Election Fraud Information, Punishment, and Political Trust: Evidence from a Survey Experiment in Colombia, Mexico, and Russia*

April 10, 2024

Abstract: In developing democracies and even electoral autocracies, credible allegations of electoral manipulation regularly lead to political interventions such as the dismissal of electoral staff or court punishments of alleged perpetrators. So far, scholars have shown that consciousness of election fraud lets individuals withdraw support from candidates, institutions and governments that are supposedly involved in manipulation. In this study, we first argue that election fraud information lets individuals extrapolate legitimacy loss even to political institutions that are unrelated to electoral events. Second, we investigate whether political interventions mitigate decays in diffuse support. Using a pre-registered online survey experiment in Colombia, Mexico, and Russia ($n = 2,057$), we (i) present empirical evidence for the spillover effect of election fraud information to political institutions unrelated to the fraud stimulus (ii) and show how spillovers largely persist even after political interventions. Our findings hold important implications for the study of developing democracies and electoral autocracies.

Keywords: *Election Fraud; Diffuse Support; Political Trust; Electoral Courts; Survey Experiment; Causal estimation.*

Which consequences do information about electoral fraud have for citizens' relationship towards their political system? Since the 'electoral revolution' that surged since the mid-twentieth century led to a dramatic increase in the number of electoral events (Norris 2014), multiparty elections have become omnipresent across new democracies and electoral authoritarian regimes worldwide. The conduct of these, however, is frequently accompanied by publicly voiced doubts about their integrity. For instance, nearly 80% of all federal elections in non-established democracies nowadays are monitored 'on the ground' by international observers (Kelley 2012*a*; Hyde 2011) and around half of all observed elections had led international missions to declare problems of moderate or high magnitude (Kelley 2012*a,b*).

Among the citizens themselves, credible information about electoral crimes can hold several behavioral and attitudinal consequences. Becoming conscious of electoral malpractice has been shown to lead to participation in popular protests and violent uprisings (Daxecker 2012). The literature has furthermore amassed a wealth of knowledge on the effects of election fraud perceptions on individuals' attitudes towards their political authorities. First, scholars have examined how information about electoral misconduct shape individuals' evaluations of the electoral process itself. For instance, Robertson (2017) shows that providing citizens with critical reports of election observation missions considerably reduces their perceived levels of electoral integrity. Second, citizens that are conscious about misbehavior withdraw support from those candidates that are allegedly involved in malpractice (Reuter and Szakonyi 2019) and express lower levels of legitimacy for the political regime that surged out of an electoral process that is perceived to be fraudulent (Williamson 2021).

This article holds two main contributions. First, we draw on theories of information processing and outline a mechanism of *attitudinal spillover* which states that individuals extrapolate specific fraud allegations to their confidence in the political system itself. The theory argues that even when information about fraud is attributed to unique political actors, citizens tend to relate these to political institutions that are unconnected to electoral administration. In contrast to the prior literature, the presence of such attitudinal spillovers predicts that consciousness about electoral misconduct will not let individuals merely detach from political authorities that can be directly linked to misbehavior and the regime that surged out of an illegitimate election process, but rather holds implications that are considerably more detrimental. Following the mechanism that we

outline, consciousness about electoral malpractice can lead citizens to withdraw approval from the political system as a whole.

Second, we investigate whether the spillover effect induced by fraud information is endogenous to the reactions of other actors of the political system. Credible allegations of electoral manipulation regularly induce political interventions in developing democracies and even electoral autocracies. In the year 2015 in Colombia, a controversial congressman was stopped by local authorities two days before the country's regional elections while transporting a sum of over 200 million Colombian *pesos*. Under allegations of vote buying, the congressman was dismissed from office (El Espectador 2015). After the Bolivian presidential elections of 2019, the Organization of American States has voiced criticisms around statistical patterns among late-counted votes (OAS 2019), shortly after which ex-president Evo Morales was asked to resign by the country's military chief and head of the police (Idrobo, Kronick and Rodríguez 2022). In Russia, rumors and allegations of vote irregularities were widespread after the country's 2016 legislative elections. After allegations culminated, Russian election chiefs sacked seven officials from polling stations in a region were contestations were most severe from their posts, proclaiming that observed irregularities were 'isolated incidents' that do not put the results of the election into doubt (Reuters 2016).

Scholars of electoral integrity have put little effort into understanding how such political interventions to fraud allegations interplay with political attitudes. Notably, attitude shifts induced by information on electoral malpractice might be *exacerbated* (amplification effect) or *mitigated* (suppression effect) by how other political actors respond. For instance, alleged perpetrators being removed from their posts in the electoral commission or public court rulings on electoral crimes may send out signals of the political system's professionalism, autonomy, and commitment to a fair electoral process (Kerr and Wahman 2020). Successful convictions of alleged perpetrators might therefore function as signals of at least some level of horizontal 'checks and balances', mitigating individuals' depressed levels of diffuse support. On the other hand, especially for those individuals that hold positive views of the government or political regime in place, interventions like these might provide legitimacy to claims that otherwise are dismissed as political rhetoric. Under this mechanism, political interventions might *induce* rather than prevent spillover effects. Grasping such dynamics is crucial for understanding the real-life impact of fraud information, as citizens are

not only exposed to disseminated information about electoral malpractice, but also perceive how actors of the political system *respond.*

We present evidence from a pre-registered online survey experiment conducted in Colombia, Mexico and Russia ($n = 2,057$) assessing (i) the presence of attitudinal spillovers of election fraud information to political institutions that are unrelated to electoral events and (ii) how punishment of alleged perpetrators exacerbate or mitigate decays in political trust. Since much of the previous literature has focused on the analysis of large cross-national survey data, our empirical analysis first showcases that these are unable to answer questions such as those that we pose here using 48,953 respondents across 48 countries from Wave 7 (2017-2020) of the World Values Survey (WVS). Even after applying a range of state-of-the-art matching algorithms for causal estimation combined with various robustness checks and a Bayesian estimation approach, we cannot distinguish spillover effects on political institutions (that are dictated by theory) from spillovers on non-political institutions (that should not be present in theory). We then present evidence from our experiment adding two original findings to the current literature. First, exposing individuals to information about electoral misconduct induces negative spillovers to trust in components of the political system that are not tied to elections among government opponents and supporters. Second, across a range of subsample analyses, we find no evidence for an amplification effect of electoral staff dismissal or court punishments among government supporters and opponents. However, for most group comparisons, spillover effects persist after disseminating information on punishments of alleged perpetrators. Finally, we show that if multiple credible punishments are in place, negative shifts in attitudes can be mitigated.

The main conclusion of our article is two-fold. On a general level, the consequences of administering election fraud for public support are even more detrimental than currently acknowledged by the literature. This is because information on electoral misconduct even induces shifts in public support towards components of the political system that are no beneficiaries of manipulation and are not related to electoral administration. Second, we cast light on the under-acknowledged role that political interventions which punish alleged perpetrators of fraud play. A combination of several credible punishments *can* mitigate (or even remove) negative attitude shifts of individuals exposed to fraud information, although this effect is far from consistent across countries and institutions.

This article hence closes on a cautious note: While the spillover effect of fraud information is consistent, the mitigating effect of effective punishments is not.

2 Election Fraud Information, Punishment, and Political Trust

In this section, we outline a theory of how the acquisition of new information about the integrity of domestic elections will affect the amount of trust that citizens place in the institutions of their broader political system. Essentially, this comes down to defining an argument of why individuals will extrapolate information about electoral misconduct to political institutions that are unrelated to electoral administration. Afterwards, we discuss how the interventions of other political actors can *amplify* or *mitigate* such spillover effects.

2.1 Election Fraud Information and Attitude Extrapolation

Scholarly contributions that examine the attitudinal nexus between citizens and the state commonly refer to the work of David Easton (1965, 1975) on 'system support' as a joint conceptual heritage. The theoretical distinction that is most relevant for our argument is the classical discrimination between diffuse and specific levels of support. Specific support refers to the relationship between members of a system and the specific actions and decisions of political authorities that reside within its institutions. As such, specific support relates to the evaluations of the day-to-day actions of political leaders, and are highest if perceived outputs match citizens' articulated demands (Easton 1975, p. 438). In contrast, diffuse support describes individuals' generalized attachment to the political system. According to Easton, *"[diffuse support] refers to evaluations of what an object is or represents [..] not of what it does. [..] Whereas specific support is extended only to the incumbent authorities, diffuse support is directed towards offices themselves as well as their individual occupants. More than that, diffuse support is support that underlies the regime as a whole and the political community."* (Easton 1975, pp. 444-445). Hence, diffuse support is *a priori* expected to be more durable than citizens' performance evaluation of specific political authorities. While positive evaluations of actors' performance is volatile and comes with consistent rise and fall, diffuse political support for the entity of the political system is in general thought to be long-lasting.

Early work on the concept of political support did almost exclusively focus on the relation between citizens and the state in the context of the United States and other advanced industrialized democracies (Easton 1965, 1975; Citrin 1974). Importantly, already in their seminal work on popular support for authoritarian regimes, Geddes and Zaller (1989) have argued that political reasoning in democracies and autocracies can be expected to operate in similar ways and a range of studies have evaluated concepts derived from the distinction of specific and diffuse support in autocratic settings as well (Reuter and Szakonyi 2019; Frye and Borisova 2019). In addition, it has been shown that measurement equivalence of the most prominent operationalizations of diffuse support holds across a variety of regime types (Schneider 2017).

In the first place, we can expect that credible fraud information evolving around electoral contests will lower citizens' confidence in such. For instance, both Robertson (2017) as well as Bush and Prather (2018) show that confronting voters with criticisms from election observer groups reduces their evaluations of electoral quality and the legitimacy of the electoral process. In the literature evolving around system support, it has long been argued that attitudes about the performance of individual objects that are commonly associated with specific support can spill over to more generalized attachments towards the political system (Bowler and Karp 2004). It is important to note that this goes contrary to an assumption by which citizens' evaluations of *political actors* are unrelated to their evaluation of their *political institutions*.

Empirically, spillover-like effects are a well-established phenomenon in various branches of attitudinal research. These can be understood as specific manifestations of a more general psychological principle commonly referred to as the 'halo effect' by which individuals ascribe characteristics to a person or an object based on their evaluation of other empirically observable object-related characteristics even if the individual traits are unrelated to each other (Thorndike 1920; Palmer and Peterson 2016). Such spurious inferences may result from individuals' inability to differentiate between different characteristics and may even occur if there is sufficient information to allow for independent assessments in the first place (Nisbett and DeCamp Wilson 1977). Regarding citizens' evaluations of actors and institutions, it has been shown that trust in national institutions transcends to trust that is placed in the international arena, extrapolating federal-level experiences to European institutions (Torcal and Christmann 2019) and international organizations (Dellmuth and Tallberg 2015). Studying attitudinal spillovers between national institutions, Bowler and Karp

(2004) show that political scandals of individual politicians have the power to erode confidence in executive institutions and the government in general. Notably, such spillover effects may either be the result of evaluating a series of repeated outputs over a long time series that can change even fundamental beliefs, or chief, salient, and decisive short-term experiences that transform into fundamental attitudes more rapidly.

We hypothesize that information about electoral fraud provide the kinds of short-term information that dis-attaches from the volatile performance of political actors and transforms into generalized evaluations even of other components of the political system. Essentially, this is based on a two-step argumentation line. First, as elections lie at the core of democratic accountability and are the one crucial element common to all and even minimalist definitions of democracy (Przeworski, Stokes and Manin 1999), systematic misbehavior that evolves around the decisive process of elections is likely to be taken as informative not only of what a specific political object does, but even towards the system that it represents. Hence, the central place of well-conducted elections in the constitution of a democratic political system lets evaluations of the electoral process fundamentally differ in their nature from perceived output that is generated through the short-term and volatile performance of individual office holders. This provides election-related information with the general *possibility* for producing spillovers. Second, it has been shown by a variety of authors that citizens tend to fail in *distinguishing* their attitudes towards individual components of the multidimensional political system. This is most evident as the political sphere is usually described to be too complex to understand even for highly informed individuals (Zaller 1992) and as citizens need lower-complexity informational cues to maneuver their perceptions of political affairs. Empirically, scholars have found that support levels for different political institutions or entities are highly correlated with each other and are often hard to disentangle within individual respondents (Hooghe and Marien 2012; Mishler and Rose 2001). It is these two observations that build the premises from which motivate the first central claim of this paper. The centrality of election-related information for citizens' evaluations of the political system which provides the possibility for spillover fused with the general tendency of individuals to fail distinguishing support for different institutions leads us to formulating the first main hypothesis:

6

Hypothesis 1: When exposed to information about electoral fraud, individuals show less confidence in institutions of the political system that are unrelated to electoral administration.

2.2 Previous Literature

While examining the empirical interrelations between operationalizations of specific and diffuse support is a decade-old endeavor, the attempt to link system support with election fraud information is rather new. Our specific research strategy tabs into a broader field of previous studies that have examined related phenomena which are relevant for our hypothesis. A branch of studies focused on the relation between 'objective' measures of electoral manipulation and average levels of diffuse support. Mauk (2019) globally assembles expert-coded judgements of federal-level electoral integrity from the Varieties of Democracy dataset and relates these to national levels of political trust, finding little evidence that objectively coded factual levels of electoral integrity are related to country-specific average values of political trust. Exploiting largely exogenous variation in a survey conducted in Moscow around the 2011 Russian *Duma* elections, Frye and Borisova (2019) reach similar conclusions and find that simply the mere event of an allegedly fraudulent election does not significantly reduce levels of diffuse support when comparing those individuals that have been surveyed after the election with the respondents whose data has been collected beforehand. These studies carry the obvious shortcoming that they calculate effects of fraud indicators that have been collected *a posteriori* on all individuals that might potentially have become aware of such information. However, as Mauk (2019) outlines, actual electoral malpractice does not necessarily need to be related closely to citizens' individual perceptions of electoral integrity (see also van Ham 2015), since these crucially depend on factors such as a sufficiently free media environment to report about electoral inferences and one's individual political interest to become informed through media channels.

A different group of authors directly exposes individuals to information about electoral malpractice and investigates how becoming aware of misbehavior affects citizens' beliefs about the electoral process (Robertson 2017; Bush and Prather 2018) and their support for candidates that are allegedly involved in malpractice (Reuter and Szakonyi 2019). These studies shed great light into individuals opinion-formation dynamics as a response to sensitive information, but restrict

7

their analyses to attitudes that are directly linked to the electoral process or to specific evaluations of office holders rather than examining underlying attachments towards the political system.

Using data from the World Values Survey, Norris (2014, 2019) exploits a cross-sectional design and shows that even when controlling for a range of attitudinal and socio-demographic factors, expert evaluations and perceptions of electoral integrity are still correlated to an array of items as wide as confidence in elected institutions such as parliaments and governments, overall satisfaction with performance of democracy and respect for human rights. Obviously, using such cross-sectional strategies, it as hard to disentangle whether perceptions of electoral integrity and institutional confidence are simply observed jointly, or if one determines the other, falling short in testing a spillover theory as outlined here. Even if a directional effect exists, the causality chain might well go into the opposite direction. It is not less reasonable to assume that stable underlying beliefs such as confidence in political authorities pre-structure individuals' evaluations of specific political events such as electoral contests. In the piece that is most relevant for our research, Williamson (2021) shows how confronting citizens with condemnations of international election monitors can reduce expressed legitimacy in the political regime that surged out of an allegedly fraudulent process. Using correlational analysis from eight Arab countries and a survey experiment conducted in authoritarian Egypt and Morocco, he shows how perceptions of electoral misconduct hinder both attitudinal and behavioral compliance with a regime's rule. This investigation of individual conformity with the direct beneficiary of misbehavior is considerably different from our spillover perspective which investigates effects even towards components of the multidimensional political system that are unrelated to fraud information as coined by the Easton's concept of diffuse support.

2.3 Electoral Crimes and Punishment

In this second part of our theoretical scrutiny, we calibrate our theoretical expectation and outline how spillover effects might be moderated if third-party system actors become active as a response to fraud allegations. While accounting for the reactions of political actors has—to the best of our knowledge—not been incorporated into any study on election fraud information so far, it is at the same time crucial for understanding attitudinal dynamics stemming from exposing cheating, as fraud allegations are never observed in isolation but are accompanied by political developments

that are either permissive or marked by intervention. How does the spillover effect of election fraud information behave against interventions from within the political system?

The intervention that is most relevant to our argumentation line here is punishment. After information on malicious behavior has been exposed, it is not unusual that functionaries in the electoral commission that are responsible for electoral administration need to step down from office or forcefully lose their posts (El Espectador 2015; Idrobo, Kronick and Rodríguez 2022; Reuters 2016). Additionally, in recent decades, the judiciary has played an increasingly important role in electoral politics. Courts have emerged as an important actor that settles electoral disputes and frequently intervenes in pre- and post-electoral stages when the electoral conduct is in doubt (Eisenstadt 2002; Kerr and Wahman 2021). The topics that are covered by electoral tribunals range from issues revolving around constructing valid and comprehensive voter registers, the confirmation of candidate or party lists and the regulation of campaign resources up to sensitive issues such as election day fraud and vote manipulation. Court rulings on electoral crimes are highly salient for the electorate as they provide citizens with key non-partisan political information which regularly makes headlines in federal newspapers. Therefore, punishments are likely to directly affect the dynamics of attitude extrapolation. We focus on two specific arguments: the *amplifying spillover* argument and the *spillover suppression* argument.

2.3.1 The Spillover Suppression Effect

The line of reasoning emphasizing the suppression potential of punishments builds on the idea that functioning punishment mechanisms within the electoral commission or interventions of the judiciary into the electoral process signals information about the quality and independence of the underlying political system. Electoral commission punishment or court rulings may be interpreted as a sign of autonomy and professionalism which goes in counter to information about electoral fraud signaling system deficiencies. Punishments may lead to individual perceptions that the system of checks and balances in the country works reasonably well and that the political system does indeed have the capacity for self-correction if elections fail to meet shared standards. In this line of argumentation, successful punishments show that it's not the political system *as a whole* that is foul, but that state institutions do have the capacity to offer counterweights to malpractice. As a

consequence, interventions by electoral commissions or courts may reduce the spillover to decays in diffuse support.

2.3.2 The Spillover Amplification Effect

On the other hand, there is reason to believe that information on election fraud can lead to an amplification of spillovers. This argumentation line is rooted in the empirical observation that electoral quality is so routinely disputed in new democracies and authoritarian regimes that, opposition parties' or the international community's protests may simply be perceived as a conventional part of the game (Kerr and Wahman 2020). From this perspective, defeated candidates are incentivized to publicly condemn the electoral process in order to avoid seeming weak in front of their voter base and to discredit the authority of the political opponent (Lindberg 2006). The potential spillover effects from acquiring information about electoral misconduct may hence be depressed by doubts whether the allegation itself is credible. When alleged perpetrators of electoral crimes are subject to punishment, the presence of real convictions in turn provide an official recognition that the election process was not free and fair and send credible signals about the trustworthiness of fraud claims. Under this logic, punishments provide individuals with detailed information about the nature and scope of electoral malpractice and may serve as a heuristic device for them to reliably evaluate electoral fairness based on the statement of third-party actors. As such, punishments can be expected to lead to a *stronger* spillover effect, as they confirm the deficiencies in the political system as suggested by information on the presence of electoral manipulation.

2.3.3 Heterogeneous Effects for Government Supporters and Opponents

While in principle, both argumentation lines can be put forward, we can expect the interplay between information on election fraud and punishment effects to vary across *supporters* and *opponents* of the government or political regime. We argue that heterogeneous punishment effects across supporters and opponents might be routed in the logic of 'Bayesian belief updating' (Bullock 2009; Hill 2017). As showcased by a wide branch of research on the winner-loser gap in political support (Nadeau, Daoust and Dassonneville 2021; Cantú and García-Ponce 2015) and literature on the opposition in authoritarian regimes (Reuter and Szakonyi 2019), government opponents do *a priori* take on considerably more negative attitudes towards the political process than supporters.

Regime supporters have *ex ante* beliefs that are considerably more in line with a well-functioning political system than regime opponents. The sources of this imbalance can be manifold.

For one, they can be a manifestation of regime supporters and opponents selectively exposing themselves to different kinds of news. In authoritarian states and developing democracies, it's safe to assume that regime supporters are considerably more exposed to pro-regime propaganda or state-owned media outlets that particularly present the government in a favourable light. These arguments relate to differences in information *acquisition* that supporters and opponents self-select into.

Additionally, the way that both groups *process* the same kind of information might lead to differences in *ex ante* beliefs about the political system. Even if regime supporters have been exposed to fraud information in the past, it is likely that these are simply discounted as anti-government agitation. Reuter and Szakonyi (2019) show that when revealing information about systematic interference, especially *regime supporters* withdraw support from regime candidates that allegedly engaged in fraud as it these respondents for which the information actually makes a difference. Opponents, on the other hand, already hold *ex ante* beliefs that elections are tainted and have already incorporated expectations about election fraud into their pre-existing belief *before* being exposed to new information about electoral manipulation. As a consequence, government opponents likely don't need to perceive official interventions to be convinced that a certain fraud allegation is credible. Rather, interventions might challenge their pre-existing belief of a foul system and mitigate—or even remove—attitudinal spillovers as a consequence of fraud.

When tracing the impact of punishment, the following hypotheses hence guide our empirical scrutiny:

> *Hypothesis 2a: The attitudinal spillover effect of election fraud information is stronger for regime supporters when they are exposed to information about within-system interventions.*
>
> *Hypothesis 2b: The attitudinal spillover effect of election fraud information is weaker for regime opponents when they are exposed to information about within-system interventions.*

3   Matching Estimates from Cross-Sectional Survey Data

Before we describe our survey experiment, we first present a placebo test using various algorithms for statistical matching and cross-sectional survey data to motivate our experimental design. As we outline in Section 2.2, much of the previous literature on the interrelation between perceptions of electoral integrity and institutional trust has exploited (cross-national) survey data to identify the dimensions of political trust for which perceptions of electoral integrity matter, relating to our *Hypothesis 1* (for instance Norris 2014, 2019). We first showcase that even when applying a range of state-of-the-art matching algorithms that balance covariates across compared groups, we cannot distinguish spillover effects on political institutions (that are dictated by theory) from spillovers on non-political institutions (that should not be present in theory). Afterwards, we turn to our survey experiment.

To showcase how observational survey data is unsuitable to study the type of questions that we pose here and to motivate our experiment, we turn to data from the World Values Survey (Haerpfer et al. 2022). Because perceptions of electoral fraud are not randomly assigned among respondents, individuals may differ from each other in ways that are related to their fraud perception as well as their diffuse support for political institutions. The World Values Survey includes a rich set of covariates that make it possible to condition on possible differences between individuals and construct balanced (sub-)samples. While survey items relating to Easton's (1965, 1975) concept of diffuse support are part of the core questionnaire and asked consistently throughout all waves, a comprehensive battery of questions assessing respondents' perceptions of their country's electoral integrity has not been introduced before Wave 6 (2010-2014, c.f. Norris 2014). In particular, Wave 7 (2017-2020) of the cumulative data file provides questions on issues that are relevant as confounders of the relationship between fraud perceptions and diffuse support. We hence exploit data from Wave 7 of the World Value Survey covering 48 different countries across democratic and electoral authoritarian regimes. Table A1 in the Appendix provides an overview of our main dependent and independent variables as well as all covariates used for sample adjustment.

Regarding the measurement of the core variables, we use the classic measure of political trust grasping confidence in different institutions on a four-point scale as our dependent variables. The fraud information treatment is asking whether the respondent believes ballots in the country to be

counted free and fairly and measured on a four-point scale, which we dichotomize for the purposes of matching, with zero being more negative perceptions of election integrity. Due to the nature of dependent variable, we present results from ordered logit models.

We do not provide a detailed discussion of the different matching algorithms that underly our analysis. The interested reader is referred to Morgan and Winship (2007) for a comprehensive overview of the statistical underpinnings. We employ three algorithms: direct exact matching; a less restrictive coarsened exact matching (King and Nielsen 2019); and the widely-used propensity score-based nearest neighbor matching (without replacement). Exact and coarsened exact matching balance observations across our multivariate vector of covariates and drastically reduce sample size to balanced datasets of of $n = 580$ (exact matching) and $n = 2,475$ (coarsened exact matching). Propensity-score models match on a pre-determined balance score and keep our sample size at $n = 42,246$.[12] For exact and coarsened exact matching, we first create balanced datasets, and then calculate ATEs using Bayesian ordered logit models using non-informative priors and $1,000$ posterior samples. Following Alvarez and Levin (2021), for our propensity-score based models, we employ Bayesian estimation in our first-stage model in order to account for uncertainty in the balance score arriving at a posterior distribution of $1,000$ propensity scores. Results presented from propensity-score models are subsequently based on $1,000$ separate matching procedures for these $1,000$ different propensity scores per individual, again yielding a posterior distribution for the ATE.

First, we estimate the spillover effect of election integrity perceptions on political institutions that are unrelated to electoral administration. Second, we perform a placebo test by estimating the same effects on dimensions of trust that cannot be expected to be subject to attitudinal spillover. If the spillover effect of election integrity perceptions is well-identified, we expect to observe effects on *political institutions* while treatment and control units necessarily are similar in their attitudes towards *non-political institutions* that do not form part of the domestic political system.

---

[1]Figures A1 and A2 in the Appendix present measures of covariate balance between those individuals falling into the fraud and no fraud perception condition before and after our sample adjustment. After matching, differences in covariates between 'treatment' and 'control groups' are almost completely removed. We obtain a very high degree of balance on all covariates.

[2]One issue with our procedure is that the attitudinal covariates are not strictly preceding the main variables of interest and might thus be affected by fraud perceptions and confidence in institutions rather than predicting 'treatment' status. We report estimates including the attitudinal measures, but note that the results are robust to model specifications which only use socio-demographic information for covariate adjustment.

Figure 1: Average treatment effects of election integrity perceptions on trust in political (upper panel) and non-political (lower-panel) institutions from matching analyses on WVS Wave 7 (2017-2020) data. Individual figures report the difference in the predicted probability to tick each of the four categories of the dependent variable between respondents who perceive ballot counting to be foul vs. fair. Respondents with negative perceptions of electoral fairness show less confidence in political and non-political institutions. Point estimates are means, with 95% credibility intervals depicted with point-ranges.

Figure 1 visualizes the average treatment effect of the variable of interest, fraud perceptions, on political trust from the ordered logit regressions across the three matched datasets for political and non-political institutions. Perceived prevalence of election day ballot fraud in the federal elections of one's country shows to robustly decrease confidence in political institutions that are unrelated to electoral administration. Yet, we also observe similar effects of fraud information for non-political institutions such as the World Bank or the WTO, which are not dictated by theory and suggest that the effect of fraud perceptions is not causally identified. These results demonstrate the unsuitability of cross-sectional survey data for the analysis of the type of questions that we pose and motivate our experimental design that we present in the following.

14

4 Survey Experiment Design

We now test our arguments on (i) the attitudinal spillover of election fraud information and (ii) the moderating role of punishments using data from a pre-registered[3] online survey experiment that was conducted in Colombia, Mexico, and Russia in April-June of 2021. We first discuss the case selection for the experimental part of our study, describe our sample and provide an overview of the questionnaire. Lastly, we elaborate on the the causal identification strategy.

4.1 Case Selection

We focus on Colombia, Mexico, and Russia as these countries, on the one hand, share a variety of features relevant for our experimental design. In particular, we study middle-income countries with party systems that have shown a sizable degree of stability throughout past decades. These countries also share a large history of public controversies around electoral fraud. On the other hand, these countries are sufficiently different from one another to observe if the mechanisms that we outline travel to diverse political and cultural contexts. Additionally, the three countries provide a temporally comparable political environment for the study of electoral misconduct, as Mexico (July 2021) and Russia (September 2021) held federal elections later in 2021 and Colombia in May 2022. More details on how the three countries compare in terms of their electoral history of fair and foul elections are discussed in the Appendix.

4.2 Survey Design and Implementation

The respondents from all three countries were recruited using the crowd-sourcing platform *Yandex.Toloka*.[4] Evidently, not all societal groups are (equally) represented on crowd-sourcing platforms (e.g., Bartneck et al. 2015; Berinsky, Huber and Lenz 2012). As a consequence, our survey was predominantly conducted among a population of urban internet users who are somewhat younger than the general population and have obtained some level of higher education[5]. Attitudes towards

---

[3]The experimental design and all hypotheses have been pre-registered using the Open Science Framework. The pre-registration plan can be accessed via `https://osf.io/jyc2n/`.

[4]As stated in the pre-analysis plan, we initially intended to use *Amazon Mechanical Turk* for recruiting participants from the two Latin American countries. As the number of available *MTurk* workers from Mexico and Colombia turned out to be far from sufficient, we deviated from the pre-analysis plan and collected all participants from *Yandex.Toloka*. All data that has been genereted through *Amazon Mechanical Turk* is not reported nor included in the analysis.

[5]Tables A2 and A3 in the Appendix report summary statistics on the socio-demographic profile of survey respondents.

incumbents and political authorities are divided enough among this group (Robertson 2017) for us to be able to test both hypotheses. Most importantly, the socio-demographic profile of our survey respondents specifically targets those population groups that are particularly important for political dynamics such as gathering and sharing sensitive regime information and boosting their publicity by carrying them to the streets in protests. This makes the adopted sampling strategy advantageous over nationally representative surveys for this particular study.

In addition, while the specific attitudes of these surveyed groups might not be representative of the population as a whole, there is a wealth of evidence amassing that the factors which shape these attitudes are. Research has shown that treatment effects within attitudinal research based on data collected using crowd-sourcing platforms is similar to those found in representative surveys (Clifford, Jewell and Waggoner 2015; Coppock 2019).[6] Hence, while we expect that our sampled group differs from the general population in terms of their socio-demographic profile descriptively, it is fair to assume that the general patterns around their reactions to the experimental stimuli hold in the general population.

The survey was designed to take about ten minutes to complete and participants were presented with the survey in either Russian or Spanish (Colombia: $n = 517$, Mexico: $n = 481$, Russia: $n = 1,334$). When constructing our questionnaire, we mimicked all question formulations from the World Values Survey Wave 7 core questionnaire as closely as possible. For all original questions, questionnaire text was validated through cognitive interviews with native speakers.

4.3 Causal Identification

After answering a range of introductory questions on political attitudes, political knowledge, party affiliation, and general trust, respondents were randomly assigned to one of four experimental conditions, i.e. they were required to read a paragraph of text related to a hypothetical election (see Table 1). The text of the first group is neutral and states some basic facts about a hypothetically upcoming election to elect the country's national legislative body, and the respondents are presented with a status quo outcome. The second text contains general information about the election scenario identical to that of the first group, but additionally explicitly exposes respondents to information

---

[6]Similar conclusions have been reached independently from each other across a variety of disciplines (Bartneck et al. 2015; Yang et al. 2015).

about malpractices that were allegedly performed on election day.[7] The information presented to the third and fourth group additionally exposes respondents to information about punishments from within the political system directed at those individuals who are allegedly responsible for electoral crimes. For the third group, the punishment is limited to actions being undertaken by the superior electoral commission, i.e. the alleged perpetrators lose their positions at the commission. For the fourth group, the punishment includes legal actions in which alleged perpetrators were legally convicted for the performance of electoral crimes as well as personal consequences within electoral commissions for which the alleged perpetrators have worked.

In order to ensure that exactly one fourth of all respondents per country are placed in each of the four experimental groups and to avoid sparse data problems that might arise from extreme scenarios under complete randomization, we apply a randomized block design separately for each of the countries. Additionally, to be able to compare the results across regime supporters and opponents, we randomize treatments within these two groups.

After being confronted with the treatment conditions, respondents are presented a battery of questions measuring their levels of political trust in individual institutions of their political system. Specifically, we ask each participant the following question: *Upon receiving this information, how much confidence would you have in the following organizations or institutions?*. The battery of items includes *(1) the armed forces, (2) the police, (3) the justice system/courts, (4) the central electoral commission, (5) the government, (6) the parliament and (7) political parties"* in randomized order for each participant. Responses are collected on a four-point scale ranging from none at all to a great deal. Similarly, we include the same list of randomly-ordered non-political institutions to validate the identification of spillover effects as a placebo test as in Section 3.

A range of important characteristics of the experimental design can be noted in relation to the spillover theory outlined above. First, the text fragments that are presented to members of the treatment groups explicitly discuss the *mechanisms* of alleged election-day fraud. This should allow us to avoid variations in the interpretations of our fraud treatment, for instance associations of vote buying in Mexico (Cantú 2019*a*) and ballot box stuffing in Russia (Klimek et al. 2012, and hence prevent heterogeneous effects related to vagueness in the treatment. Second, referring to

---

[7]We follow Reuter and Szakonyi (2019) and omit the information source to avoid convoluting the effects of *information* and *source credibility*.

Table 1: Experimental Treatments.

---

(1) Control group: Neutral information
On Sunday, *[6 June 2021/19 September 2021/13 March 2022]*, legislative elections are scheduled to be held in *[Mexico/Russia/Colombia]*. More than [2,000/2,000/1,000] registered candidates will compete for the *[500/450/280]* parliamentary seats of the *[Champer of Deputies/State Duma/Congress of Colombia]*. The results will be determined by nearly *[90/110/36]* million *[Mexicans/Russians/Colombians]*. Imagine that the elections have already passed and suppose that as in the current convocation, *eight/four/twenty* parties retained seats in the assembly.

(2) Treatment group: Fraud information
[Neutral information]
In this hypothetical scenario, suppose that after election day, it becomes known that ballot-box stuffing and alterations of vote tallies in favor of the incumbent party perpetrated by individuals working for electoral commissions were widespread. Suppose that these electoral misconducts and manipulation practices took place across several regions of the country.

(3) Treatment group: Fraud information with electoral commission punishment
[Neutral information]
[Fraud information]
Furthermore, suppose that as a consequence, individuals allegedly responsible for fraud lost their position in the electoral commissions that they served in.

(4) Treatment group: Fraud information with court and electoral commission punishment
[Neutral information]
[Fraud information]
Furthermore, suppose that as a consequence, individuals allegedly responsible for fraud lost their position in the electoral commissions that they served in. Also, legal action was brought against these individuals, who were convicted for electoral crimes by responsible courts.

'individuals working for electoral commissions', we explicitly state the alleged *perpetrator of fraud*. Mentioning the perpetrator is crucial as it allows us to directly study if trust is extrapolated to different bodies that are unrelated to the administration of elections.

Importantly, not all of the institutions included in our battery on political trust allow us to unambiguously identify spillover effects. For example, it is hard to disentangle whether fraud in new democracies or electoral authoritarian regimes constitutes the actions of micro-level agents or whether these practices are instructed from party and/or state representatives. For citizens who expect the latter, information about election day fraud may actually function as informational cues implying partisan involvement from political authorities. Among these participants, changes

in diffuse support for the institutions of the government, parliament, political parties and central electoral commission may not necessarily be the result of an attitudinal spillover following from trust extrapolation as we theorize.Instead, it can be a straightforward withdrawal of political support from the perceived perpetrator of electoral misconduct.

To counter this possibility, we ask for respondents' levels of diffuse support towards a number of institutions that are clearly exogenous to our fraud treatment. Without spillover, trust in institutions such as the armed forces, the police, or the justice system are unaffected by information on election day manipulation, as members of these institutions are unrelated to fraudulent interference practiced by individuals working for electoral commissions on voting day. In contrast, decays in diffuse support towards these institutions as a response to the fraud treatments present genuine spillover effects.

5   Experimental Results

5.1   Statistical Modeling and Estimation

To test Hypothesis 1 (the effect of exposure to the information about electoral fraud on confidence in institutions), we estimate the mean difference between the control group (1) and treatment group (2) in Table 1. To evaluate Hypotheses 2a and 2b (heterogeneous effects of punishment across supporters and opponents), we estimate the differences between group (2), which only received fraud information, and either group (3) or (4), who received punishment information on top of fraud information. To estimate varying effects across government supporters and opponents, we construct a binary moderator based on our pre-treatment measure of party affiliation, distinguishing those respondents that consider themselves supporters of a party that was in government or the opposition at the time of the survey.

Since our dependent variables comprise four ordered categories j, with $(j = 1, ..., 4)$, we again estimate a set of ordered logistic regressions with the following specifications:

$$ln\left(\frac{\Pr(y_i \leq j)}{\Pr(y_i > j)}\right) = \underbrace{\underbrace{\alpha_j - (\beta_1 \text{ Control}_i + \beta_2 \text{ Punishment}_i + \beta_3 \text{ Judicial Punishment}_i}_{\text{Main Specification (H1)}}) \times \beta_4 \text{ Opponent}_i}_{\text{Heterogeneous Effects Specification (H2)}}, \quad (1)$$

19

where y_i is the level of diffuse support of an individual i with $(i = 1, ..., n)$ for an institution, and Control, Punishment, and Judicial Punishment are binary indicators for membership in the experimental groups (1), (3), and (4) [8]. As outlined in the pre-registration plan, we pool the available data across countries and use all available observations that fulfill basic data-quality criteria for the main analysis.[9][10]

For parameter estimation, we employ a fully Bayesian framework for statistical inference as implemented in Stan (Stan Development Team 2020). Specifically, we rely on Hamiltonian Monte Carlo sampling in which priors are defined to follow Student-t distributions centered around zero and take on a sufficiently large variance to ensure that prior distributions are uninformative and do not favor any of the substantial hypotheses. We run a set of four Markov chains out of which we discard the first 10,000 as warm-up and use the following 10,000 samples to describe posterior distributions, which results in a total of 40,000 post-warmup samples. We check for model convergence using the Gelman-Rubin diagnostic and consider models converged if the discrepancy measure stays below 1.1 (Gelman et al. 2004).

5.2 Empirical Results

We start off my describing evidence for the spillover effect of fraud information as stated by Hypothesis 1 (Figure 2), validate it against a placebo test on non-political institutions (Figure 3), and trace spillover dynamics across countries (Figure A4). Secondly, we evaluate Hypotheses 2a and 2b (Figure 4). For each figure, we first calculate individuals' probabilities of selecting each of the four answer categories, from *A great deal* to *None at all*. In the figures themselves, we present the differences in probabilities for each answer category between two respective experimental conditions.

---

[8]Individuals who only received fraud information serve as the reference category in our analysis.

[9]Respondents are excluded if they completed the survey in less than three minutes.

[10]As our theory—in which attitudes towards elections *spill over* to attitudes of other political institutions—implies that it is the changes in trust in elections that are responsible for the differences in trust for other institutions of the political system, one might trace empirical effects using mediation models. In our questionnaire, we have included a question that accounts for trust in the electoral system. Figure A6 in the Appendix re-presents our main results explicitly incorporating trust in elections as a mediator, adjusting the model formulation in Equation (1). As detailed in the Appendix, all substantive conclusions are unchanged and mediation effects behave as hypothesized.

5.2.1 Spillover Effect of Fraud Information

Figure 2 summarizes evidence for the spillover effect of fraud information as stated by Hypothesis 1, comparing trust in political institutions between respondents in experimental groups (2) and (1). As we can see, respondents who received fraud information are substantially less likely to
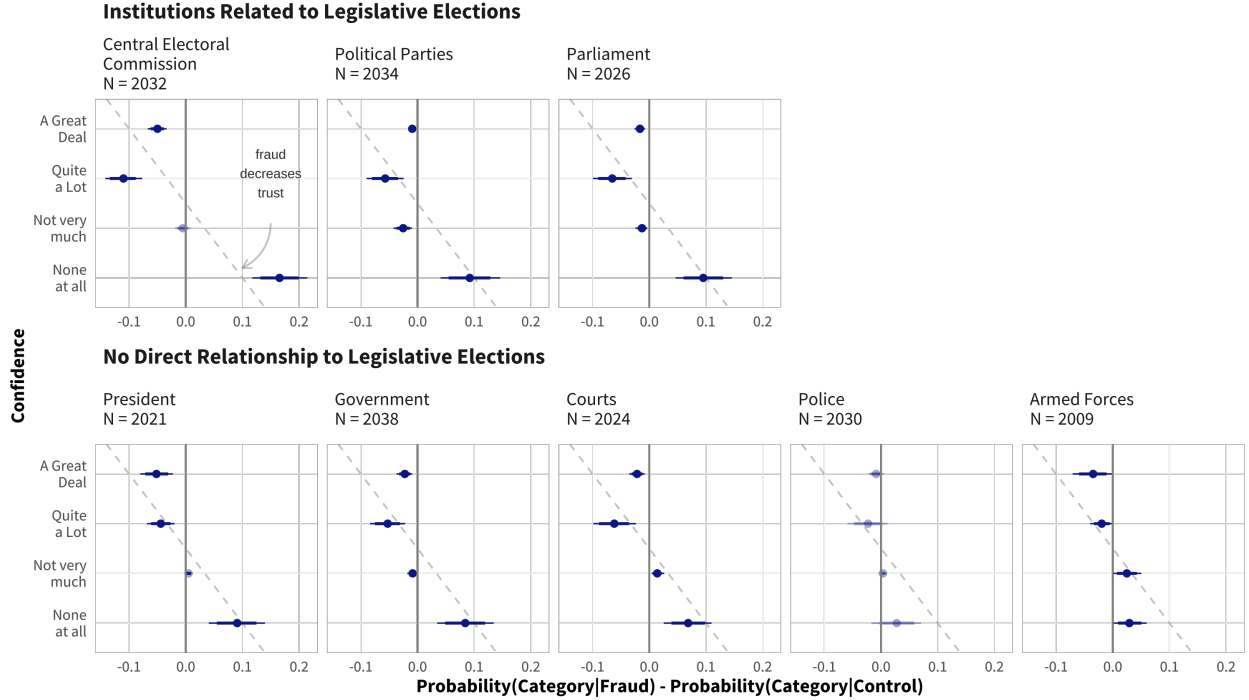
Figure 2: The Effects of Exposure to Fraud Information on Confidence in Political Institutions. Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.

voice higher degrees of confidence in political institutions than respondents of the control group. Naturally, this effect is most evident for trust in the Central Electoral Commission, whose affiliates have performed vote rigging as in our experimental treatment. However, these decreasing levels of political trust hold equally for institutions that are directly related to legislative elections and hence and might be perceived as endogenous to the fraud treatment (political parties, parliament) as well as political institutions that are not related to the electoral event and can be treated as exogenous. With the exception of the police, information about election day electoral produces

decays in political trust that spill over to political institutions which are unrelated to electoral events, providing robust empirical evidence for Hypothesis 1. Regarding the non-existing effect on the institution of the police, we would expect an average respondent to have little direct contact with most political institutions we study, yet one could argue that the police, is, potentially, the
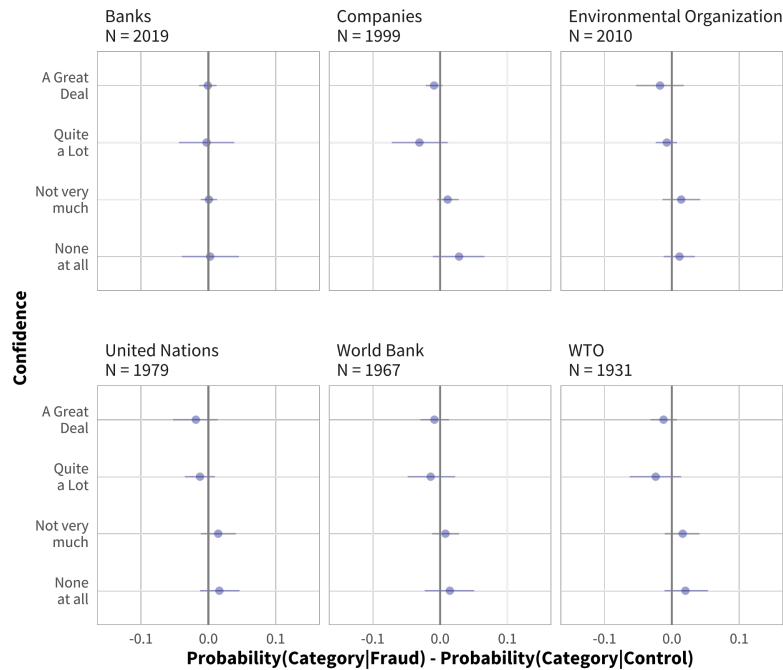
Figure 3: The Effects of Exposure to Fraud Information on Confidence in Non-political Institutions. Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.

institution with which respondents are more likely to interact in their daily life. For that reason, trust in police may be naturally stronger and more insensitive than trust in other institutions, making the spillover for the police less pronounced.

Figure 3 repeats the placebo test that we reported on in Section 3 and provides further evidence that the spillover effects of election fraud information are well identified—respondents of the two different treatment conditions do not differ in their trust in non-political institutions that are

unrelated to our outlined theory, going against the notion that the reported effects might be a methodological artefact of the survey design.

After having described main effects, we also investigate whether experimental results differ across countries. As Figure A4 shows, while *some* hypothesized effects do exist for Colombia and Russia (for the institutions of political parties, the parliament and the president), main effects in the pooled sample are mostly driven by the Mexican subsample, which shows strong effects across seven out of the eight analyzed institutions. First, we do note that there is evidence suggesting that Mexican respondents simply did a better job in grasping and revising the treatment text. If we restrict the analysis to those respondents that correctly summarized their treatment text in the follow-up attention check[11], spillover effects are present for a range of institutions in all three countries (see Figure A3.) More substantively, these country-level differences could be attributed to the straightforward mechanism of Bayesian belief updating which is restricted if prior expectations are already in line with the treatment. On the one hand, if baseline trust in political institutions (as captured in the control group in our experiment) could simply be very low, allowing very little room for any belief updating as a consequence to fraud. In other words, if there exists little trust in political institutions in the first place, information on fraudulent elections can hardly show any negative effects as little belief updating is taking place. This is likely to explain the fewer number of spillover effects in Colombia and Russia. Colombia shows the lowest levels of trust across all eights institutions *across all respondents* and *within the control group*, with government opponents making up 82% of our sample (compared to 65% in Mexico and 59% in Russia). In Russia, it is primarily regime supporters that show to be sensitive to treatment information (see Figure A5) while few effects are found for opponents who are already sceptical of the regime, likewise speaking in favor of a mechanism of Bayesian belief updating that acquires higher baseline levels of trust in general.[12]

---

[11]We manually coded respondents' summaries into correctly or incorrectly the scenario presented to them.

[12]Specifically for the case of authoritarian Russia and past cases of severely restricting opposition candidates from running in the election, a second explanation lies in the possibility of the control condition which re-states the status-quo not being interpreted as 'neutral' by regime opponents. Rather, an election in which the status quo persists might re-inforce their beliefs of an authoritarian political system, again not allowing for much updating after being confronted with additional information on electoral malpractice.

Needless to say, across all examined scenarios, it is clear that information on electoral malpractice robustly leads to decays in political trust for more than one political institution that is not tied to electoral administration.

5.2.2 System-response Effect

Regarding the moderating role of political interventions as a response to the circulation of fraud information, Hypotheses 2a and 2b expected the amplification of spillovers in comparison to the fraud
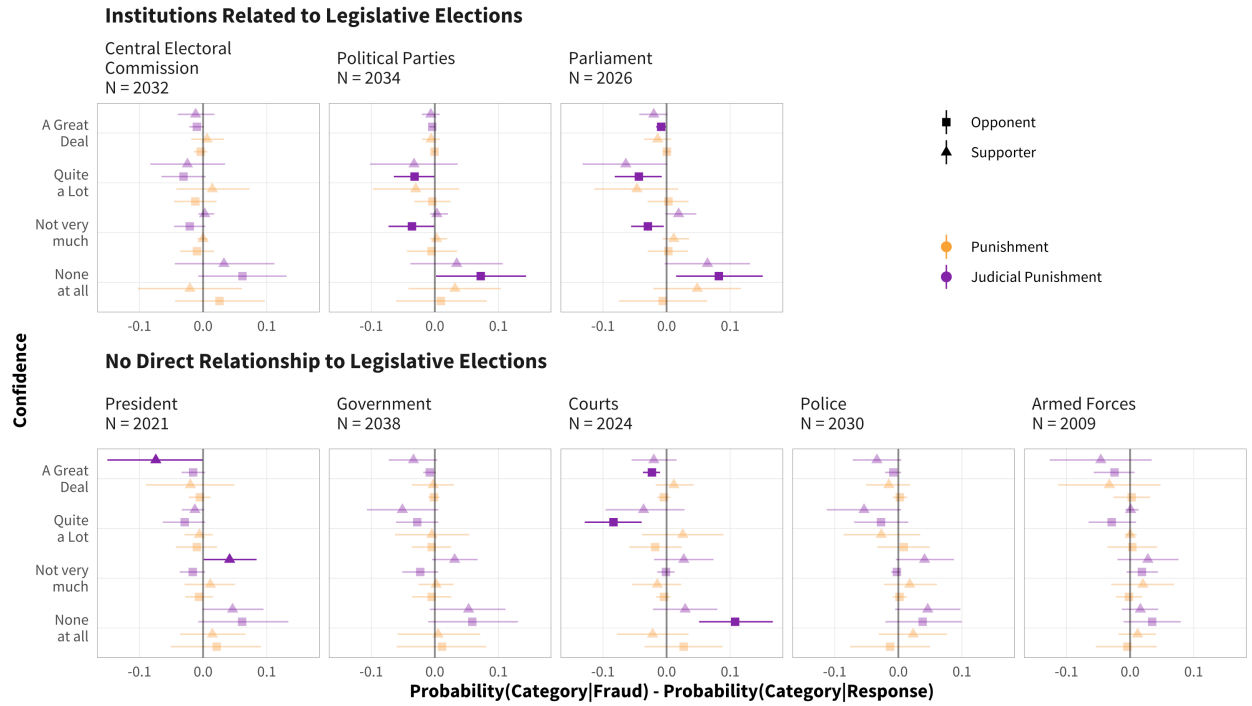
Figure 4: The Effects of Exposure to Perpetrators' Judicial Punishment Information on Confidence in Political Institutions. Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.

condition (2) for government supporters and a suppression of spillovers in comparison to the fraud condition (2) for government opponents. Figure 4 plots differences in predicted probabilities to tick any of the four trust categories between those that only received fraud information (experimental

group 2) and those that additionally received information on punishments (experimental groups 2 and 3). There are several key take aways. First, punishment information does predominantly not change spillover dynamics as for most institutions, punishment information does—on average—not alter levels of political trust in comparison to those who only received information on fraud. Second, if any, we observe evidence for spillover suppression across opponents and supporters given that experimental texts mention electoral commission *and* judicial punishment. Naturally, this effect is most evident for trust in courts, as respondents that received information on electoral commission and judicial punishment exert higher confidence in courts than those who merely perceived information on election fraud. Yet, for some unique comparisons, spillovers to other institutions are mitigated as well. These patterns persist once we allow for heterogeneous effects across countries, too, and when we restrict the sample to those individuals who clearly finished reading the treatment text.

To sum up, we find sufficient support for the first hypothesis, indicating the damaging effect of information about fraud on institutional trust across both samples. At the same time, the effect of within-system responses remains ambiguous. While adequate punishment of fraud perpetrators may decrease the negative effects of fraud information, this effect is far from universal and omnipresent across institutions.

6   Conclusion

While nation-wide multiparty electoral events are nowadays omnipresent across democracies and (electoral) autocracies alike, these are frequently accompanied by doubts about their integrity, outright accusations of fraud, and large-scale public protests. While literature on the citizen-system nexus has already amassed a wealth of knowledge on the behavioral and attitudinal consequences that becoming aware of electoral malpractice has for the citizenry, scholars have predominantly studied individuals' evaluations of the electoral process itself (Robertson 2017), the candidates that are supposedly involved in manipulation (Reuter and Szakonyi 2019) and the government that stemmed out of an allegedly fraudulent election (Williamson 2021). On the other hand, those contributions that have explicitly focused on how perceptions of election fraud influence individuals'

attitudes towards components of the political system that are not tied to electoral administration have relied on cross-sectional survey data (Norris 2014, 2019).

At the same time, widespread accusations of electoral fraud regularly induce political interventions by exactly those political elites that allegedly are the beneficiaries of manipulation such as the dismissal or judicial punishment of electoral staff (compare Reuters 2016). To our best knowledge, no contribution has so far focused on how credible within-system responses shape individuals' attitudinal responses to allegations of fraud.

In this article, we have first empirically showcased the limitations that cross-sectional survey data holds for the study of the attitudinal nexus between citizens and the political system. Even when applying a range of matching algorithms combined with a Bayesian estimation procedure, spillover effects on political institutions cannot be distinguished from placebo effects that should not be in place in theory, which casts serious doubts on the causal identification of attitudinal spillovers as a whole. Second, overcoming these limitations, we have presented results from a pre-registered online survey experiment conducted in Colombia, Mexico and Russia. Our experimental results document that disseminating credible information on election fraud does indeed induce decays in diffuse support for political institutions that are unrelated to electoral administration. Third, we showed that these attitudinal spillovers largely persist even if information on electoral manipulation is accompanied by information on credible punishments of micro-level agents of fraud.

Our article holds two main implications for developing democracies as well as contemporary authoritarianism. The first is that the negative effect of election fraud information is substantially more detrimental than currently acknowledged by the academic literature. This is because information on electoral misconduct even induces shifts in public support towards components of the political system that are no beneficiaries of manipulation and are not related to electoral administration. Second, our findings hold implications for the practice of election monitoring itself. Our article well aligns with a set of studies that have highlighted the cost among civil society when election observation missions expose cheating (Daxecker 2012). Our preliminary findings suggest that when large-scale observation missions that are perceived and framed as credible players in the field claim election malpractice to be at place, such exposure may have detrimental effects that may hinder, rather than foster, the consolidation of a democratic society. This is especially relevant against the backdrop of widespread criticism that has been voiced against recent election

observation missions proclaiming early conclusions about electoral malpractice that later do not uphold more intensive scrutiny (Idrobo, Kronick and Rodríguez 2022). While the spillover effect of fraud information is consistent, the mitigating effect of effective punishments is not. Third-party actors who monitor and evaluate the legitimacy of electoral events are hence advised to cautiously reflect negative assessments of electoral integrity before disseminating information.

References

Alvarez, Michael R. and Ines Levin. 2021. "Uncertain Neighbors: Bayesian Propensity Score Matching for Causal Inference." *Unpublished Manuscript* .

Bartneck, Christoph, Andreas Duenser, Elena Moltchanova and Karolina Zawieska. 2015. "Comparing the similarity of responses received from studies in Amazon's Mechanical Turk to studies conducted online and with direct recruitment." *PLoS ONE* 10(4):1–23.

Bejarano, Ana María and Eduardo Pizarro. 2005. From 'Restricted' to 'Besieged': The Changing Nature of the Limits to Democracy in Colombia. In *The Third Wave of Democratization in Latin America: Advances and Setbacks*, ed. Frances Hagopian and Scott P. Mainwaring. Cambridge: Cambridge University Press pp. 235–260.

Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon.com's mechanical turk." *Political Analysis* 20(3):351–368.

Bowler, Shaun and Jeffrey A. Karp. 2004. "Politicians, scandals, and trust in government." *Political Behavior* 26(3):271–287.

Bullock, John G. 2009. "Partisan bias and the bayesian ideal in the study of public opinion." *Journal of Politics* 71(3):1109–1124.

Bush, Sarah Sunn and Lauren Prather. 2018. "Who's There? Election observer identity and the local credibility of elections." *International Organization* 72(3):659–692.

Cantú, Francisco. 2014. "Identifying irregularities in Mexican local elections." *American Journal of Political Science* 58(4):936–951.

Cantú, Francisco. 2019*a*. "Groceries for votes: The electoral returns of vote buying." *Journal of Politics* 81(3):790–804.

Cantú, Francisco. 2019*b*. "The fingerprints of fraud: Evidence from Mexico's 1988 presidential election." *American Political Science Review* 113(3):710–726.

Cantú, Francisco and Omar García-Ponce. 2015. "Partisan losers' effects: Perceptions of electoral integrity in Mexico." *Electoral Studies* 39:1–14.

Citrin, Jack. 1974. "Comment: The Political Relevance of Trust in Government." *American Political Science Review* 68(3):973–988.

Clifford, Scott, Ryan M Jewell and Philip D Waggoner. 2015. "Are samples drawn from Mechanical Turk valid for research on political ideology?" *Research & Politics* 2(4):205316801562207.

Coppock, Alexander. 2019. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* 7(3):613–628.

Daxecker, Ursula E. 2012. "The cost of exposing cheating: International election monitoring, fraud, and post-election violence in Africa." *Journal of Peace Research* 49(4):503–516.

Dellmuth, Lisa Maria and Jonas Tallberg. 2015. "The social legitimacy of international organisations: Interest representation, institutional performance, and confidence extrapolation in the United Nations." *Review of International Studies* 41(3):451–475.

Duque Daza, Javier. 2019. "Democracia electoral fraudulenta. La trashumancia electoral como estrategia para ganar elecciones locales en Colombia." *Estudios Políticos (Medellín)* (55):61–86.

Easton, David. 1965. *A Systems Analysis of Political Life.* New York: Wiley.

Easton, David. 1975. "A Re-Assessment of the Concept of Political Support." *British Journal of Political Science* 5(4):435–457.

El Espectador. 2015. "A Dos Días de Elecciones, Sorprenden a Yahír Acuña Con \$200 Millones." *El Espectador, October 24* .

Enikolopov, Ruben, Vasily Korovkin, Maria Petrova, Konstantin Sonin and Alexei Zakharov. 2013. "Field experiment estimate of electoral fraud in Russian parliamentary elections." *PNAS* 110(2):448–452.

Frye, Timothy and Ekaterina Borisova. 2019. "Elections, Protest, and Trust in Government: A Natural Experiment from Russia." *The Journal of Politics* 81(3):820–832.

Geddes, Barbara and John Zaller. 1989. "Sources of Popular Support for Authoritarian Regimes." *American Journal of Political Science* 33(2):319–347.

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis.* Boca Raton: Chapman & Hall/CRC.

Greene, Kenneth F. 2007. *Why Dominant Parties Lose: Mexico's Democratization in Comparative Perspective.* Cambridge: Cambridge University Press.

Greene, Kenneth F. and Mariano Sánchez-Talanquer. 2018. Authoritarian legacies and party system stability in Mexico. In *Party Systems in Latin America: Institutionalization, Decay, and Collapse*, ed. Scott Mainwaring. Cambridge: Cambridge University Press pp. 201–226.

Haerpfer, C., R. Inglehart, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin and B. Puranen. 2022. *World Values Survey: Round Seven - Country-Pooled Datafile Version 3.0.* Madrid, Spain and Vienna, Austria: JD Systems Institute and WVSA Secretariat.

Hill, Seth J. 2017. "Learning together slowly: Bayesian learning about political facts." *Journal of Politics* 79(4):1403–1418.

Hiskey, Jonathan T. and Shaun Bowler. 2005. "Local context and democratization in mexico." *American Journal of Political Science* 49(1):57–71.

Hooghe, Marc and Sofie Marien. 2012. "A comparative analysis of the relation between political trust and forms of political participation in Europe." *European Societies* 15(1):131–152.

Hyde, Susan D. 2011. "Catch Us If You Can: Election Monitoring and International Norm Diffusion." *American Journal of Political Science* 55(2):356–369.

Idrobo, Nicolás, Dorothy Kronick and Francisco Rodríguez. 2022. "Do Shifts in Late-Counted Votes Signal Fraud? Evidence From Bolivia." *Journal of Politics* 84(4):Online Only.

Inglehart, R., C. Haerpfer, A.. Moreno, C. Welzel, K. Kizilova, J. Diez-Moreno, M. Lagos, P. Norris and E. & Puranen B. et al. Ponarin. 2020. "World Values Survey: Round Seven - Country-Pooled Datafile 2017-2020." *Madrid: JD Systems Institute* .

Jimenez, Raúl, Manuel Hidalgo and Peter Klimek. 2017. "Testing for voter rigging in small polling stations." *Science Advances* 3(6):1–8.

Kelley, Judith. 2012*a*. "The Good, the Bad, and the Ugly: Rethinking Election Monitoring." *International Institute for Democracy and Electoral Assistance* .

Kelley, Judith G. 2012*b*. *Monitoring Democracy.* Princeton and Oxford: Princeton University Press.

Kerr, Nicholas and Michael Wahman. 2020. "Electoral Rulings and Public Trust in African Courts and Elections." *Comparative Politics* .

King, Gary and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27(4):435–454.

Klimek, Peter, Yuri Yegorov, Rudolf Hanel and Stefan Thurner. 2012. "Statistical Detection of Systematic Election Irregularities." *PNAS* 109(41):16469–16473.

Kobak, Dmitry, Sergey Shpilkin and Maxim S. Pshenichnikov. 2016*a*. "Integer percentages as electoral falsification fingerprints." *Annals of Applied Statistics* 10(1):54–73.

Kobak, Dmitry, Sergey Shpilkin and Maxim S. Pshenichnikov. 2016*b*. "Statistical fingerprints of electoral fraud?" *Significance* 13(4):20–23.

Kobak, Dmitry, Sergey Shpilkin and Maxim S. Pshenichnikov. 2018. "Putin's peaks: Russian election data revisited." *Significance* 15(3):8–9.

Lankina, Tomila and Katerina Tertytchnaya. 2020. "Protest in electoral autocracies: a new dataset." *Post-Soviet Affairs* 36(1):20–36.
URL: *https://www.tandfonline.com/doi/full/10.1080/1060586X.2019.1656039*

Lindberg, Staffan I. 2006. Tragic Protest: Why Do Opposition Parties Boycott Elections? In *Electoral Authoritarianism: The Dynamics of Unfree Competition*, ed. Andreas Schedler. Lynne Rienner Publishers.

Mauk, Marlene. 2019. "When does cheating matter? Electoral malpractice, perceptions of electoral fairness, and political trust." *Working Paper* .

Mishler, William and Richard Rose. 2001. "What Are the Origins of Political Trust?: Testing Institutional and Cultural Theories in Post-communist Societies." *Comparative Political Studies* 34(1):30–62.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactual and Causal Inference— Methods and Principles for Social Research.* Cambridge: Cambridge University Press.

Myagkov, Mikhail, Peter C. Ordeshook and Dimitri Shakin. 2009. *The Forensics of Election Fraud: Russia and Ukraine.* Cambridge: Cambridge University Press.

Nadeau, Richard, Jean-François Daoust and Ruth Dassonneville. 2021. "Winning, Losing, and the Quality of Democracy." *Political Studies* p. Online Only.

Nisbett, Richard E. and Timothy DeCamp Wilson. 1977. "The Halo Effect: Evidence for Unconscious Alteration of Judgements." *Journal of Personality and Social Psychology* 35(4):250–256.

Norris, Pippa. 2014. *Why Electoral Integrity Matters.* Cambridge: Cambridge University Press.

Norris, Pippa. 2019. "Do perceptions of electoral malpractice undermine democratic satisfaction? The US in comparative perspective." *International Political Science Review* 40(1):5–22.

OAS. 2019. "Electoral Integrity Analysis: General Elections in the Plurinational State of Bolivia, October 20, 2019, Final Report." *Organization for American States (OAS), Secretariat for Strengthening Democracy (SSD), Department of Electoral Cooperation and Observation* .

Palmer, Carl L and Rolfe D Peterson. 2016. "Halo Effects and the Attractiveness Premium in Perceptions of Political Expertise." *American Politics Research* 44(2):353–382.

Przeworski, Adam, Susan C. Stokes and Bernard Manin. 1999. *Democracy, Accountability, and Representation.* Cambridge: Cambridge University Press.

Reuter, Ora John and David Szakonyi. 2019. "Electoral Manipulation and Regime Support: Survey Evidence from Russia." *Unpublished manuscript* .

Reuters. 2016. "Russian Local Election Officials Fired Over Vote Irregularities." *Reuters, November 2* .

Robertson, Graeme. 2017. "Political Orientation, Information and Perceptions of Election Fraud: Evidence from Russia." *British Journal of Political Science* 47(3):589–608.

Rozenas, Arturas. 2017. "Detecting Election Fraud from Irregularities in Vote-Share Distributions." *Political Analysis* 25(1):41–56.

Sartori, Giovanni. 1976. *Parties and Party Systems.* New York: Cambridge University Press.

Schneider, Irena. 2017. "Can We Trust Measures of Political Trust? Assessing Measurement Equivalence in Diverse Regime Types." *Social Indicators Research* 133(3):963–984.

Stan Development Team. 2020. "RStan: the R interface to Stan.".
URL: *http://mc-stan.org/*

Thorndike, E.L. 1920. "A Constant Error in Psychological Ratings." *Journal of Applied Psychology* 4(1):25–29.

Torcal, Mariano and Pablo Christmann. 2019. "Congruence, national context and trust in European institutions." *Journal of European Public Policy* 26(12):1779–1798.

van Ham, Carolien. 2015. "Getting elections right? Measuring electoral integrity." *Democratization* 22(4):714–737.

Williamson, Scott. 2021. "Elections, legitimacy, and compliance in authoritarian regimes: evidence from the Arab world." *Democratization* 0(0):1–22.

Yang, Hung Chia, Sally M. Donovan, Scott J. Young, Jeffery B. Greenblatt and Louis Benoit Desroches. 2015. "Assessment of household appliance surveys collected with Amazon Mechanical Turk." *Energy Efficiency* 8(6):1063–1075.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion.* Cambridge: Cambridge University Press.

Appendices

A. Cross-National Evidence: Matching Analysis

a. Descriptive Statistics

Table A1 contains the summary statistics for complete cases, World Values Survey Wave 7 (2017-2020).

Table A1: Summary Statistics of Key Variables, World Values Survey Wave 7 (2017-2020).

	N	Mean	St. Dev.	Min	25%	75%	Max
Confidence in Armed Forces	42,586	2.865	0.945	1	2	4	4
Confidence in the Police	42,586	2.629	0.943	1	2	3	4
Confidence in Courts	42,586	2.568	0.941	1	2	3	4
Confidence in Parliament	42,586	2.187	0.926	1	1	3	4
Confidence in Government	42,586	2.384	0.984	1	2	3	4
Confidence in Parties	42,586	2.034	0.874	1	1	3	4
Confidence in Companies	42,586	2.398	0.851	1	2	3	4
Confidence in UN	42,586	2.445	0.943	1	2	3	4
Confidence in Banks	42,586	2.563	0.930	1	2	3	4
Confidence in WTO	42,586	2.439	0.911	1	2	3	4
Confidence in World Bank	42,586	2.419	0.948	1	2	3	4
Fraud Perception	42,246	2.841	1.028	1	2	4	4
Fraud Perception (binary)	42,586	0.651	0.477	0	0	1	1
Political Interest	42,586	2.390	0.955	1	2	3	4
Generalized Trust	42,246	0.202	0.401	0	0	0	1
Female	42,246	0.495	0.500	0	0	1	1
Age	42,586	41.634	15.832	16	28	54	103
Family Savings	42,586	2.049	0.902	1	1	2	4
Rural	42,246	0.339	0.473	0	0	1	1
Perceived Political Corruption	42,586	7.736	2.423	1	6	10	10

- Confidence in political institutions is measured on a 1 to 4 scale, with 1 meaning *None at all* and 4 depicting *A great deal.*

- Political interest is measured on a 1 to 4 scale, with 1 meaning *Not at all interested* and 4 depicting *Very interested.*

- Generalized trust is a binary indicator, with 0 meaning *Need to be very careful* and 1 depicting *Most people can be trusted.*

- Fraud perception is measured in WVS on a 1 to 4 scale based on the agreement with the following statement: "How often in country's elections: Votes are counted fairly", with 1 meaning *Not at all often* and 4 depicting *Very often.*

- Fraud perception (binary) is the variable used in our matching analysis and is a binary indicator, with 0 being comprised of *Not at all often* and *Not often* and 1 of *Fairly often* and *Very often* answers to the respective WVS question.

- Family savings during past year is measured on a 1 to 4 scale, with 1 indicating Save money, 2—*Just get by*, 3—*Spent some savings and borrowed money*, and 4—*Spent savings and borrowed money.*

- Perceived Political Corruption is measured on a 1 to 10 scale, with 1 meaning *There is no corruption in my country* and 10 depicting *There is abundant corruption in my country.*

b. Covariate Balance

Figures present measures of covariate balance between those individuals falling into the 'fraud' and 'no fraud' perception condition before and after our sample adjustment. For each covariate that we match on, we report the standardized bias as measured by the difference in means between both groups of individuals scaled by the pooled standard deviation. Dots to the right (left) of the dashed vertical line are indicative of a higher incidence of respective characteristics among those individuals that perceive the elections of their country as unfair (fair). As indicated by the grey circles, fraud perceptions are most prevalent among those individuals that report to never turn out to vote, obtain low levels of interpersonal trust, and perceive political corruption at large scale in their country.
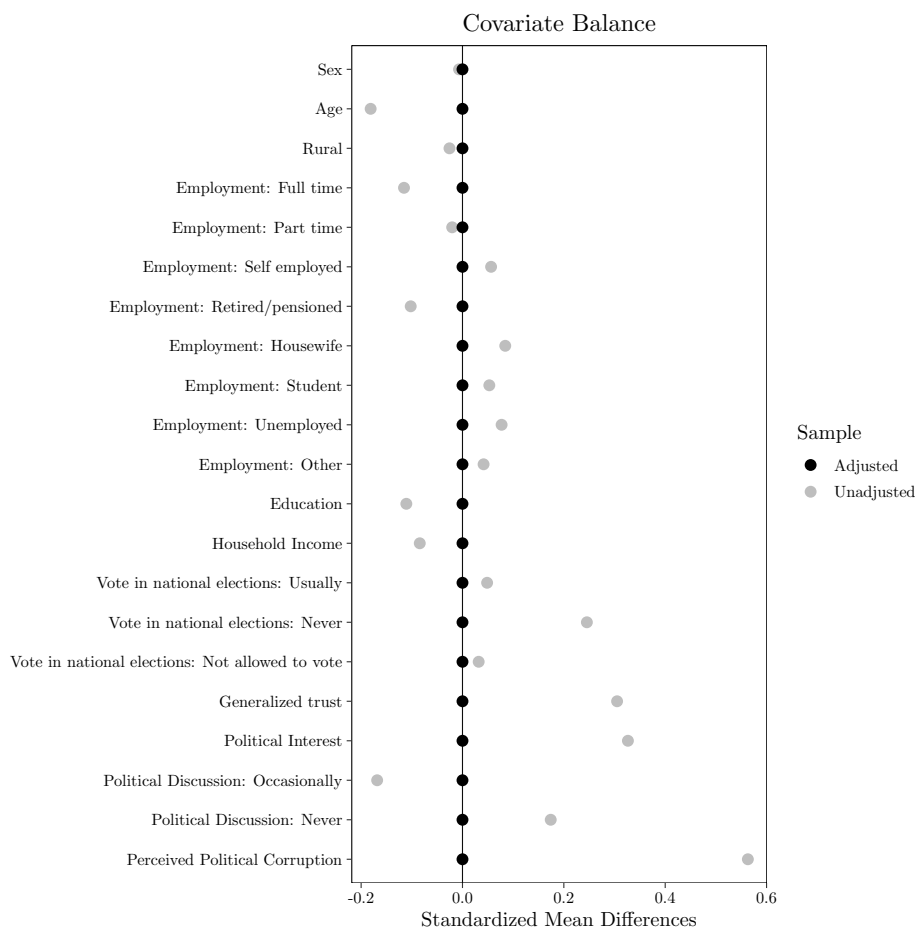
Figure A1: Covariate balance before and after sample adjustment using direct exact matching.
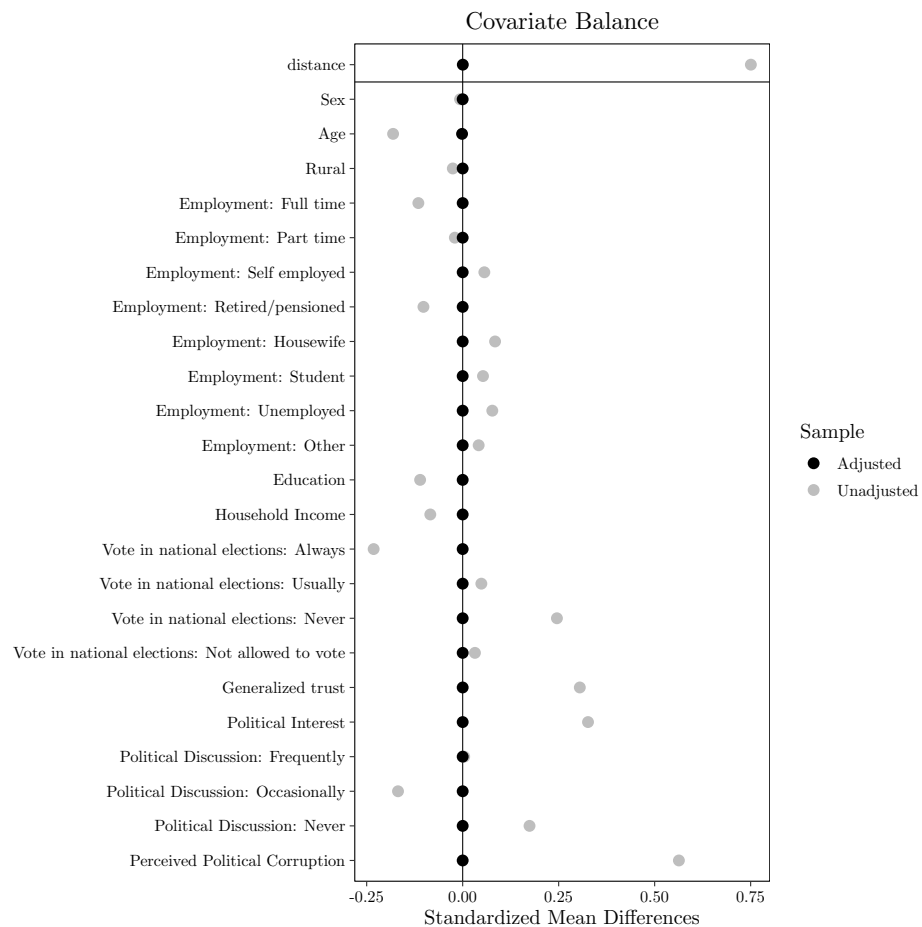
Figure A2: Covariate balance before and after sample adjustment using coarsened exact matching.

B. Details on Case Selection

In the post-World War II era, Mexico underwent a gradual and often described as pendular democratization process (Cantú and García-Ponce 2015; Hiskey and Bowler 2005). Since the end of the Mexican Revolution in 1917, elections were held regularly in six-year intervals and political opposition was granted passive voting rights. Yet the authoritarian rule of the Institutional Revolutionary Party (PRI) effectively consolidated a hegemonic one-party party system (Sartori 1976). Popular distrust in the legitimacy of Mexican elections roots in the experience of PRI's one-party rule, which was notoriously engaged in attempts of electoral manipulation against parties from both sides of the political spectrum. PRI's strategies in balancing out authoritarianism with competitive elections manifested in unrecognized victories of the right-wing National Action Party (PAN) in a multitude of subnational elections in the 1980s and 1990s (Greene 2007; Cantú and García-Ponce 2015), systematic repression against the left-winged Democratic Revolution Party's (PRD) candidates (Greene 2007), and election-day fraud such as the manual alteration of vote tallies in multiple regional and national-level contests (c.f. Cantú 2019*b*). It was not before the 1980s that electoral competition led to more inclusive electoral contests which produced changing majorities. First, recognized opposition victories occurred in state-level and local elections and only recently culminated in the first national-level contest since the Mexican Revolution of 1920 being decided in favor of the political challenger when the National Action Party's (PAN) candidate defeated PRI's Francisco Labastida in July 2000. Today, Mexico's political party system shows a remarkable level of institutionalization when compared to other new democracies and locates the country on the upper end on the scale of party system stability in Latin America (Greene and Sánchez-Talanquer 2018). Notably, the historical baggage of electoral maladministration and attempts of manipulation persists and reaches forward up until the country's most recent electoral events (Cantú 2014; Cantú 2019*a*).

Colombia's history of democratization holds several paradoxes. Formally, regular multi-party elections are held since the 1830s. In practice, electoral events consistently trigger fraud accusations both from the citizenry and academic literature (Duque Daza 2019) and candidates, politicians and journalists are regularly assassinated (Bejarano and Pizarro 2005). These and other observations have led scholars of the Colombian case to describe the country as a 'besieged' (Bejarano and Pizarro 2005) or 'fraudulent' electoral democracy (Duque Daza 2019), in which the institutional design and democratic practice of the country diverge strongly, which resembles the Mexican case to a large extent.

In Russia, we study a context of institutionalized authoritarian rule. After the dissolution of the Soviet Union in 1991, meaningful opposition has effectively been banned since the beginning of Vladimir Putin's administration in 1999 and several observers note that election-day fraud has metasized since in the earlier 2000s (Myagkov, Ordeshook and Shakin 2009). Consistently accompanied by widespread fraud allegations, Russian elections are also occasionally followed by anti-regime protests (Lankina and Tertytchnaya 2020; Robertson 2017). Election monitors routinely condemn Russian elections and a whole range of scholarly contributions focuses on highly unusual patterns in published voting returns that are hard to explain with processes other than manual alteration of vote counts (Rozenas 2017; Klimek et al. 2012; Myagkov, Ordeshook and Shakin 2009; Jimenez, Hidalgo and Klimek 2017; Kobak, Shpilkin and Pshenichnikov 2016*a*,*b*, 2018). Studies suggest that in the 2011 parliamentary elections, the vote share of the incumbent United Russia party was at least 11 percentage points lower than documented by official figures (Enikolopov et al. 2013).

Today, only about 15.5% of Mexicans, 20.5% of Colombians, and about 39.8% of Russians say that election officials are fair and that votes are counted free and fairly.[13]

C. Survey Experiment

a. Descriptive Statistics

Tables A2 and A3 contain summary statistics for the survey responses in Russia and Latin America for the observations we used in the main analysis, for each treatment group. As per the pre-registration analysis plan, we exclude any unfinished interviews and cases where the interviews were finished in under 3 minutes to control for data quality. To represent the target population, i.e. potential voters, we also remove responses from respondents under 18 and non-citizens of target countries, which was also set as a restriction for participation at the recruitment stage. Our sample consists of primarily young, urban, middle-to-high educated people in both Russia and Latin America.

---

[13]Source: World Values Survey Wave 7 (Inglehart et al. 2020), 2017-2021.

Table A2: Summary Statistics of Key Variables, Survey Data for Russia

		Fraud (N=297)		Control (N=310)		Punishment (N=309)		Judicial Punishment (N=307)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Confidence in Armed Forced		2.7	1.0	2.8	1.0	2.8	1.0	2.9	0.9
Confidence in Police		2.3	0.9	2.3	0.9	2.3	0.9	2.4	0.9
Confidence in CEC		1.9	0.9	2.2	1.0	1.9	0.9	2.0	1.0
Confidence in Government		2.1	1.0	2.2	1.0	2.1	0.9	2.3	1.0
Confidence in Parties		1.9	0.8	2.1	0.8	2.0	0.9	2.1	0.8
Confidence in Parliament		2.0	0.9	2.1	0.9	2.0	0.9	2.2	0.9
Confidence in Courts		2.2	0.9	2.3	0.9	2.2	0.8	2.4	1.0
Confidence in President		2.4	1.1	2.5	1.1	2.4	1.0	2.6	1.1
Confidence in Elections		1.9	0.9	2.3	0.9	1.9	0.9	2.2	1.0
Confidence in Companies		2.1	0.9	2.2	0.8	2.2	0.8	2.3	0.9
Confidence in Banks		2.1	0.9	2.2	0.9	2.2	0.9	2.2	0.9
Confidence in Environmental Organizations		2.5	0.9	2.6	0.9	2.5	0.9	2.6	1.0
Confidence in UN		2.3	0.9	2.4	0.9	2.3	0.9	2.4	1.0
Confidence in World Bank		2.2	0.9	2.2	0.9	2.2	0.9	2.3	0.9
Confidence in WTO		2.2	0.9	2.2	0.9	2.2	0.8	2.3	0.9
Political Interest		2.6	0.8	2.6	0.8	2.7	0.8	2.6	0.8
Age		36.2	10.6	36.1	10.7	36.6	10.8	36.1	10.4
Pol. Corruption		7.9	2.1	8.2	1.9	8.1	1.9	8.0	1.9
Involvement		2.7	0.7	2.7	0.8	2.8	0.7	2.7	0.7
		N	%	N	%	N	%	N	%
Generalized Trust	Most people can be trusted	55	18.5	55	17.7	62	20.1	64	20.8
	Need to be very careful	227	76.4	244	78.7	240	77.7	228	74.3
Sex	Female	124	41.8	158	51.0	155	50.2	151	49.2
	Male	173	58.2	152	49.0	154	49.8	156	50.8
Education	Lower	7	2.4	10	3.2	6	1.9	7	2.3
	Middle	133	44.8	132	42.6	154	49.8	131	42.7
	Higher	153	51.5	166	53.5	147	47.6	167	54.4
Empl. Status	Paid employment	182	61.3	167	53.9	167	54.0	182	59.3
	Retired/pensioned	15	5.1	17	5.5	17	5.5	16	5.2
	Housewife	34	11.4	45	14.5	41	13.3	31	10.1
	Student	13	4.4	23	7.4	19	6.1	23	7.5
	Unemployed	38	12.8	42	13.5	48	15.5	33	10.7
	Other	15	5.1	16	5.2	17	5.5	21	6.8
Empl. Sector	Government or public institution	56	18.9	74	23.9	78	25.2	65	21.2
	Private business or industry	184	62.0	180	58.1	184	59.5	180	58.6
	Private non-profit organization	24	8.1	25	8.1	20	6.5	25	8.1
Savings: Last year	Saved money	46	15.5	55	17.7	48	15.5	66	21.5
	Just got by	148	49.8	150	48.4	164	53.1	157	51.1
	Spent some savings	59	19.9	63	20.3	57	18.4	52	16.9
	Spent savings and borrowed money	44	14.8	42	13.5	40	12.9	32	10.4
Set. Type	Rural	25	8.4	28	9.0	20	6.5	18	5.9
	Urban	272	91.6	282	91.0	289	93.5	289	94.1
Opponent	No	121	40.7	125	40.3	129	41.7	127	41.4
	Yes	176	59.3	185	59.7	180	58.3	180	58.6

Table A3: Summary Statistics of Key Variables, Survey Data for Latin America

		Fraud (N=222)		Control (N=218)		Punishment (N=216)		Judicial Punishment (N=216)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Confidence in Armed Forced		2.3	0.9	2.5	0.9	2.2	0.9	2.3	0.9
Confidence in Police		1.8	0.7	1.9	0.7	1.8	0.8	1.9	0.8
Confidence in CEC		1.8	0.8	2.3	0.8	1.9	0.8	1.9	0.8
Confidence in Government		1.7	0.8	2.0	0.8	1.7	0.7	1.7	0.8
Confidence in Parties		1.6	0.7	1.8	0.7	1.6	0.6	1.6	0.6
Confidence in Pariament		1.7	0.7	2.0	0.8	1.7	0.7	1.8	0.8
Confidence in Courts		1.9	0.8	2.2	0.8	1.9	0.7	2.1	0.9
Confidence in President		1.7	0.8	2.0	1.0	1.7	0.9	1.8	0.9
Confidence in Elections		1.8	0.8	2.3	0.8	1.9	0.8	2.0	0.8
Confidence in Companies		2.2	0.8	2.3	0.9	2.2	0.8	2.3	0.9
Confidence in Banks		2.3	0.9	2.2	0.9	2.1	0.8	2.3	0.9
Confidence in Environmental Organizations		3.0	0.8	3.1	0.8	2.9	0.8	2.8	0.9
Confidence in UN		2.8	0.9	3.0	0.8	2.8	0.9	2.9	1.0
Confidence in World Bank		2.5	0.9	2.6	0.9	2.5	0.9	2.5	0.9
Confidence in WTO		2.6	0.8	2.7	0.8	2.5	0.8	2.5	0.8
Political Interest		3.0	0.8	2.9	0.7	3.1	0.7	3.0	0.8
Age		26.3	7.4	27.1	7.9	27.6	8.8	27.1	8.0
Pol. Corruption		9.1	1.5	9.2	1.4	9.1	1.5	9.2	1.1
Involvement		2.8	0.6	2.9	0.5	2.8	0.6	2.8	0.5
		N	%	N	%	N	%	N	%
Generalized Trust	Most people can be trusted	29	13.1	19	8.7	29	13.4	25	11.6
	Need to be very careful	192	86.5	194	89.0	184	85.2	190	88.0
Sex	Female	101	45.5	95	43.6	106	49.1	95	44.0
	Male	121	54.5	123	56.4	110	50.9	121	56.0
Education	Lower	3	1.4	2	0.9	5	2.3	7	3.2
	Middle	151	68.0	141	64.7	131	60.6	137	63.4
	Higher	66	29.7	71	32.6	72	33.3	68	31.5
Empl. Status	Paid employment	71	32.0	99	45.4	83	38.4	74	34.3
	Retired/pensioned	0	0.0	1	0.5	2	0.9	1	0.5
	Housewife	13	5.9	14	6.4	16	7.4	19	8.8
	Student	78	35.1	68	31.2	68	31.5	67	31.0
	Unemployed	51	23.0	29	13.3	33	15.3	42	19.4
	Other	9	4.1	7	3.2	14	6.5	12	5.6
Empl. Sector	Government or public institution	13	5.9	18	8.3	14	6.5	19	8.8
	Private business or industry	119	53.6	121	55.5	119	55.1	108	50.0
	Private non-profit organization	16	7.2	14	6.4	15	6.9	15	6.9
Savings: Last year	Saved money	37	16.7	36	16.5	33	15.3	41	19.0
	Just got by	92	41.4	92	42.2	97	44.9	95	44.0
	Spent some savings	43	19.4	40	18.3	47	21.8	38	17.6
	Spent savings and borrowed money	49	22.1	50	22.9	38	17.6	42	19.4
Set. Type	Rural	17	7.7	16	7.3	9	4.2	15	6.9
	Urban	205	92.3	202	92.7	207	95.8	201	93.1
Opponent	No	56	25.2	54	24.8	53	24.5	56	25.9
	Yes	166	74.8	164	75.2	163	75.5	160	74.1

b. Data Quality

Since our target population of study are the voting-eligible population in the countries of Mexico, Columbia, and Russia, our sampling strategy implies using these characteristics as pre-requisites for participation in the survey. Hence, we excluded all the cases where the reported age of the respondent was below 18 years old, which still occurred despite the crowd-sourcing platform's official age restriction which excludes underage individuals, and the cases where the respondent claimed to not be a citizen of the respective country, also occurring despite the restrictions set on the crowd-sourcing platform.

We are limiting the analysis to complete cases only, which requires us to have a closer look at the dropout and sample response when discussing data quality. As we have relied on the crowd-sourcing platforms for participant recruitment, participants were paid for completing the survey. Given that participants only received the payment upon completion of the survey, we obtained a 95% completion rate in Russia and 92% in Latin American countries.

Apart from the absolute number of completed surveys, a valid concern regarding the dropout rate could be that it varies over our experimental conditions, introducing bias into the sampling procedure and completed cases. As evidenced by the tables below, there seem to be no systematic differences in the dropout rates across the experimental conditions, i.e. dropout rates right after reading the treatment condition.

Table A4: Dropout Rates right after Reading the Treatment across Experimental Conditions, Survey Data in Russia.

Condition		Continued	Dropped Out	All
Fraud	N	317	4	321
	% row	98.75	1.25	100.00
Control	N	342	5	347
	% row	98.56	1.44	100.00
Punishment	N	325	5	330
	% row	98.48	1.52	100.00
Judicial Punishment	N	323	4	327
	% row	98.78	1.22	100.00
All	N	1324	18	1342
	% row	98.66	1.34	100.00

Furthermore, as per the pre-analysis plan, we have excluded cases where the survey was completed in under 3 minutes, as reading and answering all the questions meaningfully in a shorter amount of time seems to be unrealistic. For reference, we provide summary statistics for the time required for completing the treatment text page for the finished interviews, which required summarizing the text, and statistics for institutional trust evaluation page. While small variations in completing the treatment page may be related to small differences in the length of treatment conditions, we can conclude that the lower bound of 3 minutes seems reasonable for each condition for any of them.

Table A5: Dropout Rates right after Reading the Treatment across Experimental Conditions, Survey Data in Latin America

Condition		Continued	Dropped Out	All
Fraud	N	243	3	246
	% row	98.78	1.22	100.00
Control	N	236	3	239
	% row	98.74	1.26	100.00
Punishment	N	235	0	235
	% row	100.00	0.00	100.00
Judicial Punishment	N	250	1	251
	% row	99.60	0.40	100.00
All	N	991	7	998
	% row	99.30	0.70	100.00

Table A6: Time required for completion in seconds (finished interviews), Survey Data for Russia.

	Fraud (N=308)		Control (N=327)		Punishment (N=316)		Jud. Punishment (N=311)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Total	501.16	223.98	432.18	174.26	502.97	212.07	464.21	171.22
Treatment text	168.56	132.93	131.45	89.66	188.01	148.41	174.52	170.61
Trust eval.	85.50	49.52	82.72	52.99	89.13	49.01	85.41	83.13

Table A7: Time required for completion in seconds (finished interviews), Survey Data for Latin America

	Fraud (N=228)		Control (N=226)		Punishment (N=224)		Jud. Punishment (N=230)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Total	483.97	211.53	502.75	243.18	503.88	205.97	520.13	245.08
Treatment text	153.98	160.88	157.73	151.82	157.23	105.26	186.15	211.76
Trust eval.	80.54	41.68	93.14	92.87	96.40	138.83	92.97	122.23

We have also manually coded the summaries to differentiate the responses by the levels of engagement to allow for robustness checks. We have distinguished between *complete summaries*, which contained the condition-specific part of the treatment, *incomplete summaries*, where the responses described only a part of the treatment message but lack the crucial info for that condition; *responses* to the texts which allowed us to expect that the treatment text was read; *copy-paste* of one or multiple paragraphs from the treatment text in the questionnaire, and *unacceptable answer*, which does not allow us to observe any meaningful engagement with the treatment (e.g., a sequence of random characters). While we avoid removing any data based on this criterion in the main analysis due to potential introduction of post-treatment bias though case selection, we report the analysis on the restricted samples in robustness section, and the results show that our findings hold also when relying on a more restrictive approach to data preparation. Furthermore, for the cases where we received complete summaries, i.e. we can be certain that the treatment was read, the effects are particularly pronounced.

Table A8: Data quality categories based on manual coding for Russian sample

Attention check	N	%
Summary	667	54.54
Incomplete	419	34.26
Response	101	8.26
Unacceptable	22	1.80
Copy-paste	14	1.14

Table A9: Data quality categories based on manual coding for Latin American sample

Attention check	N	%
Summary	376	43.12
Response	317	36.35
Incomplete	151	17.32
Unacceptable	25	2.87
Copy-paste	3	0.34

c.   Raw Regression Results

Dependent Variable: Political Institutions

Table A10: Ordinal logistic regression estimates of treatment effects for Latin American pooled sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	-1.44^*	-0.59^*	-0.34^*	-0.10
	$[-1.72; -1.18]$	$[-0.85; -0.33]$	$[-0.60; -0.09]$	$[-0.35; 0.16]$
Intercept$_2$	0.33^*	1.67^*	1.58^*	1.93^*
	$[0.07; 0.58]$	$[1.40; 1.97]$	$[1.29; 1.86]$	$[1.65; 2.23]$
Intercept$_3$	2.25^*	3.55^*	3.50^*	3.75^*
	$[1.95; 2.57]$	$[3.12; 4.02]$	$[3.11; 3.91]$	$[3.31; 4.22]$
Control	0.29	0.31	1.13^*	0.68^*
	$[-0.06; 0.63]$	$[-0.05; 0.67]$	$[0.78; 1.49]$	$[0.33; 1.04]$
Punishment	-0.30	-0.18	0.23	0.05
	$[-0.65; 0.05]$	$[-0.53; 0.19]$	$[-0.12; 0.60]$	$[-0.31; 0.40]$
Judicial Punishment	-0.07	0.12	0.24	0.18
	$[-0.42; 0.27]$	$[-0.24; 0.48]$	$[-0.11; 0.60]$	$[-0.18; 0.54]$
Observations	849	851	860	861
	Political Parties	Parliament	Courts	President
Intercept$_1$	-0.01	-0.25	-0.83^*	0.10
	$[-0.28; 0.24]$	$[-0.50; 0.00]$	$[-1.09; -0.57]$	$[-0.16; 0.35]$
Intercept$_2$	2.56^*	1.95^*	1.30^*	1.64^*
	$[2.23; 2.90]$	$[1.66; 2.25]$	$[1.03; 1.58]$	$[1.36; 1.92]$
Intercept$_3$	4.81^*	3.95^*	3.48^*	3.02^*
	$[4.16; 5.58]$	$[3.49; 4.46]$	$[3.08; 3.92]$	$[2.67; 3.39]$
Control	0.55^*	0.64^*	0.55^*	0.69^*
	$[0.18; 0.91]$	$[0.28; 1.00]$	$[0.20; 0.91]$	$[0.33; 1.05]$
Punishment	0.11	0.07	-0.06	0.07
	$[-0.25; 0.48]$	$[-0.28; 0.43]$	$[-0.41; 0.30]$	$[-0.28; 0.43]$
Judicial Punishment	0.16	0.34	0.42^*	0.21
	$[-0.21; 0.52]$	$[-0.02; 0.70]$	$[0.07; 0.78]$	$[-0.15; 0.56]$
Observations	863	858	850	857

* Null hypothesis value outside 95% credible interval. Reported are medians and 95% credible intervals.
Fraud treatment group serves as the baseline.

Table A11: Ordinal logistic regression estimates of treatment effects for Russian sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	-1.67^*	-1.27^*	-0.36^*	-0.74^*
	$[-1.86; -1.47]$	$[-1.46; -1.09]$	$[-0.54; -0.18]$	$[-0.92; -0.56]$
Intercept$_2$	-0.45^*	0.29^*	1.02^*	0.72^*
	$[-0.62; -0.27]$	$[0.11; 0.47]$	$[0.84; 1.21]$	$[0.54; 0.90]$
Intercept$_3$	1.23^*	2.39^*	2.70^*	2.50^*
	$[1.04; 1.41]$	$[2.17; 2.62]$	$[2.47; 2.95]$	$[2.28; 2.73]$
Control	0.21	0.02	0.55^*	0.19
	$[-0.03; 0.45]$	$[-0.22; 0.26]$	$[0.31; 0.79]$	$[-0.05; 0.43]$
Punishment	0.23	0.10	-0.06	0.02
	$[-0.01; 0.48]$	$[-0.15; 0.34]$	$[-0.29; 0.19]$	$[-0.21; 0.26]$
Judicial Punishment	0.43^*	0.29^*	0.23	0.37^*
	$[0.20; 0.68]$	$[0.05; 0.54]$	$[-0.01; 0.48]$	$[0.13; 0.61]$
Observations	1167	1186	1180	1185
	Political Parties	Parliament	Courts	President
Intercept$_1$	-0.63^*	-0.66^*	-1.01^*	-0.95^*
	$[-0.81; -0.45]$	$[-0.84; -0.49]$	$[-1.20; -0.83]$	$[-1.13; -0.77]$
Intercept$_2$	1.04^*	0.94^*	0.54^*	0.10
	$[0.86; 1.23]$	$[0.76; 1.13]$	$[0.36; 0.72]$	$[-0.08; 0.28]$
Intercept$_3$	3.54^*	3.15^*	2.62^*	1.46^*
	$[3.26; 3.86]$	$[2.89; 3.42]$	$[2.39; 2.86]$	$[1.28; 1.66]$
Control	0.32^*	0.30^*	0.26^*	0.24^*
	$[0.08; 0.57]$	$[0.06; 0.55]$	$[0.02; 0.51]$	$[0.00; 0.48]$
Punishment	0.08	0.07	0.07	0.04
	$[-0.15; 0.33]$	$[-0.17; 0.32]$	$[-0.17; 0.31]$	$[-0.19; 0.29]$
Judicial Punishment	0.33^*	0.40^*	0.43^*	0.31^*
	$[0.08; 0.58]$	$[0.15; 0.65]$	$[0.18; 0.69]$	$[0.07; 0.56]$
Observations	1179	1176	1182	1171

* Null hypothesis value outside 95% credible interval. Reported are medians and 95% credible intervals.
Fraud treatment group serves as the baseline.

Table A12: Ordinal logistic regression estimates of treatment effects for Latin American pooled sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	−2.19*	−1.05*	−0.33	−0.61*
	[−2.68; −1.70]	[−1.57; −0.54]	[−0.82; 0.17]	[−1.12; −0.10]
Intercept$_2$	−0.37	1.24*	1.59*	1.48*
	[−0.83; 0.10]	[0.73; 1.76]	[1.09; 2.11]	[0.97; 2.01]
Intercept$_3$	1.62*	3.13*	3.52*	3.33*
	[1.13; 2.11]	[2.53; 3.77]	[2.95; 4.11]	[2.72; 3.97]
Control	0.20	0.23	1.05*	0.81*
	[−0.48; 0.89]	[−0.47; 0.94]	[0.35; 1.76]	[0.09; 1.53]
Punishment	−0.47	−0.24	0.27	0.17
	[−1.14; 0.21]	[−0.98; 0.48]	[−0.43; 0.97]	[−0.54; 0.89]
Judicial Punishment	−0.26	0.04	0.16	0.37
	[−0.91; 0.40]	[−0.68; 0.76]	[−0.54; 0.86]	[−0.34; 1.07]
Opponent	−0.96*	−0.60*	0.02	−0.66*
	[−1.51; −0.42]	[−1.18; −0.03]	[−0.55; 0.59]	[−1.24; −0.07]
Control × Opponent	0.16	0.11	0.11	−0.15
	[−0.62; 0.96]	[−0.71; 0.92]	[−0.69; 0.91]	[−0.96; 0.69]
Punishment × Opponent	0.23	0.08	−0.05	−0.16
	[−0.55; 1.02]	[−0.75; 0.92]	[−0.86; 0.76]	[−0.99; 0.67]
Judicial Punishment × Opponent	0.25	0.10	0.10	−0.25
	[−0.51; 1.02]	[−0.72; 0.93]	[−0.70; 0.91]	[−1.07; 0.57]
Observations	849	851	860	861
	Political Parties	Parliament	Courts	President
Intercept$_1$	−0.17	−0.53*	−1.12*	−0.88*
	[−0.68; 0.34]	[−1.03; −0.03]	[−1.62; −0.61]	[−1.36; −0.39]
Intercept$_2$	2.43*	1.70*	1.03*	0.82*
	[1.89; 2.99]	[1.19; 2.21]	[0.54; 1.53]	[0.34; 1.32]
Intercept$_3$	4.69*	3.71*	3.21*	2.33*
	[3.90; 5.57]	[3.08; 4.36]	[2.63; 3.82]	[1.81; 2.88]
Control	0.94*	0.76*	0.28	1.05*
	[0.20; 1.68]	[0.04; 1.47]	[−0.44; 0.99]	[0.33; 1.77]
Punishment	0.25	0.39	−0.10	0.62
	[−0.47; 0.97]	[−0.31; 1.09]	[−0.79; 0.60]	[−0.07; 1.31]
Judicial Punishment	−0.11	0.11	0.34	0.33
	[−0.83; 0.61]	[−0.58; 0.82]	[−0.36; 1.04]	[−0.34; 1.02]
Opponent	−0.21	−0.37	−0.38	−1.27*
	[−0.80; 0.39]	[−0.94; 0.19]	[−0.95; 0.20]	[−1.84; −0.69]
Control × Opponent	−0.51	−0.14	0.37	−0.34
	[−1.35; 0.33]	[−0.96; 0.69]	[−0.45; 1.20]	[−1.17; 0.49]
Punishment × Opponent	−0.18	−0.42	0.06	−0.67
	[−1.02; 0.65]	[−1.23; 0.39]	[−0.75; 0.87]	[−1.48; 0.14]
Judicial Punishment × Opponent	0.36	0.30	0.13	−0.16
	[−0.47; 1.19]	[−0.51; 1.11]	[−0.69; 0.95]	[−0.96; 0.65]
Observations	863	858	850	857

* Null hypothesis value outside 95% credible interval.. Repoted are medians and 95% credible intervals.

Table A13: Ordinal logistic regression estimates of treatment effects for Russian sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	−2.39*	−1.74*	−1.09*	−1.56*
	[−2.69; −2.09]	[−2.03; −1.46]	[−1.37; −0.81]	[−1.85; −1.28]
Intercept$_2$	−1.11*	−0.12	0.39*	0.03
	[−1.39; −0.83]	[−0.39; 0.15]	[0.11; 0.66]	[−0.24; 0.31]
Intercept$_3$	0.68*	2.07*	2.15*	1.92*
	[0.41; 0.96]	[1.77; 2.38]	[1.84; 2.46]	[1.62; 2.23]
Control	0.24	0.13	0.57*	0.31
	[−0.14; 0.63]	[−0.25; 0.51]	[0.19; 0.96]	[−0.08; 0.68]
Punishment	0.37	0.29	−0.27	−0.07
	[−0.02; 0.75]	[−0.09; 0.68]	[−0.64; 0.11]	[−0.45; 0.30]
Judicial Punishment	0.44*	0.49*	0.15	0.35
	[0.06; 0.81]	[0.11; 0.89]	[−0.24; 0.52]	[−0.04; 0.73]
Opponent	−1.07*	−0.72*	−1.16*	−1.28*
	[−1.42; −0.72]	[−1.07; −0.36]	[−1.52; −0.80]	[−1.65; −0.93]
Control × Opponent	−0.02	−0.17	−0.02	−0.17
	[−0.54; 0.46]	[−0.66; 0.33]	[−0.51; 0.48]	[−0.66; 0.33]
Punishment × Opponent	−0.19	−0.31	0.32	0.16
	[−0.70; 0.30]	[−0.81; 0.18]	[−0.16; 0.82]	[−0.33; 0.65]
Judicial Punishment × Opponent	0.02	−0.29	0.15	0.05
	[−0.48; 0.52]	[−0.80; 0.19]	[−0.35; 0.65]	[−0.45; 0.56]
Observations	1167	1186	1180	1185
	Political Parties	Parliament	Courts	President
Intercept$_1$	−1.43*	−1.29*	−1.81*	−1.96*
	[−1.72; −1.13]	[−1.56; −1.01]	[−2.10; −1.52]	[−2.25; −1.68]
Intercept$_2$	0.41*	0.44*	−0.19	−0.79*
	[0.13; 0.69]	[0.16; 0.71]	[−0.47; 0.09]	[−1.06; −0.52]
Intercept$_3$	3.01*	2.73*	1.97*	0.74*
	[2.66; 3.39]	[2.40; 3.07]	[1.66; 2.28]	[0.47; 1.02]
Control	0.41*	0.44*	0.09	0.26
	[0.03; 0.79]	[0.07; 0.81]	[−0.30; 0.48]	[−0.11; 0.63]
Punishment	0.11	0.22	−0.23	−0.17
	[−0.28; 0.49]	[−0.17; 0.60]	[−0.61; 0.14]	[−0.53; 0.20]
Judicial Punishment	0.28	0.49*	0.16	0.39*
	[−0.10; 0.67]	[0.11; 0.87]	[−0.23; 0.55]	[0.02; 0.77]
Opponent	−1.23*	−0.98*	−1.26*	−1.55*
	[−1.59; −0.86]	[−1.33; −0.63]	[−1.62; −0.90]	[−1.91; −1.19]
Control × Opponent	−0.13	−0.22	0.31	0.00
	[−0.63; 0.38]	[−0.71; 0.27]	[−0.18; 0.81]	[−0.49; 0.49]
Punishment × Opponent	−0.08	−0.22	0.50*	0.36
	[−0.58; 0.43]	[−0.71; 0.29]	[0.01; 1.00]	[−0.12; 0.83]
Judicial Punishment × Opponent	0.10	−0.14	0.47	−0.08
	[−0.42; 0.60]	[−0.65; 0.35]	[−0.04; 0.97]	[−0.57; 0.41]
Observations	1179	1176	1182	1171

* Null hypothesis value outside 95% credible interval.. Repoted are medians and 95% credible intervals.

Dependent Variable: Non-Political Institutions

We ask the respondents to evaluate their trust in non-political institutions as well as in political upon reading the treatment text, which allows us to evaluate the limitations of the spillover theory. As we can observe from the estimates in the tables below, the treatment groups do not differ significantly in their confidence in non-political institutions in both Latin America and Russia when only presented with the fraud information in comparison to the status quo outcome, while the only exception for the main effect seems to the Russian sample and the effect of court intervention as response to fraud in comparison to fraud alone (i.e., the baseline condition). The milder version of system response, i.e. responsible members' exclusion from the commissions, has no effect on trust lost after the fraud information was released. Judicial punishment information seems to increase trust in companies, banks, environmental organisations, and the World Trade Organisation, but not in the United Nations or the World Bank. One could argue that this could be an evidence of a spillover effect, which goes beyond the political sphere and exists for entities that are not portrayed negatively and "pro-Western" in the Russian media. The fact that we observe it specifically in Russia and only in response to the message about the intervention of another political institution could also be interpreted as evidence for the belief updating mechanism in evaluations. As this information may go against the general expectations of the systems' responses to election manipulation in an autocratic state, judicial interventions' consistent restorative effect on trust for various institutions indicates that respondents seem to have updated their beliefs rather than disregard the information that goes counter their current perceptions.

Table A14: Ordinal logistic regression estimates of treatment effects for Latin American pooled sample.

	Companies	Banks	Environmental Organizations
Intercept$_1$	−1.37*	−1.28*	−2.88*
	[−1.59; −1.16]	[−1.50; −1.07]	[−3.18; −2.59]
Intercept$_2$	0.57*	0.42*	−1.19*
	[0.36; 0.77]	[0.21; 0.62]	[−1.40; −0.97]
Intercept$_3$	2.71*	2.49*	0.90*
	[2.44; 2.99]	[2.23; 2.75]	[0.69; 1.11]
Control	0.27	−0.03	0.09
	[−0.01; 0.56]	[−0.32; 0.25]	[−0.19; 0.38]
Punishment	0.04	−0.27	−0.19
	[−0.24; 0.32]	[−0.54; 0.01]	[−0.48; 0.09]
Judicial Punishment	0.21	0.09	−0.29*
	[−0.07; 0.50]	[−0.19; 0.37]	[−0.58; −0.01]
Observations	856	854	852
	United Nations	World Bank	WTO
Intercept$_1$	−2.35*	−1.86*	−2.21*
	[−2.61; −2.09]	[−2.10; −1.63]	[−2.47; −1.95]
Intercept$_2$	−0.67*	−0.08	−0.39*
	[−0.87; −0.46]	[−0.28; 0.13]	[−0.60; −0.18]
Intercept$_3$	1.10*	1.92*	1.95*
	[0.89; 1.31]	[1.68; 2.16]	[1.71; 2.19]
Control	0.27	0.16	0.26
	[−0.00; 0.55]	[−0.13; 0.45]	[−0.03; 0.56]
Punishment	−0.13	−0.08	−0.17
	[−0.41; 0.14]	[−0.36; 0.20]	[−0.46; 0.12]
Judicial Punishment	0.18	−0.01	−0.25
	[−0.10; 0.47]	[−0.29; 0.28]	[−0.54; 0.04]
Observations	850	841	825

* Null hypothesis value outside 95% credible interval.
Fraud treatment group serves as the baseline.

Table A15: Ordinal logistic regression estimates of treatment effects for Russian sample.

|	Companies	Banks	Environmental Organizations
Intercept$_1$	−1.08*	−0.97*	−1.48*
	[−1.26; −0.89]	[−1.16; −0.78]	[−1.68; −1.29]
Intercept$_2$	0.61*	0.61*	−0.14
	[0.42; 0.79]	[0.43; 0.80]	[−0.32; 0.04]
Intercept$_3$	3.11*	3.04*	2.06*
	[2.84; 3.38]	[2.78; 3.30]	[1.85; 2.27]
Control	0.13	0.10	0.19
	[−0.12; 0.37]	[−0.14; 0.35]	[−0.04; 0.43]
Punishment	0.19	0.09	0.12
	[−0.05; 0.43]	[−0.15; 0.34]	[−0.13; 0.36]
Judicial Punishment	0.27*	0.25*	0.27*
	[0.03; 0.53]	[0.01; 0.50]	[0.02; 0.52]
Observations	1151	1173	1166
	United Nations	World Bank	WTO
Intercept$_1$	−1.27*	−0.97*	−1.11*
	[−1.46; −1.07]	[−1.15; −0.78]	[−1.30; −0.92]
Intercept$_2$	0.13	0.39*	0.46*
	[−0.05; 0.31]	[0.21; 0.57]	[0.28; 0.64]
Intercept$_3$	2.19*	2.50*	2.78*
	[1.97; 2.41]	[2.27; 2.73]	[2.54; 3.04]
Control	0.06	0.10	0.12
	[−0.18; 0.31]	[−0.15; 0.35]	[−0.13; 0.36]
Punishment	−0.04	0.06	0.01
	[−0.28; 0.21]	[−0.19; 0.30]	[−0.24; 0.26]
Judicial Punishment	0.15	0.18	0.29*
	[−0.09; 0.40]	[−0.06; 0.44]	[0.03; 0.54]
Observations	1137	1134	1114

* Null hypothesis value outside 95% credible interval.
Fraud treatment group serves as the baseline.

Robustness Checks

This section contains the results of various robustness checks and investigation of further observable implications.

Data Quality Restrictions

Our manual coding of attention checks allowed us to closely examine the cases where the respondents seemed to have read the treatment text closely and in full, as their responses include the treatment-specific scenario details. For such individuals, we should observe strongest effects should our theory hold, though this effect could be counteracted by the fact of smaller sample sizes. Still, we observe stronger effects for more political institutions than in the full sample. In fact, we find strong evidence for the spillover effects of fraud in all countries for this subsample.
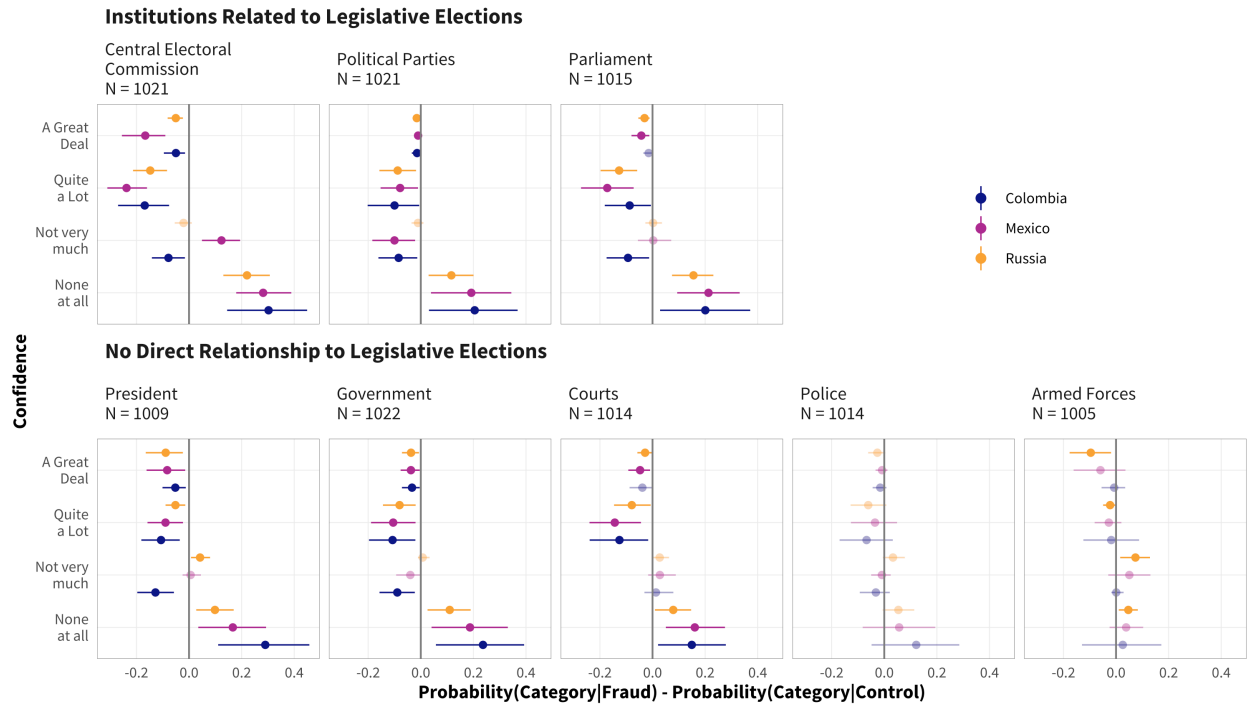
Figure A3: The Effects of Exposure to Fraud Information on Confidence in Political Institutions (only complete summaries of treatment texts). Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.

Subsample Analysis for Fraud Effect
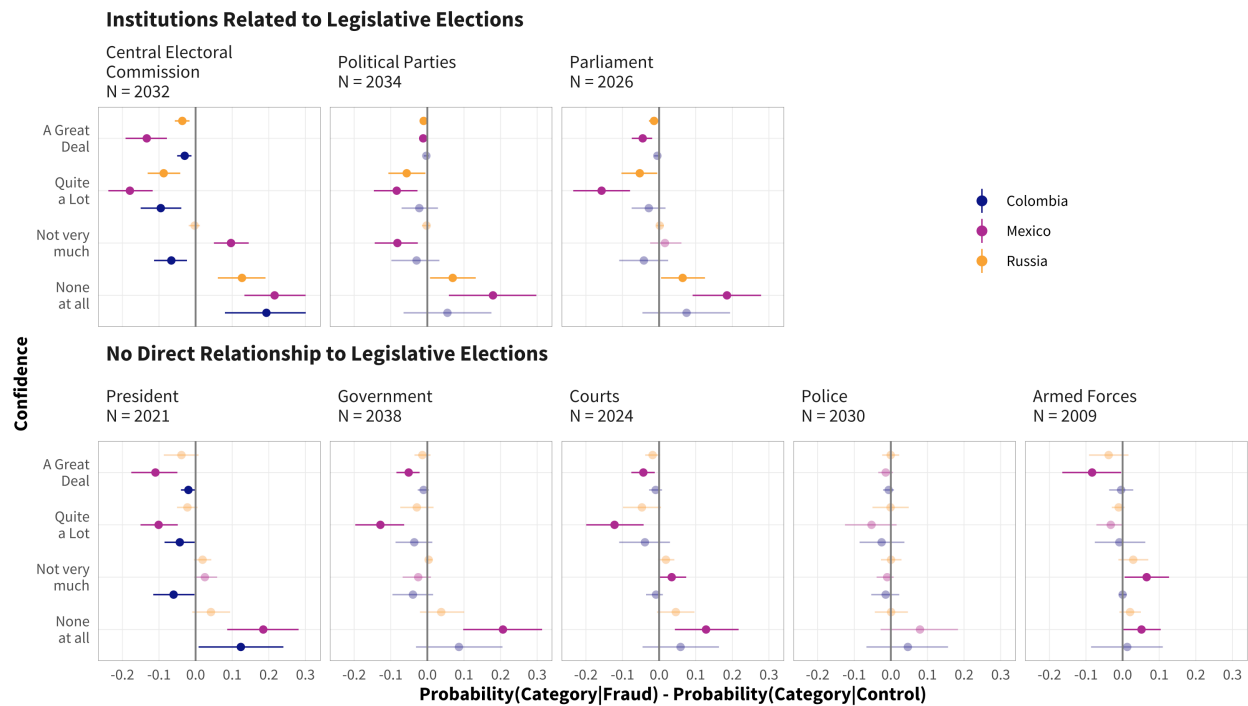
Figure A4: The Effect of Exposure to Fraud and Punishment Information on Confidence in Political Institutions Across Countries. Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.
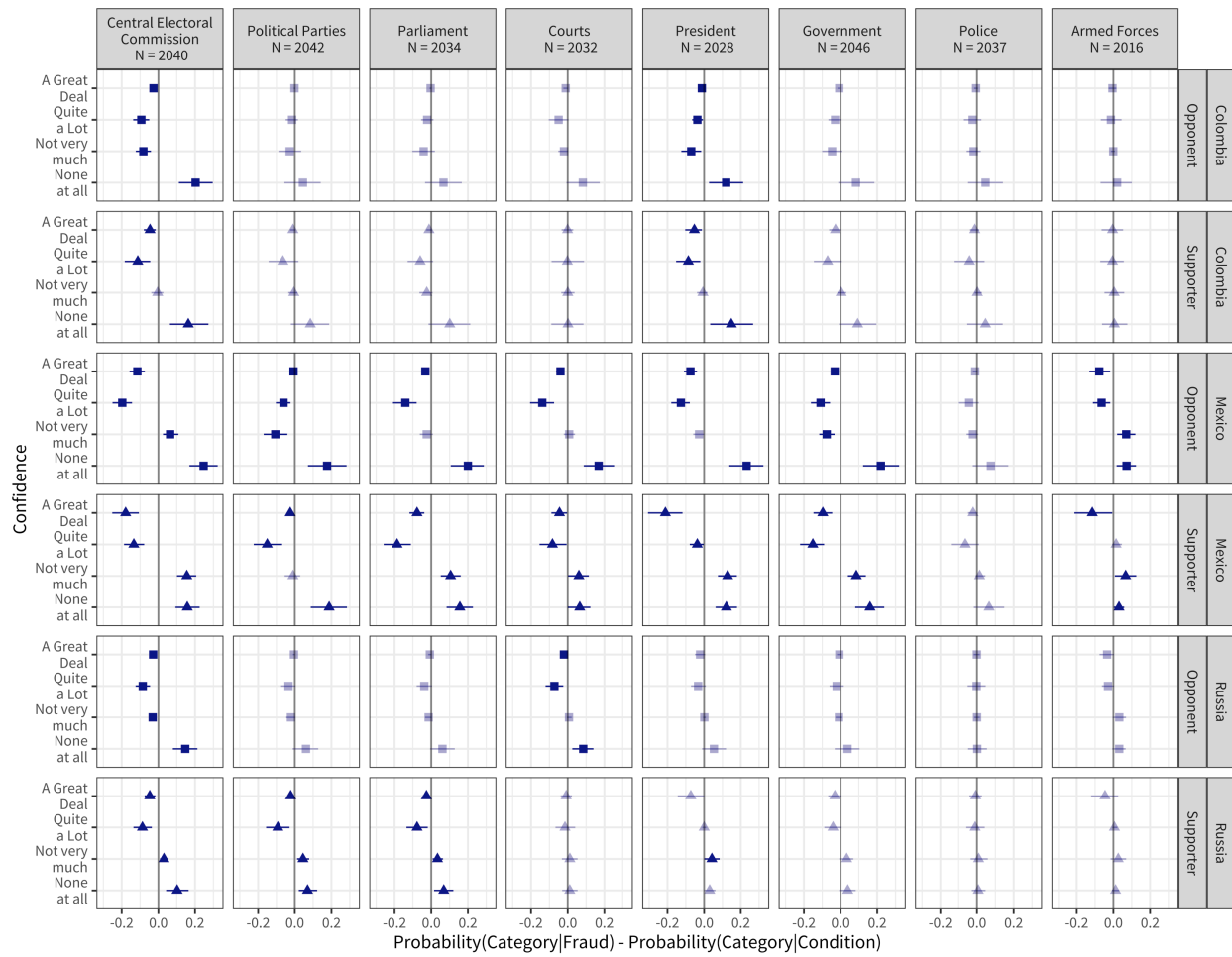
Figure A5: The Effects of Exposure to Fraud Information on Confidence in Political Institutions (across countries, opponents and supporters). Plots depict medians and 83% (bold) and 95% (thin) highest-density intervals of differences in probabilities for choosing respective categories based on draws from the posterior predictive distributions. Transparent point ranges include zero in the 95% HDCI. Probabilities are calculated based ordered logit model estimates as lined out in Equation 1. The dashed line schematically depicts the hypothesized relationship between categories.

Mediation Analysis

In this section we further investigate the mechanisms of attitudes' updating via mediation analysis. The spillover theory implies that it is the changes in trust in elections and electoral process that are responsible for the differences in the confidence in institutions of the political system. We thus have included a separate question that accounts for trust in election system after the treatment (the exact phrasing is: *In this hypothetical scenario, how much confidence would you have in elections in [Mexico/Colombia/Russia]?*) We use answers to this question as a mediator to trace the effects of the fraud and punishment information on political attitudes. Figure below presents the argument graphically:
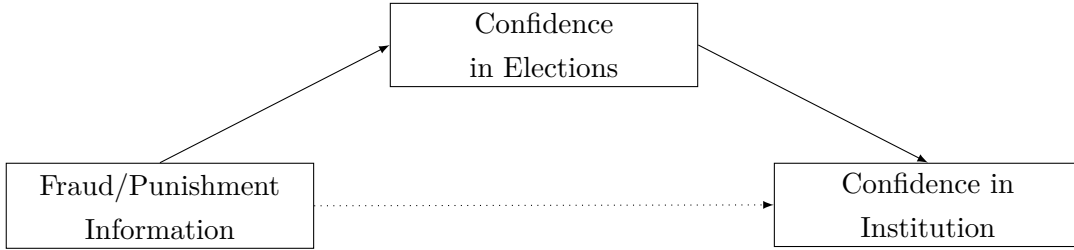
We estimate the ordered logit models specified as follows.

Outcome model:

$$ln\left(\frac{\Pr(y_i \leq j)}{\Pr(y_i > j)}\right) = \alpha_j - \underbrace{(\beta_1 \text{ Control}_i + \beta_2 \text{ Punishment}_i + \beta_3 \text{ Judicial Punishment}_i)}_{\text{Treatment Variables}} + \underbrace{\beta_4 \text{ Trust in Elections}_i}_{\text{Mediator}},$$

(2)

Mediator model:

$$ln\left(\frac{\Pr(\text{Trust in Elections}_i \leq j)}{\Pr(\text{Trust in Elections}_i > j)}\right) = \gamma_j - (\delta_1 \text{ Control}_i + \delta_2 \text{ Punishment}_i + \delta_3 \text{ Judicial Punishment}_i)$$

(3)

where y_i is the level of diffuse support of an individual i with $(i = 1, ..., n)$ for an institution, and Control, Punishment, and Judicial Punishment are binary indicators for membership in the experimental groups (2), (3), and (4) [14]. Trust in Elections$_i$ is the level of trust in electoral process in the country and is on the same scale as the main outcome variable, institutional trust. As in the main analysis, we analyze Russian and Latin American samples separately, using all available observations that fulfill the basic data-quality criteria, such as completing the questionnaire in lower than 3 minutes.

For each political institution, we estimate an ordered logit model for trust in elections and model institutional trust using the estimates from the ordered logit in the previous step. It is our expectation that *direct effect* of treatment on institutional trust, once the model includes trust in elections, is close to zero, while the indirect effect, the *average causal mediation effect*, would be different from zero and positive: as trust categories are both measured on the same scale of 1 to 4, we should expect a positive relationship between them, and in comparison to fraud alone, status

---

[14]Individuals who only received fraud information serve as the reference category in our analysis.

quo information and the punishment would be expected to raise the trust. Figure A6 presents the results of the analysis for Latin American and Russian samples.

When we look at the effect of fraud information alone, we can observe that across all institutions, indirect effect, *ACME*, is significantly different from zero and positive, which is in line with our expectations: increased (decreased) trust in elections is associated with increased (decreased) trust in political institutions. At the same time, the direct effect, i.e. the effect of treatment alone on confidence in institutions, is for most institutions, not different from zero, meaning that most of the changes in institutional confidence in response to treatment seem to be driven by the declines in trust in the elections. For the punishment information, we only observe the effects of treatment once the courts are reported to intervene , and the effect again is primarily existing via trust in elections. For punishment information alone, with no judicial intervention, we observe only the relationship between the trust in elections and trust in

Latin America
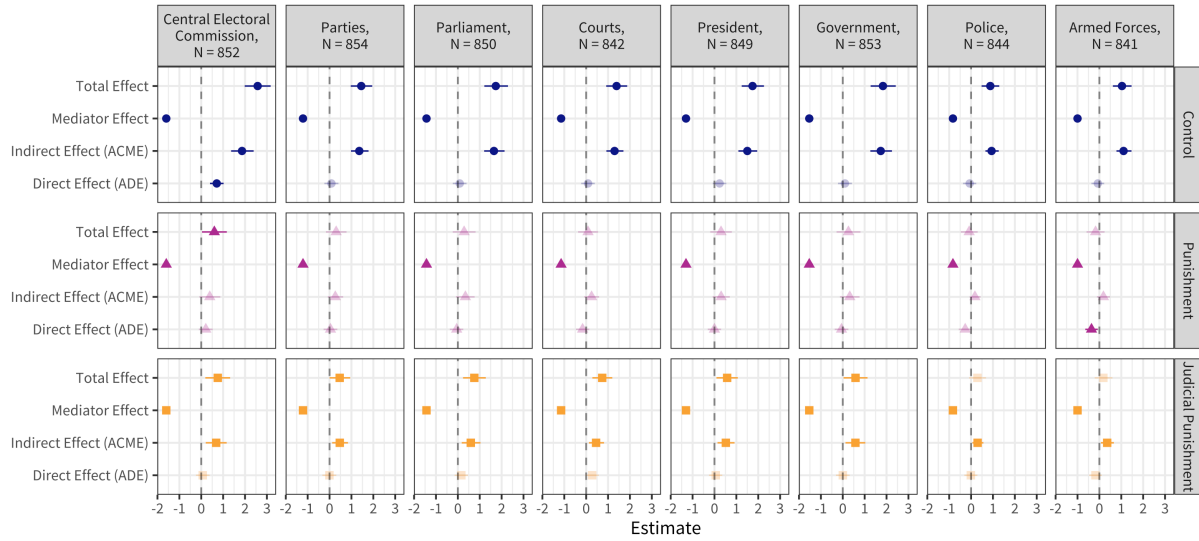
Russia
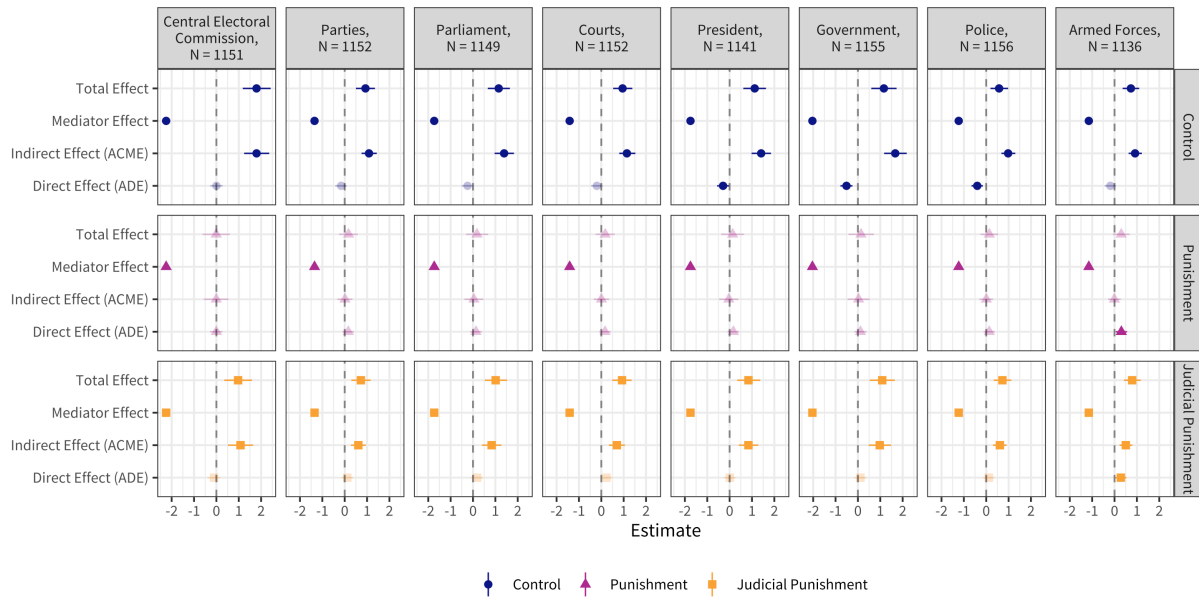
Figure A6: The Results of Mediation Analysis. Depicted are the *direct effect* (median value and 95% highest-density continuous interval (HDCI) of posterior samples from treatment of the institutional trust model), *mediator effect* (median value and 95% HDCI of posterior samples from mediator, trust in elections, of the institutional trust model), *indirect effect* (median value and 95% HDCI of the multiplication of the posterior samples from mediator, trust in elections, of the institutional trust model and the posterior samples from treatment of the trust in elections' model) and the *total effect* (median value and 95% HDCI of sums of posterior samples used for the direct and indirect effect).

Models with Controls

This section contains tables with estimates for the models from our survey experiment, but this time with control variables (tables A16 and A17). While random assignment allows us to drop the controls in our main analysis (for balance assessment across groups see tables A3 and A2), we replicate the results using control variables as well. We use sociodemographic (measured post-treatment) and political attitudes (measured pre-treatment) variables. We have attempted to include the same variables as in our matching analysis for the sake of uniformity and, to a certain extent, comparability. We thus control for

- *generalized trust*, as it may be directly related to confidence in political system institutions;

- *political interest*, as we may expect a relationship between trust and investment into the topic with causal effects pointing in either direction;

- *political affiliation*, as opposition to the regime may decrease trust in its institutions as of itself;

- *age (logged)*, as we may expect younger respondents to show less trust to political institutions;

- *education*, as it may proxy the critical thinking skills and potentially, differences in degrees of sophistication in the evaluations;

- *employment status*, as it may impact the overall government performance evaluation as well as impact the socialisation and information channels available to respondents;

- *employment sector*, as working for the government may be associated with changes in political attitudes;

- *savings*, as economic security and income are known to impact the performance of (and, potentially, trust in) government institutions;

- *political corruption*, as perceptions of corruption are likely directly related to trust in political institutions.

None of these variables are expected to be systematically related to the treatment variables due to random assignment. As a result, the estimates differ marginally with their significance and signs, and follow the patterns we observe in the main analysis.

Table A16: Ordinal logistic regression estimates of treatment effects for Latin American pooled sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	-7.43^*	-4.53^*	-7.89^*	-6.00^*
	$[-10.76; -4.12]$	$[-7.82; -1.27]$	$[-11.18; -4.67]$	$[-9.42; -2.59]$
Intercept$_2$	-5.45^*	-2.31	-6.45^*	-4.06^*
	$[-8.77; -2.17]$	$[-5.55; 0.94]$	$[-9.72; -3.25]$	$[-7.41; -0.66]$
Intercept$_3$	-3.31^*	0.33	-4.22^*	-1.82
	$[-6.61; -0.05]$	$[-2.95; 3.64]$	$[-7.46; -1.00]$	$[-5.19; 1.56]$
Control	0.69^*	0.48	1.58^*	1.16^*
	$[0.20; 1.18]$	$[-0.02; 0.99]$	$[1.08; 2.09]$	$[0.65; 1.68]$
Punishment	-0.35	-0.59^*	0.23	-0.12
	$[-0.86; 0.15]$	$[-1.10; -0.08]$	$[-0.28; 0.73]$	$[-0.66; 0.41]$
Judicial Punishment	-0.28	0.24	0.59^*	0.44
	$[-0.78; 0.23]$	$[-0.28; 0.77]$	$[0.06; 1.11]$	$[-0.09; 0.97]$
Opponent	-0.58^*	-0.28	0.51^*	-1.05^*
	$[-0.96; -0.21]$	$[-0.66; 0.10]$	$[0.13; 0.90]$	$[-1.44; -0.66]$
Political Interest	-0.10	-0.17	0.09	-0.01
	$[-0.36; 0.15]$	$[-0.43; 0.09]$	$[-0.16; 0.34]$	$[-0.28; 0.25]$
General Trust	-0.52	-0.09	-0.32	-1.07^*
	$[-1.08; 0.03]$	$[-0.66; 0.46]$	$[-0.86; 0.20]$	$[-1.66; -0.50]$
Age (log)	-1.11^*	-0.88	-2.29^*	-0.80
	$[-2.02; -0.19]$	$[-1.79; 0.04]$	$[-3.21; -1.38]$	$[-1.71; 0.11]$
Male	0.11	0.35	0.18	0.46^*
	$[-0.27; 0.49]$	$[-0.06; 0.75]$	$[-0.20; 0.56]$	$[0.06; 0.87]$
Education	-0.08	0.02	0.17	0.11
	$[-0.48; 0.31]$	$[-0.37; 0.41]$	$[-0.20; 0.54]$	$[-0.27; 0.49]$
Employment: retired	-0.11	1.35	3.47^*	1.75
	$[-3.28; 3.00]$	$[-1.69; 4.48]$	$[0.35; 6.65]$	$[-1.34; 4.84]$
Employment: housewife	-0.50	-0.24	-0.35	-0.91^*
	$[-1.21; 0.21]$	$[-1.02; 0.52]$	$[-1.07; 0.34]$	$[-1.71; -0.12]$
Employment: student	0.01	-0.52	-0.96^*	-0.20
	$[-0.56; 0.59]$	$[-1.11; 0.07]$	$[-1.52; -0.40]$	$[-0.78; 0.38]$
Employment: unemployed	-0.05	0.19	-0.39	0.52
	$[-0.61; 0.53]$	$[-0.38; 0.75]$	$[-0.97; 0.19]$	$[-0.09; 1.10]$
Employment: other	-1.18^*	-0.42	-0.79	-0.92
	$[-2.16; -0.21]$	$[-1.39; 0.51]$	$[-1.76; 0.13]$	$[-2.01; 0.09]$
Sector: Private Business	0.51	-0.34	-0.32	-0.38
	$[-0.09; 1.11]$	$[-0.93; 0.26]$	$[-0.87; 0.26]$	$[-0.95; 0.18]$
Sector: Non-profit	0.98^*	-0.26	0.35	0.14
	$[0.14; 1.80]$	$[-1.11; 0.54]$	$[-0.46; 1.18]$	$[-0.66; 0.94]$
Savings	-0.10	-0.01	0.11	-0.24^*
	$[-0.29; 0.09]$	$[-0.21; 0.18]$	$[-0.08; 0.29]$	$[-0.44; -0.05]$
Observations	296	295	301	298

* Null hypothesis value outside 95% credible interval.. Repoted are medians and 95% credible intervals.
Respective baseline categories are Fraud, Supporter, Employment: paid employment, Sector: Government or public institution

Ordinal logistic regression estimates of treatment effects for Latin American pooled sample. (cont.)

	Political Parties	Parliament	Courts	President
Intercept$_1$	−10.39*	−8.50*	−6.18*	−7.62*
	[−13.98; −6.96]	[−11.83; −5.19]	[−9.52; −3.02]	[−10.89; −4.25]
Intercept$_2$	−8.16*	−6.22*	−4.07*	−5.79*
	[−11.69; −4.76]	[−9.50; −2.95]	[−7.37; −0.94]	[−9.02; −2.45]
Intercept$_3$	−5.21*	−3.73*	−1.46	−3.98*
	[−8.78; −1.70]	[−6.99; −0.45]	[−4.75; 1.69]	[−7.21; −0.64]
Control	0.88*	1.33*	1.08*	1.36*
	[0.35; 1.41]	[0.82; 1.87]	[0.58; 1.59]	[0.84; 1.88]
Punishment	−0.31	−0.07	0.04	−0.14
	[−0.86; 0.25]	[−0.58; 0.45]	[−0.48; 0.54]	[−0.66; 0.37]
Judicial Punishment	0.17	0.74*	1.01*	0.26
	[−0.37; 0.71]	[0.20; 1.29]	[0.47; 1.55]	[−0.27; 0.79]
Opponent	−0.37	−0.46*	−0.31	−1.88*
	[−0.75; 0.02]	[−0.85; −0.07]	[−0.68; 0.07]	[−2.30; −1.47]
Political Interest	−0.22	−0.07	−0.10	−0.22
	[−0.49; 0.04]	[−0.33; 0.18]	[−0.36; 0.15]	[−0.48; 0.04]
General Trust	−0.94*	−0.51	−0.21	−0.69*
	[−1.49; −0.38]	[−1.07; 0.04]	[−0.74; 0.32]	[−1.24; −0.13]
Age (log)	−2.06*	−1.55*	−1.37*	−0.87
	[−3.04; −1.09]	[−2.47; −0.62]	[−2.30; −0.48]	[−1.78; 0.04]
Male	−0.05	0.05	0.17	0.51*
	[−0.47; 0.36]	[−0.34; 0.44]	[−0.21; 0.55]	[0.11; 0.91]
Education	−0.19	−0.19	0.16	−0.27
	[−0.59; 0.20]	[−0.58; 0.20]	[−0.22; 0.54]	[−0.66; 0.12]
Employment: retired	2.94	1.62	1.25	1.22
	[−0.28; 6.11]	[−1.57; 4.82]	[−1.85; 4.35]	[−1.94; 4.23]
Employment: housewife	−0.22	−0.82*	0.05	−0.50
	[−0.96; 0.52]	[−1.58; −0.08]	[−0.67; 0.77]	[−1.29; 0.29]
Employment: student	−1.18*	−0.92*	−0.14	−0.35
	[−1.83; −0.54]	[−1.50; −0.35]	[−0.72; 0.43]	[−0.91; 0.20]
Employment: unemployed	0.04	−0.11	0.10	−0.06
	[−0.54; 0.64]	[−0.68; 0.46]	[−0.48; 0.68]	[−0.66; 0.54]
Employment: other	−0.13	−1.28*	−0.95	−1.17*
	[−1.12; 0.82]	[−2.33; −0.29]	[−1.96; 0.02]	[−2.33; −0.10]
Sector: Private Business	−0.37	−0.55	−0.50	−0.39
	[−0.97; 0.23]	[−1.14; 0.05]	[−1.04; 0.06]	[−0.98; 0.18]
Sector: Non-profit	−0.09	−0.26	−0.35	0.07
	[−0.96; 0.79]	[−1.09; 0.54]	[−1.13; 0.45]	[−0.71; 0.86]
Savings	−0.09	−0.09	0.01	−0.16
	[−0.29; 0.11]	[−0.28; 0.11]	[−0.18; 0.20]	[−0.35; 0.04]
Observations	299	297	296	295

* Null hypothesis value outside 95% credible interval. Repoted are medians and 95% credible intervals.

Respective baseline categories are Fraud, Supporter, Employment: paid employment, Sector: Government or public institution

Table A17: Ordinal logistic regression estimates of treatment effects for Russian sample.

	Armed Forces	Police	Central EC	Government
Intercept$_1$	−2.02*	−3.71*	−0.66	−3.14*
	[−3.74; −0.25]	[−5.47; −1.95]	[−2.39; 1.06]	[−4.92; −1.44]
Intercept$_2$	−0.71	−2.03*	0.91	−1.48
	[−2.42; 1.06]	[−3.79; −0.28]	[−0.80; 2.61]	[−3.24; 0.23]
Intercept$_3$	1.22	0.30	2.76*	0.50
	[−0.48; 3.00]	[−1.45; 2.06]	[1.02; 4.46]	[−1.25; 2.19]
Control	0.23	−0.04	0.64*	0.25
	[−0.03; 0.49]	[−0.31; 0.23]	[0.38; 0.91]	[−0.02; 0.52]
Punishment	0.26*	0.04	−0.04	0.03
	[0.00; 0.52]	[−0.22; 0.31]	[−0.31; 0.23]	[−0.23; 0.29]
Judicial Punishment	0.47*	0.31*	0.24	0.32*
	[0.21; 0.74]	[0.05; 0.59]	[−0.03; 0.52]	[0.06; 0.60]
Opponent	−1.08*	−0.87*	−1.04*	−1.25*
	[−1.28; −0.88]	[−1.06; −0.68]	[−1.24; −0.84]	[−1.45; −1.05]
Political Interest	−0.13*	0.02	0.01	−0.08
	[−0.26; −0.00]	[−0.11; 0.14]	[−0.12; 0.13]	[−0.21; 0.04]
General Trust	−0.47*	−0.58*	−0.58*	−0.49*
	[−0.70; −0.24]	[−0.81; −0.35]	[−0.82; −0.34]	[−0.72; −0.25]
Age (log)	0.62*	0.12	0.78*	0.21
	[0.22; 1.05]	[−0.29; 0.54]	[0.36; 1.20]	[−0.22; 0.62]
Male	0.31*	−0.20	0.04	−0.07
	[0.11; 0.52]	[−0.40; 0.01]	[−0.17; 0.25]	[−0.28; 0.14]
Education	−0.16	−0.21	−0.29*	−0.15
	[−0.39; 0.07]	[−0.44; 0.01]	[−0.50; −0.07]	[−0.36; 0.07]
Employment: retired	0.32	−0.24	0.38	0.54*
	[−0.17; 0.81]	[−0.71; 0.24]	[−0.10; 0.86]	[0.05; 1.03]
Employment: housewife	0.02	−0.07	−0.13	−0.23
	[−0.29; 0.31]	[−0.37; 0.24]	[−0.45; 0.19]	[−0.54; 0.09]
Employment: student	−0.00	0.55	0.26	−0.40
	[−0.77; 0.75]	[−0.16; 1.26]	[−0.61; 1.11]	[−1.21; 0.40]
Employment: unemployed	−0.28	−0.18	0.47*	0.13
	[−0.59; 0.02]	[−0.49; 0.11]	[0.17; 0.77]	[−0.17; 0.43]
Employment: other	0.43	0.38	0.52*	0.69*
	[−0.04; 0.88]	[−0.06; 0.85]	[0.06; 0.97]	[0.22; 1.17]
Sector: Private Business	−0.10	−0.32*	−0.01	−0.13
	[−0.31; 0.13]	[−0.54; −0.09]	[−0.23; 0.21]	[−0.35; 0.10]
Sector: Non-profit	−0.44*	−0.61*	−0.03	−0.27
	[−0.80; −0.06]	[−1.00; −0.23]	[−0.41; 0.35]	[−0.64; 0.11]
Savings	−0.12*	−0.13*	−0.32*	−0.34*
	[−0.23; −0.02]	[−0.23; −0.03]	[−0.43; −0.21]	[−0.45; −0.23]
Observations	1021	1033	1032	1034

* Null hypothesis value outside 95% credible interval.. Repoted are medians and 95% credible intervals.
Respective baseline categories are Fraud, Supporter, Employment: paid employment, Sector: Government or public institution

Ordinal logistic regression estimates of treatment effects for Russian sample. (cont.)

	Political Parties	Parliament	Courts	President
Intercept$_1$	−4.09*	−4.13*	−4.68*	−1.10
	[−5.89; −2.33]	[−5.92; −2.38]	[−6.41; −2.98]	[−2.82; 0.62]
Intercept$_2$	−2.19*	−2.31*	−3.03*	0.12
	[−3.96; −0.45]	[−4.07; −0.57]	[−4.76; −1.35]	[−1.61; 1.83]
Intercept$_3$	0.47	0.09	−0.76	1.77*
	[−1.33; 2.22]	[−1.68; 1.83]	[−2.47; 0.93]	[0.04; 3.50]
Control	0.34*	0.33*	0.32*	0.31*
	[0.07; 0.61]	[0.06; 0.59]	[0.06; 0.59]	[0.04; 0.58]
Punishment	0.12	0.09	0.12	−0.02
	[−0.14; 0.39]	[−0.18; 0.35]	[−0.14; 0.38]	[−0.28; 0.23]
Judicial Punishment	0.31*	0.36*	0.47*	0.28*
	[0.03; 0.58]	[0.09; 0.63]	[0.19; 0.74]	[0.02; 0.55]
Opponent	−1.11*	−1.07*	−0.93*	−1.43*
	[−1.31; −0.91]	[−1.27; −0.88]	[−1.13; −0.73]	[−1.63; −1.23]
Political Interest	−0.16*	0.02	0.10	0.00
	[−0.29; −0.03]	[−0.11; 0.14]	[−0.02; 0.23]	[−0.12; 0.13]
General Trust	−0.43*	−0.58*	−0.49*	−0.41*
	[−0.68; −0.19]	[−0.82; −0.34]	[−0.72; −0.25]	[−0.64; −0.17]
Age (log)	−0.14	−0.07	−0.48*	0.71*
	[−0.56; 0.29]	[−0.49; 0.34]	[−0.89; −0.08]	[0.29; 1.12]
Male	−0.25*	−0.16	−0.16	−0.11
	[−0.46; −0.04]	[−0.36; 0.06]	[−0.37; 0.05]	[−0.31; 0.10]
Education	−0.23*	−0.23*	−0.10	−0.06
	[−0.45; −0.01]	[−0.45; −0.00]	[−0.32; 0.13]	[−0.27; 0.16]
Employment: retired	0.36	0.26	−0.04	0.23
	[−0.12; 0.84]	[−0.20; 0.72]	[−0.49; 0.43]	[−0.25; 0.72]
Employment: housewife	−0.07	−0.10	−0.17	−0.05
	[−0.39; 0.24]	[−0.41; 0.22]	[−0.47; 0.15]	[−0.35; 0.26]
Employment: student	0.50	−0.05	0.06	0.26
	[−0.27; 1.27]	[−0.79; 0.70]	[−0.65; 0.78]	[−0.47; 1.00]
Employment: unemployed	−0.17	0.15	−0.01	0.19
	[−0.47; 0.13]	[−0.14; 0.45]	[−0.31; 0.28]	[−0.11; 0.49]
Employment: other	0.15	0.40	0.38	0.77*
	[−0.29; 0.60]	[−0.06; 0.86]	[−0.08; 0.82]	[0.29; 1.26]
Sector: Private Business	0.17	−0.12	−0.21	−0.06
	[−0.06; 0.40]	[−0.34; 0.11]	[−0.43; 0.01]	[−0.27; 0.16]
Sector: Non-profit	0.11	−0.20	−0.50*	−0.36*
	[−0.28; 0.50]	[−0.57; 0.18]	[−0.88; −0.13]	[−0.73; −0.00]
Savings	−0.19*	−0.33*	−0.09	−0.31*
	[−0.30; −0.08]	[−0.44; −0.22]	[−0.20; 0.01]	[−0.42; −0.20]
Observations	1028	1027	1032	1025

* Null hypothesis value outside 95% credible interval. Repoted are medians and 95% credible intervals.
Respective baseline categories are Fraud, Supporter, Employment: paid employment, Sector: Government or public institution