the questions about the text to make sure they are engaging with the presented scenario before they moved to the questions measuring the dependent variables (see AO6 for details)
- As the survey company by mistake added an open-ended question after the vignette asking participants to reflect on the scenario (which was not a part of the experimental design). additional data was collected without this question. The hypotheses in the first batch of data (presented in this paper) were only tested after the collection of additional data was completed. There were no obvious ways in which the open question could have affected the participants. On one hand it could have encouraged participants to think about the real situation in Poland and attribute the reforms to particular parties. which would make detecting the effects more difficult. On the other. it could have stimulated participants to think deeper about the presented scenario. The results of the additional data collection are reported in AO13.

Minor deviations from pre-registration:

- The numbers of hypotheses were changed. 'signal' was replaced with 'cue' and 'coalition' with 'party'. 'motivation' with 'intention'. and "skeptical about" with "vigilant of" for consistency (See Table 6A). None of these wording adjustments have substantive consequences for the hypotheses.
- The formulation of H3a was adjusted without changing the intention of the hypothesis. The original formulation was underspecified.

Table 26A. Comparison of pre-registered hypotheses with hypotheses presented in this article. Changes are marked in red.

As pre-registered	As reported in the article
H1: A signal of out-group coalition makes citizens more sceptical about the motivations behind the reforms of checks and balances (by comparison to no coalition signal and in-group coalition signal)	General expectation: A cue of out-group party makes citizens more skeptical about the intentions behind the reforms of checks and balances (by comparison to no party cue and in-group party cue)

H1a: A signal of in-group coalition increases the perception of benign intentions behind a reform altering the system of checks and balances (by comparison with no coalition signal and out-group coalition signal)

H1a: A cue of in-group party increases the perception of benign intentions behind a reform altering the system of checks and balances (by comparison with no party cue and out-group party cue)

H1b: A signal of out-group coalition increases the perception of malicious intentions behind a reform altering the system of checks and balances (by comparison with no coalition condition and in-group coalition signal)

H1b: A cue of out-group party increases the perception of malicious intentions behind a reform altering the system of checks and balances (by comparison with no party cue and in-group party cue)

H2: A signal of out-group coalition increases negative evaluation of the reform of checks and balances it (by comparison to a signal of in-group and no coalition signal)

H2: A cue of out-group party increases negative evaluation of the reform of checks and balances (by comparison to a cue of in-group and no party cue)

H3: Citizens respond differently to the reforms altering checks and balances than to administrative (non-political) reforms.

General expectation: Citizens respond differently to the reforms altering checks and balances than to administrative (non-political) reforms.

H3a: Citizens are more divided about the motivations behind the reforms of checks and balances than behind the administrative reforms.

H3: The effect of party-cue is expected to be stronger for the reforms with higher stakes (checks and balances) than reforms with lower stakes (administrative reform)

AO13. Analysis of differential treatment effects across party voters

To test for potential differential treatment effects across different party voters. I conducted factorial ANOVAs for each checks and balances reform (appointment of judges. protest registration. and presidential term limit) and for all three dependent variables (support. benign intentions. and malicious intentions). The analyses showed that the party vote (Law & Justice. Civic Coalition. Other. No vote) was significantly associated with each of these dependent variables meaning that the voters of different parties differ in their level of support for reforming checks and balances. In no party condition. the voters of Civic Coalition expressed the lowest support for each of the reforms. while the voters of Law and Justice expressed the highest support. Importantly. in none of the models was the interaction between the condition (no party. in-party. out-party) and party significant. showing that the treatment affected all party voters similarly. As an example. I present the results for the analyses of support for different checks and balances.

Table 27A. Mean support for the reforms of different checks and balances per party vote. Standard deviations in the brackets and number of participants per cell (N)

Reform	Treatment	PiS (Law & Justice)	KO (Civic Coalition)	Other party	No vote
		Vote			
Appointment of judges	No party	4.83 (1.77) N = 100	2.44 (1.94) N = 87	3.48 (2.04) N = 84	3.33 (1.82) N = 81
	In-party	5.45 (1.76) N = 83	2.88 (2.07) N = 85	3.47 (2.10) N = 95	3.85 (2.00) N = 82
	Out-party	4.28 (1.61) N = 77	1.94 (1.58) N = 101	3.02 (1.92) N = 90	2.93 (1.88) N = 75
Presidential term limit	No party	3.56 (2.08) N = 100	2.06 (1.74) N = 87	2.79 (1.93) N = 84	2.56 (1.59) N = 81
	In-party	4.66 (2.05) N = 83	2.58 (1.98) N = 85	2.96 (1.94) N = 95	3.30 (1.86) N = 82
	Out-party	3.62 (1.84) N = 77	1.99 (1.58) N = 101	2.53 (1.81) N = 90	2.87 (1.70) N = 75
Protest registration	No party	3.81 (2.04) N = 100	1.86 (1.45) N = 87	2.57 (1.95) N = 84	2.51 (1.48) N = 81
	In-party	4.72 (1.99) N = 83	2.08 (1.62) N = 85	2.64 (1.89) N = 95	3.25 (1.78) N = 82
	Out-party	4.08 (1.97) N = 77	1.86 (1.53) N = 101	2.23 (1.54) N = 90	2.83 (1.77) N = 82

Table 28A. Results of factorial ANOVAs for the support for the reforms of different checks and balances.

Reform	Predictor	Sum of Squaes	df	Mean Square	F	p
Appointment of judges	Intercept	974.33	11	88.58	25.05	<.001
	Condition	130.15	2	65.08	18.41	<.001
	Party vote	800.29	3	266.76	75.45	<.001

	Party vote x condition	14.82	6	2.47	0.70	.651
	Adj R² = 0.20 N = 1040					
Presidential term limit	Intercept	526.88	11	47.90	13.99	<.001
	Condition	90.40	2	45.20	13.20	<.001
	Party vote	411.91	3	137.30	40.09	<.001
	Party vote x condition	24.98	6	4.16	1.22	.296
	Adj R² = 0.12 N = 1040					
Protest registration	Intercept	748.50	11	68.05	22.53	<.001
	Condition	56.40	2	28.20	9.34	<.001
	Party vote	662.88	3	220.96	73.17	<.001
	Party vote x condition	29.14	6	4.86	1.61	.142
	Adj R² = 0.19 N = 1040					

AO14. Analysis of additional data collection

Checks

As in the main study. two manipulation- and one comprehension-check questions were asked. If a participant answered one of the three questions incorrectly. they were prompted to read the vignette text carefully again and answer the questions correctly to proceed with the study. Irrespective of the responses at the second try. participants were able to continue the survey. Table 29A shows the frequencies of the correct and incorrect answers to all there questions across the different party and reform conditions. As demonstrated in the Table the rate of correct answers improved at the second try and ranged from 59% (out-party proposed the tax statement deadline reform) to 70% (in-party proposed the reform of courts/appointment of judges).

Table 29A. Frequencies for answers on the manipulation and comprehension check questions. The 1st try pass/fail indicates the number of participants who got all three answers correct/who got one of the three answers incorrect at the first attempt. The total pass/total fail indicates the total number of participants who answered each of the questions correctly/incorrectly for each reform vignette after the second try.

Condition	1st try pass/fail	2nd try pass/fail	Reform	Question	total pass	total fail

no party

N = 125

72/53	84/41	courts	party	90	35
58%/42%	67%/33%		reform	104	21
			actor	100	25
73/52	82/43	term limit	party	86	39
58%/42%	66%/34%		reform	97	28
			actor	103	22
72/53	76/49	protest	party	79	46
58%/42%	61%/39%		reform	96	29
			actor	99	26
79/46	85/40	tax deadline	party	90	35
63%/37%	68%/32%		reform	102	23
			actor	99	26

in-party

N = 103

66/37	72/31	courts	party	83	20
64%/36%	70%/30%		reform	87	16
			actor	86	17
65/38	69/34	term limit	party	78	25
63%/37%	67%/33%		reform	84	19
			actor	89	14
60/43	68/35	protest	party	80	23
58%/42%	66%/34%		reform	87	16
			actor	88	15
66/37	68/35	tax deadline	party	78	25

64%/36%	66%/34%		reform	85	18
			actor	84	19

out-party

N = 122

68/54	74/48	courts	party	85	37
56%/44%	61%/39%		reform	102	20
			actor	96	26
65/57	74/48	term limit	party	83	39
53%/47%	61%/39%		reform	103	19
			actor	96	26
68/54	74/48	protest	party	84	38
56%/44%	61%/39%		reform	99	23
			actor	93	29
64/58	72/50	tax deadline	party	80	42
52%/48%	59%/41%		reform	104	18
			actor	96	26

Demographics

The total sample collected in the rerun of the study was N = 350 Polish citizens. The sample was 50 % female. with average age of 44 years old (SD = 15.1) and minimum of 18 and maximum of 69 years old. Around 5 % of participants indicated an education level below upper secondary. 56 % upper secondary. and around 40 % tertiary education.

In terms of voting behaviour. 27 % of participants indicated that they voted for Law and Justice. 26 % for Civic Coalition. 11 % for the Left. 8 % for Confederation. and 5 % for Polish Coalition. The same as in the main study. 13 % answered that they did not vote in elections.

These descriptive statistics show that that the two samples (in the main study and in the rerun) are very similar in terms of demographic characteristics.

Measures

The design of the re-run study was the same as of the main study. Because the sample in the re-run is much smaller than in the main study and. as a consequence. more measurement error is to be expected in the analyses. the answers about the three checks and balances reforms were pulled to create a more reliable dependent variable. Three scales averaging answers to questions about dependent variables across the different reforms:

- *Reform support scale* (Cronbach's alpha = .90): appointment of judges reform justified. agree with the reform of appointment of judges. presidential term limit reform justified. agree with the reform of presidential term limit. protest registration reform justified. agree with the reform of protest registration
- *Benign motivations scale* (Cronbach's alpha = .83): reform of courts is to assure independence of judges. reforms of courts is to strengthen democracy. reform of presidential term limit is to assure that the president can check the government in line with the mechanisms of checks and balances of power. reform of presidential term limit is to strengthen democracy. reform of protest registration is to assure that citizens can protest in a safe way. reform of protest registration is to strengthen democracy
- *Malicious motivations scale* (Crobach's alpha = .84): reform of courts is to have fewer limitations on their power. so the government can do whatever it wants. reforms of courts is to undermine democracy. reform of presidential term limit is to is to have fewer limitations on their power. so the government can do whatever it wants. reform of presidential term limit is to undermine democracy. reform of protest registration is to is to have fewer limitations on their power. so the government can do whatever it wants. reform of protest registration is to undermine democracy

For the tax statement deadline reform. the dependent variables were kept the same as in the main study:

- *Reform support scale* (Cronbach's alpha = .92): tax statement deadline reform justified. agree with the tax statement deadline reform
- *Benign motivations* (Cronbach's alpha = .83): tax statement deadline is to make it easier for citizens to file taxes. tax statement deadline reform is to improve the functioning of state administration
- *Malicious motivations* (Cronbach's alpha = .81): tax statement deadline is to make it more difficult for citizens to file taxes. tax statement deadline reform is to worsen the functioning of state administration
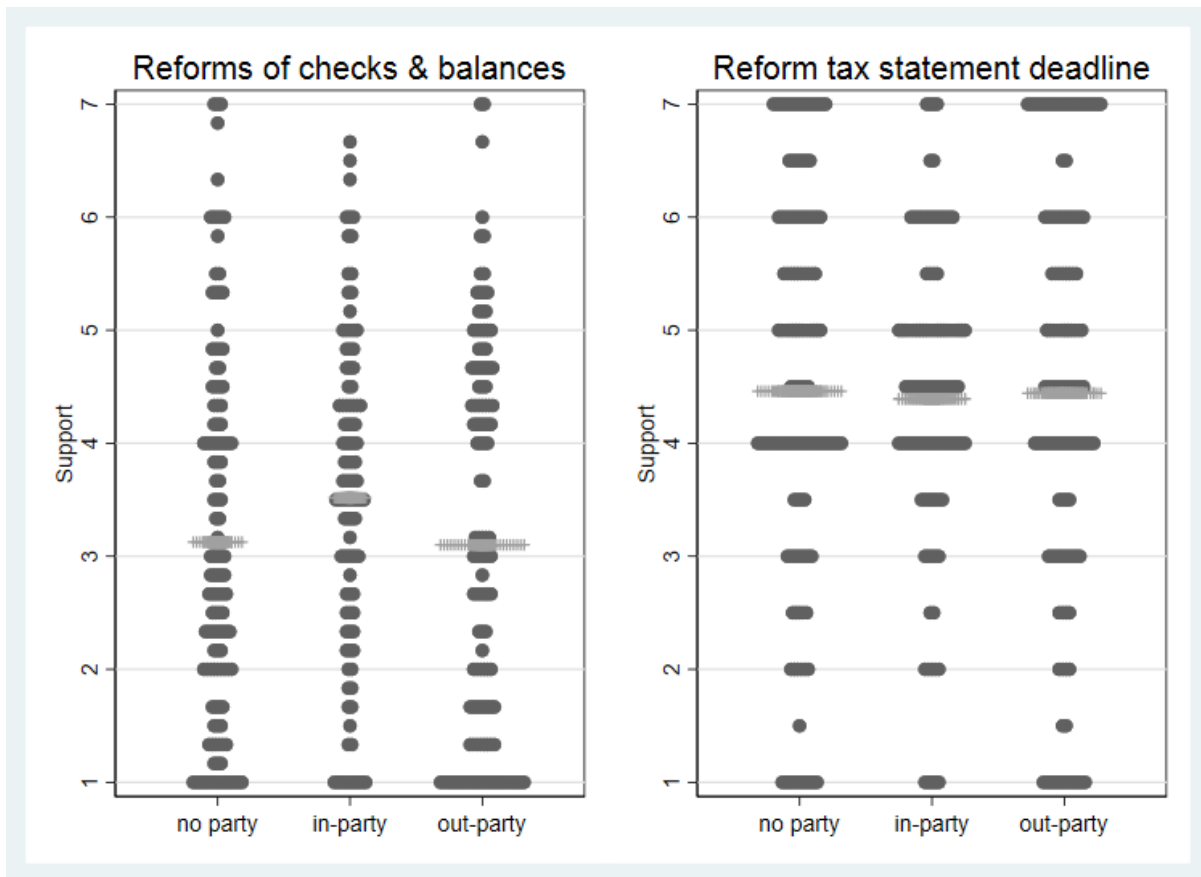
Results

Three ANOVA analyses were conducted to test the effects of party cues on each of the dependent variables related to the checks and balances reforms. and three ANOVA analyses tested the effect of party cues on the dependent variables related to the tax statement deadline reforms. Figures 1A. 2A. and 3A show pairs of graphs with the effects of party cues on support for. attribution of benign

motivations. and attribution of malicious motivations to checks and balances reforms and to tax statement deadline reform.
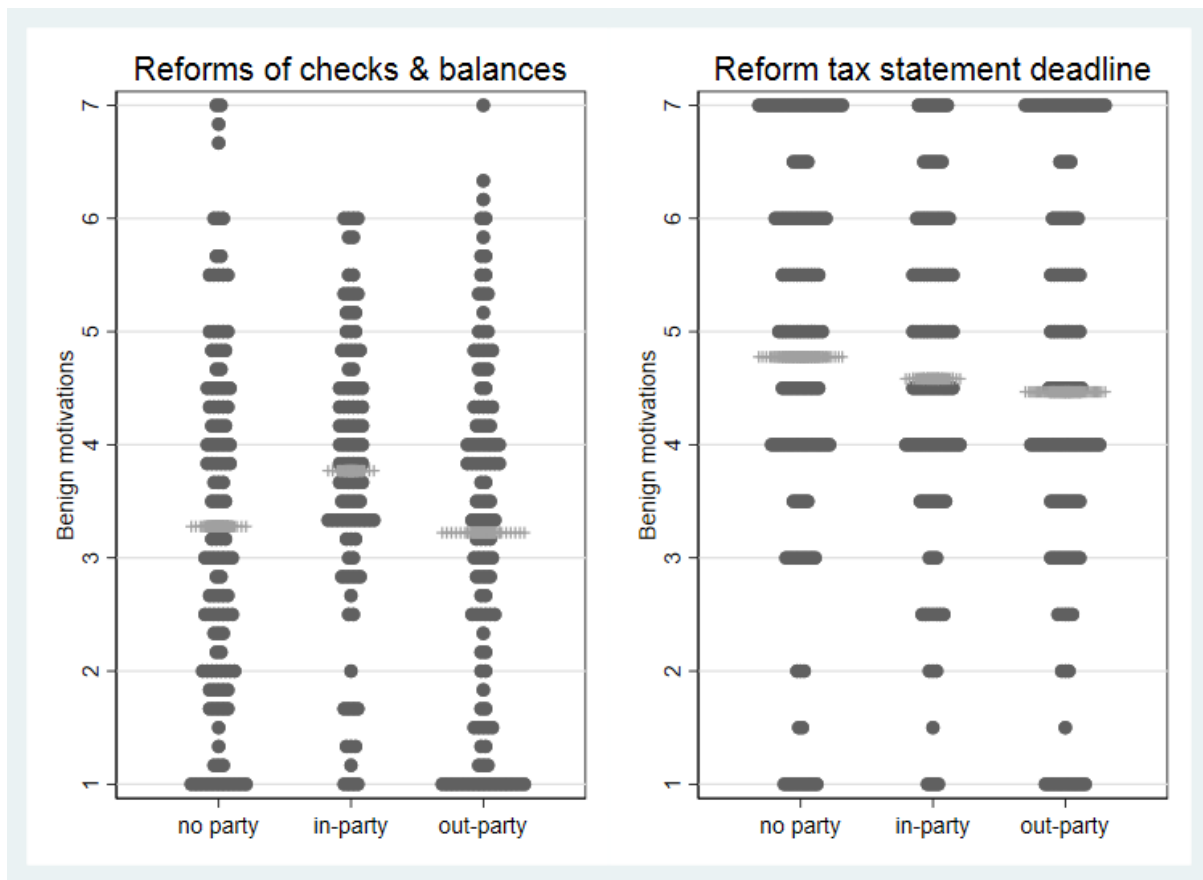
The results of the ANOVA analysis showed no statistically significant effect of party cues on the *support* for checks and balances reforms. F (2. 347) = 2.30. p = .10. Given this analysis relies on a much smaller sample. the power to detect this effect was much lower than in the planned main study. Therefore. it is important to investigate the data patterns more closely. The data shows a very similar pattern to the results of the main study. Participants who received an in-party cue supported the reforms of checks and balances more (M = 3.52. SD = 1.45) than participants in no party (M = 3.13. SD = 1.63) and out-party condition (M = 3.10. SD = 1.70). There was no statistically significant effect of party cues on the support for the tax statement deadline reform (F (2. 347) = 0.04. p = .96). Figure 1A shows that on average reforms of checks and balances are supported less than tax statement deadline reform and the differences between means across party cues conditions are larger in the case of checks and balances reforms.

Figure 7A. Support for the reforms of checks and balances and for the reform of tax statement deadline. The dotplots show the data distribution for support of the reforms across the three party-cue conditions. The grey lines indicate the means.

The results of the ANOVA analysis showed statistically significant effect of party cues on the *attribution of benign intentions* to the reforms of checks and balances. $F_{(2, 347)} = 4.61$. $p = .01$. Participants who received an in-party cue attributed significantly more benign motivations to the reforms of checks and balances (M = 3.77. SD = 1.24) than participants in no party (M = 3.28. SD = 1.55) and out-party condition (M = 3.22. SD = 1.55). There was no statistically significant effect of party cues on the attribution of benign motivations to the tax statement deadline reform ($F_{(2, 347)} = 0.99$. $p = .37$). Figure 2A also shows that on average the motivations for reforms of checks and balances are considered as less benign than motivations for the reform of tax statement deadline. Also. the differences between means across party-cues conditions are larger in the case of checks and balances reforms. Importantly. we see much higher concentration of low scores indicating low level of benign motivations attributed to the reforms of checks and balances and a much higher concentration of high scores indicating high level of benign motivations attributed to the reform of tax statement deadline.

Figure 8A. Attribution of benign motivations to the reforms of checks and balances and to the reform of tax statement deadline. The dotplots show the data distribution of benign motivations of the reforms across the three party-cue conditions. The grey lines indicate the means.

The results of the ANOVA analysis showed statistically significant effect of party cues on the *attribution of malicious intentions* to the reforms of checks and balances. F (2. 347) = 4.62. p = .01. Participants who received an in-party cue attributed significantly less malicious motivations to the reforms of checks and balances (M = 4.32. SD = 1.36) than participants in no party (M = 4.88. SD = 1.51) and out-party condition (M = 4.80. SD = 1.51). There was no statistically significant effect of party cues on the attribution of malicious motivations to the tax statement deadline reform (F (2. 347) = 0.31. p = .74). Figure 3A shows a reversed patterns of data to the patterns for benign motivations. On average the motivations for reforms of checks and balances are considered as more malicious than motivations for the reform of tax statement deadline with more high scores for malicious motivations recorded in the case of checks and balances reforms and more low scores recorded in the case of tax statement deadline.

Figure 9A. Attribution of malicious motivations to the reforms of checks and balances and to the reform of tax statement deadline. The dotplots show the data distribution of benign motivations of the reforms across the three party cue conditions. The grey lines indicate the means.
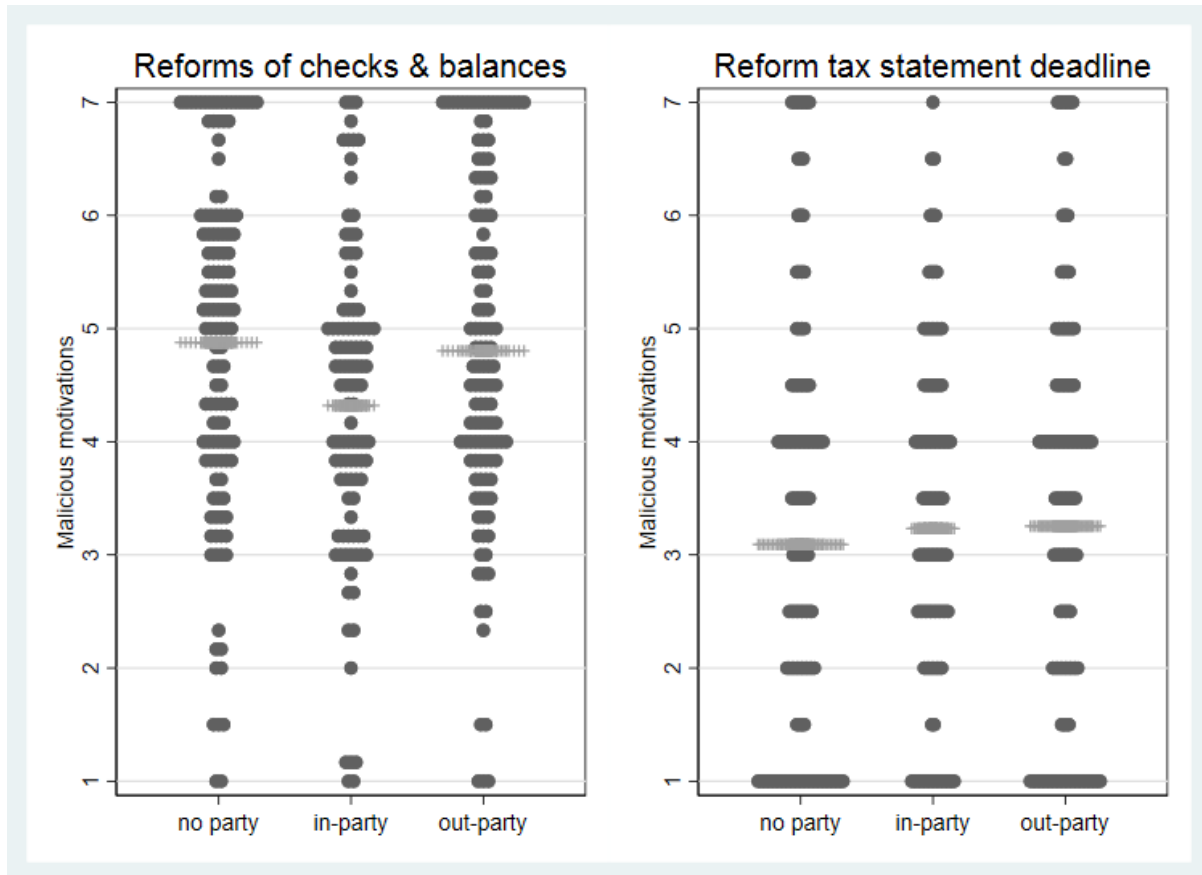
AO15. Analysis of the restricted sample

To check whether the manipulation check prompt affected the results of the experiment, I analyzed the data with the sample restricted to the participants wo answered all the checks questions correctly and were not prompted to read the text again and answer the check questions again. This restricted sample resulted in N = 601 (no-party n = 196, in-party n= 201, out-party n = 204). Tables 30A, 31A, and 32A present the results of the mixed-design ANOVAs. There are no substantive differences between the results of the analysis of the full sample presented in the manuscript and the results of the restricted sample.

Table 30A. Results of the mixed ANOVA for *reform support* (sphericity not assumed), restricted sample (N = 601)

Factor	Test	Type III SS	df	M²	F	p	Partial η²
Reform (within effect)	Sphericity Assumed	1012.33	3	337.44	158.426	<.001	0.209
	Greenhouse-Geisser	1012.33	2.850	355.16	158.426	<.001	0.209
	Huynh-Feldt	1012.33	2.875	352.11	158.426	<.001	0.209
	Lower-bound	**1012.33**	**1.000**	**1012.33**	**158.426**	**<.001**	**0.209**
Party-cue (between effect)	Intercept	22681.99	1	22681.99	3041.323	<.001	0.836
	Effect	212.00	2	106.002	14.213	**<.001**	0.045
	Error	4459.84	598	7.458			
Reform * party-cue (within effect)	Sphericity Assumed	51.507	6	8.585	4.030	0.001	0.013
	Greenhouse-Geisser	51.507	5.701	9.035	4.030	0.001	0.013
	Huynh-Feldt	51.507	5.750	8.958	4.030	0.001	0.013
	Lower-bound	**51.507**	**2.000**	**25.754**	**4.030**	**0.018**	**0.013**
Reform * party-cue (contrast)	Linear	36.33	2	18.167	7.376	**0.001**	0.024
	Quadratic	6.97	2	3.482	1.648	0.193	0.005
	Cubic	8.21	2	4.105	2.263	0.105	0.008

Table 31A. Results of the mixed ANOVA for *benign intentions* attribution (sphericity not assumed), restricted sample (N = 601)

Factor	Test	Type III SS	df	M²	F	p	Partial η²
Reform (within effect)	Sphericity Assumed	1056.89	3	352.298	176.313	<.001	0.228
	Greenhouse-Geisser	1056.89	2.664	396.746	176.313	<.001	0.228
	Huynh-Feldt	1056.89	2.686	393.498	176.313	<.001	0.228
	Lower-bound	**1056.89**	**1.000**	**1056.893**	**176.313**	**<.001**	**0.228**
Party-cue (between effect)	Intercept	26440.88	1	26440.882	3920.872	<.001	0.868
	Effect	172.31	2	86.155	12.776	**<.001**	0.041
	Error	4032.69	598	6.744			
Reform * party-cue (within effect)	Sphericity Assumed	18.41	6	3.068	1.535	0.163	0.005
	Greenhouse-Geisser	18.41	5.328	3.455	1.535	0.171	0.005
	Huynh-Feldt	18.41	5.372	3.427	1.535	0.171	0.005
	Lower-bound	18.41	2.000	9.204	1.535	0.216	0.005
	Linear	15.29	2	7.646	3.015	**0.050**	0.010

| Reform * party-cue (contrast) | Quadratic | 1.31 | 2 | 0.656 | 0.343 | 0.710 | 0.001 |
| | Cubic | 1.81 | 2 | 0.902 | 0.584 | 0.558 | 0.002 |

Table 32A. Results of the mixed ANOVA for *malicious intentions* attribution (sphericity not assumed), restricted sample (N = 601)

Factor	Test	Type III SS	df	M^2	F	p	Partial η^2
Reform (within effect)	Sphericity Assumed	2170.77	3	723.59	347.050	<.001	0.367
	Greenhouse-Geisser	2170.77	2.698	804.55	347.050	<.001	0.367
	Huynh-Feldt	2170.77	2.721	797.91	347.050	<.001	0.367
	Lower-bound	**2170.77**	**1.000**	**2170.77**	**347.050**	**<.001**	**0.367**
Party-cue (between effect)	Intercept	49052.90	1	49052.90	7391.744	<.001	0.925
	Effect	195.44	2	97.72	14.726	**<.001**	0.047
	Error	3968.43	598	6.64			
Reform * party-cue (within effect)	Sphericity Assumed	25.52	6	4.25	2.040	0.057	0.007
	Greenhouse-Geisser	25.52	5.396	4.73	2.040	0.065	0.007
	Huynh-Feldt	25.52	5.441	4.69	2.040	0.064	0.007
	Lower-bound	25.52	2.000	12.76	2.040	0.131	0.007
Reform * party-cue (contrast)	Linear	14.79	2	7.40	3.114	**0.045**	0.010
	Quadratic	0.87	2	0.43	0.192	0.826	0.001
	Cubic	9.87	2	4.93	3.056	0.048	0.010

Figure 8A. Violin plots show the support for reforms (restricted sample) across conditions of party-cue (in-party, no party, and out-party). The white dot shows the median, the dark bar shows the interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density.
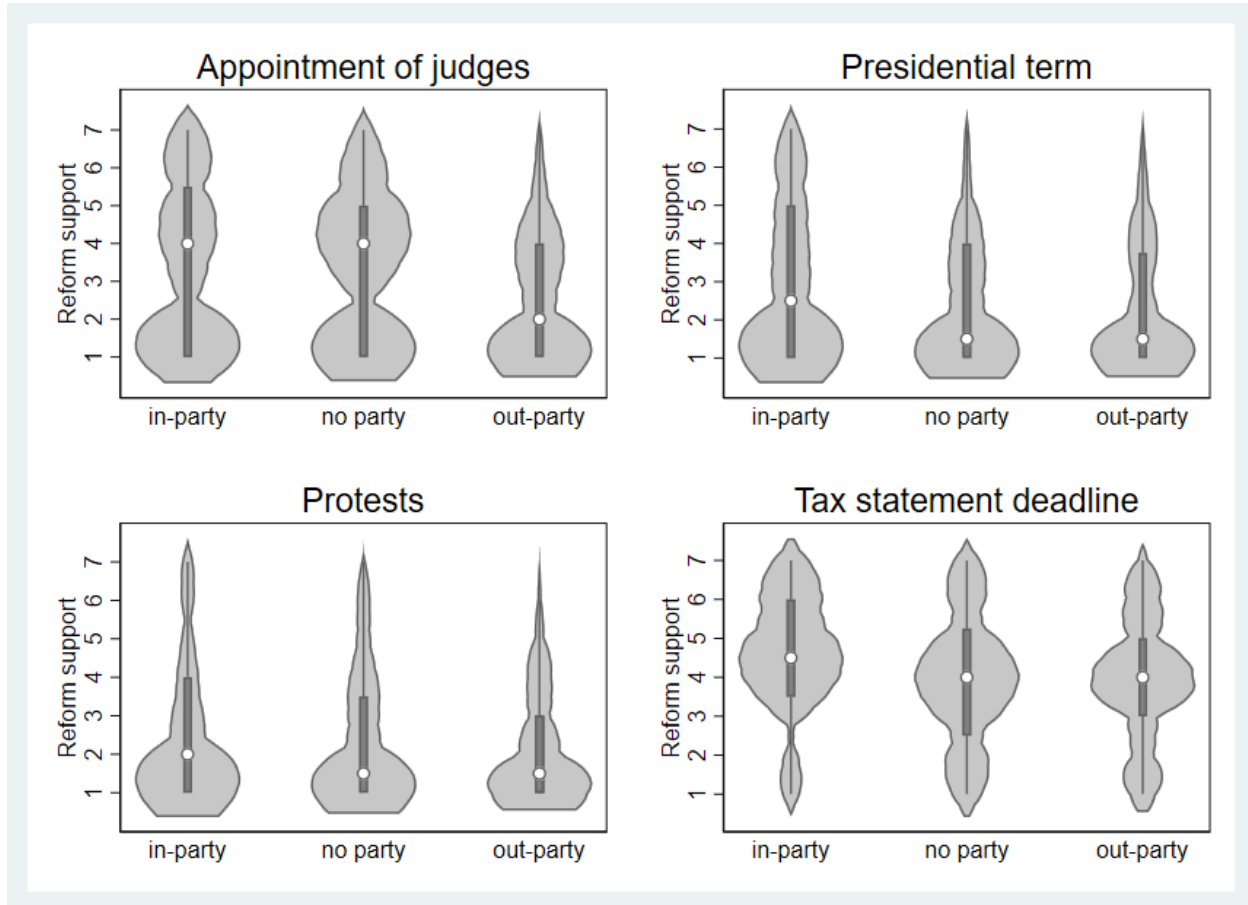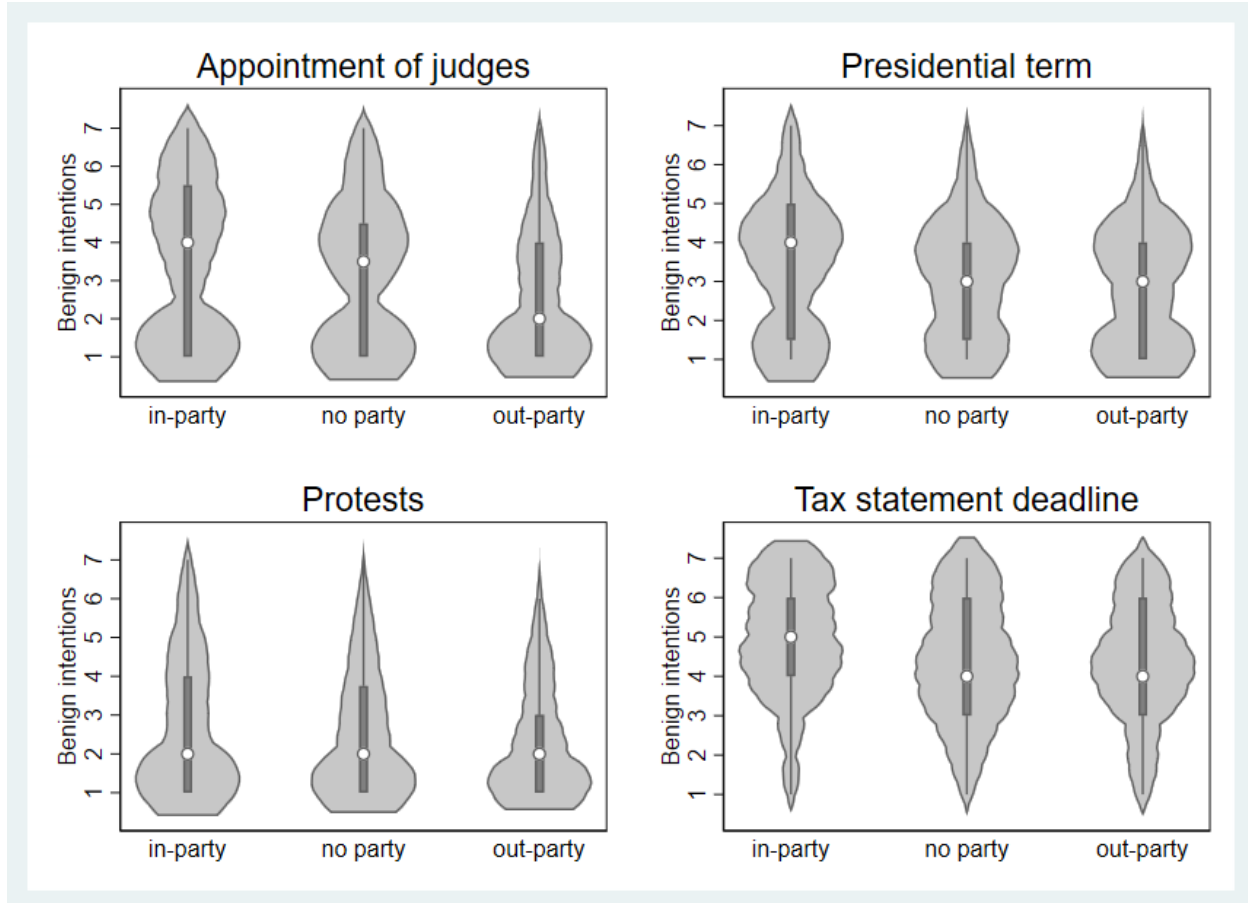
Figure 9A. Violin plots show the attribution of benign intentions (restricted sample) across conditions of party-cue (in-party, no party, and out-party). The white dot shows the median, the dark bar shows the interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density.
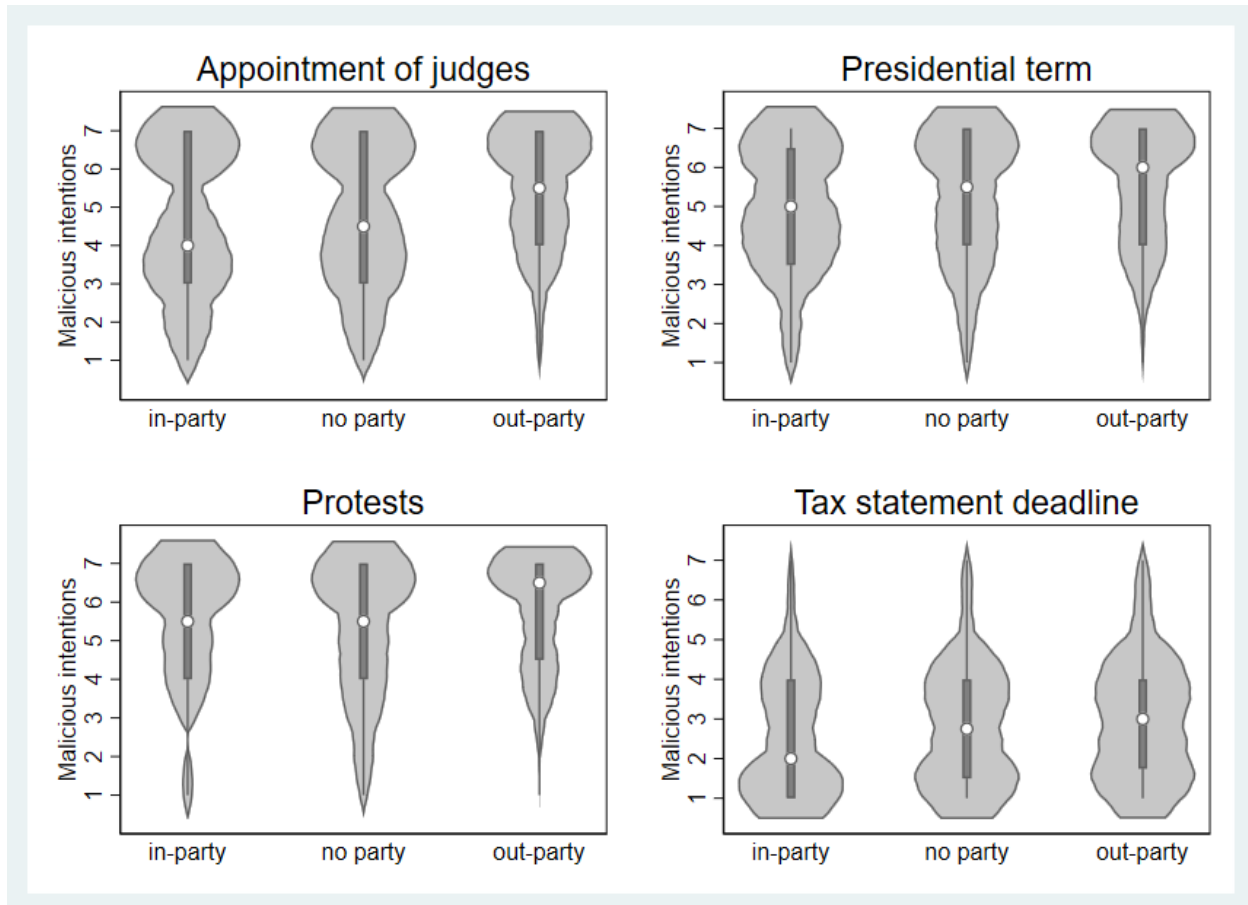
Figure 10A. Violin plots show the attribution of malicious intentions (restricted sample) across conditions of party-cue (in-party, no party, and out-party). The white dot shows the median, the dark bar shows the interquartile range, and the vertical line extends to the highest and lowest value. The violin shows the estimated kernel density.