Independent media and political engagement in an autocracy: Evidence from Kazakhstan*

By Dina O'Brien†
University of Toronto

Draft: August 22, 2025

Abstract

Durable authoritarian regimes often rely on censorship to maintain control. This study investigates how access to locally sourced, uncensored media influences beliefs, attitudes, and political engagement within an autocratic context. In collaboration with an independent media organization in Kazakhstan, I conducted a field experiment in which individuals were randomly assigned to receive non-mainstream political information via a fact-checking service. The results show that exposure to truthful information prompts individuals to update their beliefs about the country's limited democratic practices and increases their support for protesters. They become more polarized on the views of the state and opposition and want less information control. However, on average, it also leads to reduced political engagement. While disengagement signals eroding support for the regime, it also suggests a diminished likelihood of regime transition.

*I thank my research partners in Kazakhstan: International Center for Journalism MediaNet and Public Policy Research Center PaperLab. This research is funded by the Weiss Fund for Research in Development Economics at the University of Chicago and Department of Economics, University of Toronto. The study is approved by University of Toronto Research Ethics Board Protocol # 43249 on September 11, 2023. The study was registered in the AEA RCT Registry, RCT ID AEARCTR-0013912. Link to registration: https://www.socialscienceregistry.org/trials/13912. I am grateful to my supervisors, Ceren Baysan, Gustavo Bobonis and Arthur Blouin, and seminar participants at the University of Toronto for their helpful comments and guidance.

†Department of Economics, University of Toronto. Email: dina.obrien@mail.utoronto.ca.

1 Introduction

In an era marked by rising authoritarianism and widespread state-controlled narratives, access to truthful information has become a crucial element in democratic transition. While scholars and policymakers alike have emphasized the empowering role of independent media and information transparency, the behavioral and attitudinal consequences of such access remain complex and underexplored. This study provides new empirical evidence on how individuals respond when exposed to truthful information in a manipulated information environment.

First, I investigate whether there is demand for truthful information in this context. Chen and Yang (2019) find in China that such demand is generally low but can be increased with appropriate encouragement. In contrast, this study lowers the search costs of acquiring truthful information by delivering daily fact-checked messages directly to respondents via a popular social media platform. I assess participants' demand for the service after they have experienced it and become familiar with its content.

Second, I examine whether providing truthful political information in an environment saturated with state-sponsored disinformation affects support for both the autocratic regime and the opposition. Much of the existing research on the persuasive effects of disinformation and efforts to correct it has focused on democratic contexts (Barrera et al., 2020; Zhuravskaya et al., 2020; Henry et al., 2020; Van Duyn and Collier, 2018). However, the stakes of information control are arguably higher in autocracies, where state manipulation is often more pervasive and entrenched. While the spread of social media has improved access to political information (Enikolopov et al., 2020; Manacorda and Tesei, 2020), it has also intensified citizens' exposure to government-controlled narratives regarding state performance and the broader political environment. By randomizing access to independently sourced political information, this study seeks to identify how such exposure influences the political attitudes, beliefs, and behaviors of individuals living under autocratic rule.

The main finding of this study is that greater access to truthful information prompts individuals to revise their beliefs about the state of democracy by 6% and increases their support for protesters by 5%, signaling the potential for greater political awareness. Mirroring these effects, individuals become more polarized in their views of the state and the opposition due to exposure to state manipulation. They also show a higher shared appreciation of independent sources of information as they become less polarized on information control. However, truthful information also leads to significant political disengagement by 15%. Using theory to interpret the disengagement, I conclude that this withdrawal reflects the waning support for the regime. However, given the context of limited opportunities for political activity, the disengagement also diminishes the prospects for collective action and meaningful political change.

Upon closer examination of heterogeneity, disengagement appears most pronounced among individuals with low commitment to democratic norms. In contrast, those who strongly support democratic principles tend to update their beliefs and remain politically engaged. However,

their relatively small numbers are insufficient to counterbalance the broader pattern of political disengagement.

Furthermore, this research demonstrates that although autocrats can shape public beliefs and attitudes through propaganda, even a relatively short intervention involving independent information can alter those beliefs and affect behavior. This suggests that the primary obstacle to democratization may not be citizens' inability to recognize or correct for propaganda, but rather their willingness and ability to act on newly acquired information.

I contribute to the growing literature on the effects of independent media on political outcomes in autocratic settings. Existing research shows that exposure to independent information tends to reduce support for authoritarian regimes, though the effects vary across populations. In Russia, for example, exposure to a new online TV channel decreased support for the regime among traditional media users, while producing a polarizing effect among social media users (Enikolopov et al., 2023). Similarly, in Turkey, independent media exposure was found to increase political polarization (Baysan, 2022). In China, Chen and Yang (2019) find that while demand for uncensored information is generally low, it rises with encouragement, and those exposed to Western media become persistently less satisfied with their government.

Consistent with these findings, I also observe lower support for the regime and higher support for protesters following exposure to independent information. However, it differs in key ways. Unlike Chen and Yang (2019), the corrective information in this study is delivered by a domestic source, addressing concerns over the trust of foreign media. Unlike Enikolopov et al. (2023), this study focuses on the effect of specific fact-checked messages rather than general media exposure. Taken together, these distinctions allow this research to examine the conditions under which accurate political information leads to changes in political engagement. It highlights the value of verifying government narratives while also pointing to the limitations of information-based strategies in driving democratization.

This research also addresses a relatively unexplored question: the role of fact-checking in autocratic contexts. While the impact of independent media in autocracies has been widely studied, this study offers a novel contribution by focusing specifically on the correction of political misinformation that exposes state manipulation. This focus is particularly important for advancing research on authoritarian information control, as it illustrates how independent media strategically leverage social media platforms to enhance their reach and influence (Allcott et al., 2020; Chen and Yang, 2019; Levy, 2021, among others).

2 Context

Kazakhstan is a state with de facto one-party rule, where a long-serving president was succeeded in 2019 by a loyal party member. This transition of presidential power sparked mass protests in 2019, which persisted for several years. The largest of these protests, in January 2022, was brutally suppressed but resulted in increased activity among opposition parties.

While a certain degree of civil liberties exists, allowing for some government criticism, collective action remains restricted – similar to patterns observed online in China (King et al., 2013). Since 2019, citizens have engaged politically in forms other than protests, such as participating in political discussions online and offline, signing petitions, voting and encouraging others to vote and be politically active, and following the campaigns of newly established movements and parties.

The partner organization, International Center for Journalism MediaNet, is a non-governmental media organization in Kazakhstan that focuses on fact-checking through its online platform, "Factcheck.kz," which is available as both a website and on social media. Their emphasis on fact-checking is driven by the significant amount of disinformation in the country, including disinformation disseminated by the state. For instance, an opinion survey conducted in January 2023 found that the people of Kazakhstan are generally unaware of disinformation, with instances of political repression in 2022 particularly susceptible to disinformation. Through its fact-checking efforts, MediaNet addresses sensitive political topics and positions itself as an independent media source, employing a rigorous, internationally approved methodology to evaluate the truthfulness of information from both the government and opposition.

3 Experimental Design

3.1 Intervention

The experiment leverages an existing MediaNet messaging service within its Factcheck.kz project, with Telegram serving as one of the primary platforms. For the purposes of this study, four dedicated Telegram channels were created, separate from MediaNet's main channel, to isolate political content from other topics produced by MediaNet. These channels include one treatment and one control group for each of the two official languages, Kazakh and Russian. Communication on these channels is one-way: participants can react to messages they receive but cannot interact with one another. The intervention was randomized at the individual level and stratified by respondents' preferred language for receiving information. The experiment was conducted over six weeks, from July 3 to August 16, 2024.

The control group was exposed to the type of information they typically consume through mainstream media. Participants in this group received one news item per day. This design ensured that control group members remained active on the same platform as the treatment group, were similarly exposed to political information available through mass media—including state-sponsored disinformation—and maintained engagement throughout the study period.

Figure 1 presents an example message received by the control group.

The treatment group received the same daily news content as the control group, supplemented by an intervention from the International Center for Journalism "MediaNet." This intervention involved daily fact-based messages delivered by MediaNet journalists, covering top-

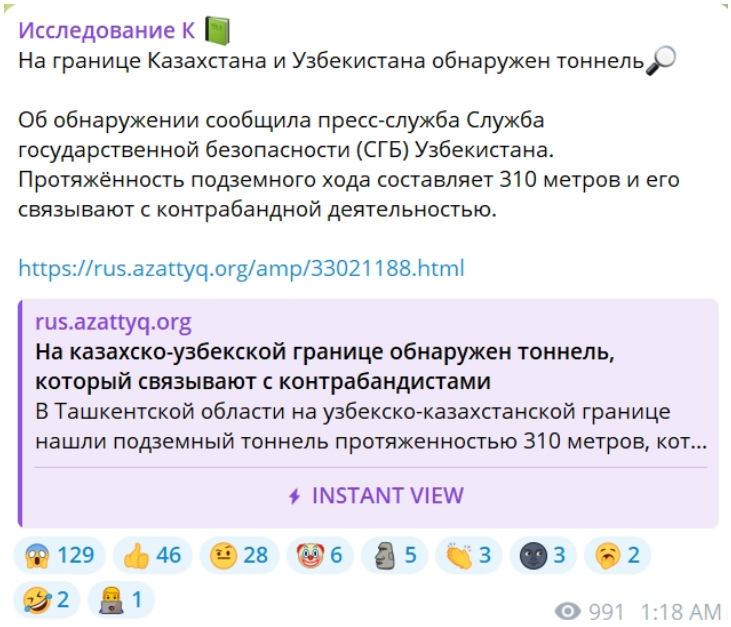


Figure 1: News taken from mainstream media

Translation: *A tunnel has been discovered on the border between Kazakhstan and Uzbekistan, according to a statement from the press corps of the National Security Service of Uzbekistan. The tunnel, which is 310 meters long, is believed to be linked to contraband activities.*

ics related to domestic politics, governance, accountability, and civil liberties. Specifically, the messages focused on deconstructing propaganda while highlighting manipulations by state and non-state actors (e.g., installation of surveillance cameras, control of digital spaces), public monitoring of government activities (such as public spending), and tracking government promises (e.g., annual Presidential addresses), with evidence-based verdicts on instances of state manipulation. On social media, MediaNet summarized these articles and provided links to their website for more detailed evidence and assessments of each claim.

There are two types of intervention messages: (i) new non-mainstream political information, and (ii) fact-checked content that issues verdicts—such as True, Fake, Manipulation, or Technical Error—on pieces of disinformation, including those distributed by the state.

Figures 2 and 3 present examples of the first and second types of treatment messages respectively. The topics covered during the intervention were selected by MediaNet journalists and primarily focused on public spending, personal data breaches, and political action. Appendix Figure A.2 presents the distribution of all topics covered.

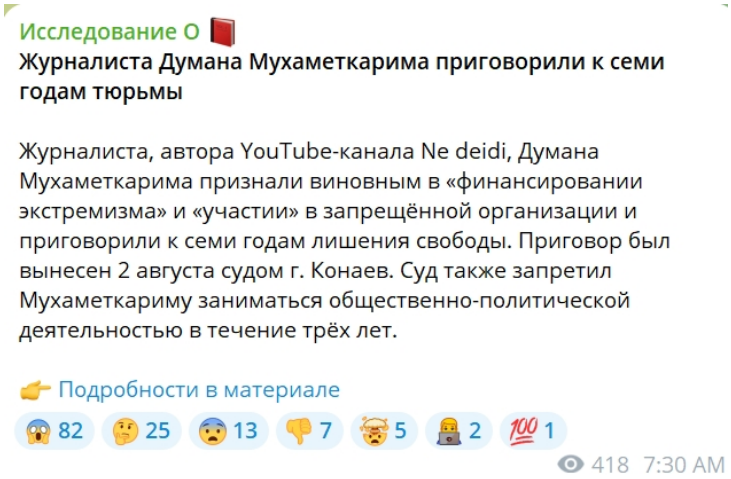


Figure 2: MediaNet information unavailable in mainstream media

Translation: *Journalist Duman Mouhametkarim has been sentenced to seven years in prison. The author of the YouTube channel "Ne deidi", Mouhametkarim was found guilty of "financing extremism" and "participating" in an illegal organization. The sentence was handed down on August 2 in a court in Konaev. Additionally, the court imposed a three-year ban on Mouhametkarim from engaging in any social or political activities.*

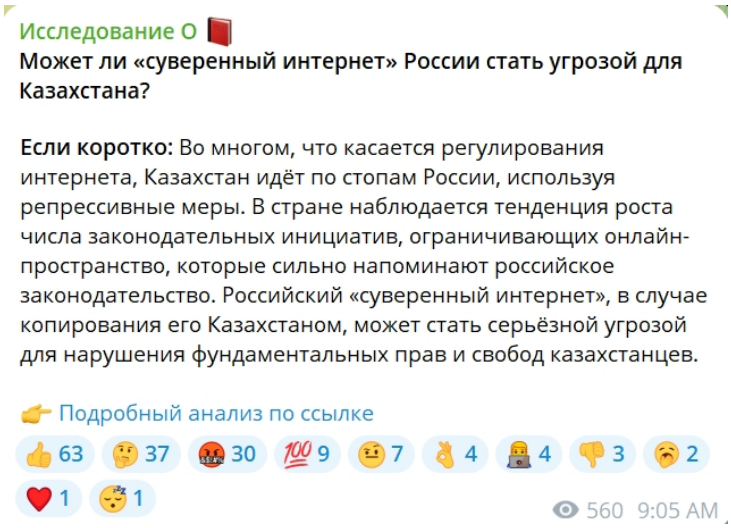


Figure 3: MediaNet information highlighting manipulation

Translation: *Kazakhstan follows Russia's lead on internet regulation by employing repressive measures. The country is seeing an increase in legislation that restricts the online space in a manner similar to Russia. If Kazakhstan adopts Russia's "sovereign internet" model, it could pose a serious threat to the fundamental rights and freedoms of its citizens.*

3.2 Randomization and Procedures

Recruitment: Recruitment took place between May and June 2024, targeting a sample size of 1,400 respondents. The process was conducted in both Russian and Kazakh. Initially, 600 participants were recruited from the survey firm's existing pool of subjects, all of whom had previously agreed to participate in social science research. After exhausting this pool, the firm engaged regional representatives across 20 regions of Kazakhstan to recruit additional participants. This strategy helped reach individuals less likely to participate in online research, thereby better reflecting the broader population of Kazakhstan. This regional recruitment effort added 1,087 new subjects. However, many participants dropped out after recruitment but before completing the baseline survey. To compensate, a supplementary recruitment campaign was launched on the survey firm's Instagram page, resulting in a final total of 1,461 respondents who completed the baseline survey.

Exclusions: The recruitment criteria required participants to be at least 18 years old, reside in Kazakhstan, and provide informed consent to participate in the study. Additionally, participants could not already be subscribed to receive fact-checked information from Factcheck.kz and had to agree to complete an initial recruitment survey consisting of 10 questions.

Randomization: After recruitment, participants completed a baseline survey, the results of which were used to randomize them into treatment and control groups. Randomization was stratified according to the language participants chose for receiving information during the study. MediaNet provides fact-checked content in both Russian and Kazakh. While the information is generally similar across the two languages, it is not always identical, as different journalists produce the content. Nevertheless, many messages are available in both languages.

Among the 1,461 respondents, 296 preferred to receive information in Kazakh, resulting in an imbalance in language distribution. Given this imbalance, stratifying on additional covariates such as urbanity and gender would create very small strata, reducing the effectiveness of stratification. Therefore, while balance on urbanity, gender, and other covariates is checked, it is not explicitly targeted during randomization.

Incentives: Participants in both the treatment and control groups are compensated for their time as follows: approximately 6 CAD (2000 KZT) for completing surveys, 6 CAD (2000 KZT) for subscribing to receive news for six weeks, and 6 CAD (2000 KZT) for receiving political information from MediaNet for six weeks. Additionally, participants may gain non-pecuniary benefits from contributing to a study aimed at enhancing the effectiveness of MediaNet's fact-checking service.

A potential concern is that the treatment group received more political information overall, in addition to it being truthful, while the control group may have had less exposure to political news. To address this, I compared political knowledge between the treatment and control groups. As shown in Table A.1, political knowledge is largely similar across both groups. Although one item shows a statistically significant difference, it is actually higher in the control

group, further supporting the conclusion that the treatment group did not gain more political knowledge overall.

Target Population: The experiment targets individuals on the margins of political participation. In the autocratic context, this group typically does not include the elite and comprises a substantial portion of economically disadvantaged individuals living outside major urban centers. Historically, these individuals have been more likely to engage in protests and labor strikes. Moreover, the study focuses on a critical segment of society: those who are interested in accurate political information but have not yet committed to active political participation.

The sample consists of individuals who read local news daily, predominantly consume information online—unlike the majority of the population, which relies more on television—and are active on popular social media platforms. They report encountering disinformation regularly and are relatively sophisticated consumers: about one-third of pilot respondents acknowledge the difficulty of identifying disinformation, compared to less than 20% of the general population. They are also more likely to verify information when they suspect it may be false and value independent media as a check on government narratives. While most participants are concerned about political issues such as the state of democracy, only about one-third express a strong personal interest in politics.

Attrition and Balance Checks: Figure 4 presents the sample formation process. Out of the 1,461 respondents who were recruited and completed the baseline survey, 788 subscribed to a study channel, and 753 respondents completed the endline survey. Fewer participants in the treatment groups chose to subscribe. However, there are no statistically significant differences in observable characteristics between the treatment and control groups among those who chose not to subscribe or dropped out before the endline survey (attritors).

Out of 788 who subscribed to the channels and received treatment, 69 left the study before completing the endline survey. Among them, similar numbers are in the treated and control groups; therefore, the difference in numbers is not statistically significant. Additionally, the covariates used in the analysis and baseline outcomes of interest do not differ between the treatment and control groups in this subgroup.

Table 1: Balance Check

	(1) Control Mean (SD)		(2) Treatment Mean (SD)		(3) Difference Mean (SE)	
Urban	0.94	(0.23)	0.92	(0.27)	0.02	(0.02)
Male	0.32	(0.47)	0.31	(0.46)	0.02	(0.03)
Age	34.89	(10.67)	34.07	(10.26)	0.82	(0.76)
High School	0.08	(0.27)	0.10	(0.30)	-0.02	(0.02)
Vocational Education	0.20	(0.40)	0.19	(0.40)	0.01	(0.03)
Bachelors Degree	0.57	(0.50)	0.58	(0.49)	-0.00	(0.04)
Masters and up	0.14	(0.35)	0.13	(0.34)	0.01	(0.03)
Democratically governed	1.80	(0.77)	1.81	(0.75)	-0.01	(0.06)
Support protesters	1.76	(0.73)	1.80	(0.74)	-0.04	(0.05)
Searched pol info	0.43	(0.50)	0.42	(0.49)	0.02	(0.04)
Read others	0.42	(0.49)	0.40	(0.49)	0.02	(0.04)
Sent request	0.31	(0.46)	0.31	(0.46)	0.00	(0.03)
Encourage voting	0.31	(0.46)	0.34	(0.47)	-0.03	(0.03)
News consumption frequency	2.65	(1.63)	2.79	(1.69)	-0.14	(0.12)
Internet news consumption frequency	2.31	(1.64)	2.49	(1.70)	-0.18	(0.12)
Top news source: Friends and family	0.49	(0.50)	0.51	(0.50)	-0.02	(0.04)
Top news source: TV	0.23	(0.42)	0.23	(0.42)	-0.00	(0.03)
Top news source: Internet	0.65	(0.48)	0.60	(0.49)	0.05	(0.04)
Top news source: Print media and radio	0.09	(0.28)	0.06	(0.23)	0.03	(0.02)
Top news source: Telegram channels	0.33	(0.47)	0.36	(0.48)	-0.03	(0.03)
Top news source: Other social media	0.74	(0.44)	0.75	(0.43)	-0.02	(0.03)
Top news source: Other	0.03	(0.18)	0.03	(0.16)	0.01	(0.01)
News origin Kazakhstan	0.91	(0.29)	0.90	(0.31)	0.01	(0.02)
Discuss issues frequency	4.80	(1.78)	4.70	(1.77)	0.09	(0.13)
Local politics source: Friends and family	0.47	(0.50)	0.49	(0.50)	-0.03	(0.04)
Local politics source:: TV	0.15	(0.36)	0.16	(0.37)	-0.01	(0.03)
Local politics source:: Internet	0.49	(0.50)	0.39	(0.49)	0.10***	(0.04)
Local politics source:: Print media and radio	0.08	(0.26)	0.03	(0.18)	0.04**	(0.02)
Local politics source:: Telegram channels	0.23	(0.42)	0.23	(0.42)	-0.00	(0.03)
Local politics source:: Other social media	0.56	(0.50)	0.58	(0.49)	-0.02	(0.04)
Local politics source:: Other	0.09	(0.29)	0.08	(0.27)	0.01	(0.02)
N	397		356		753	

Notes: Columns (1) and (2) show means with standard deviations in parentheses. Column (3) shows the difference in means with the standard error of the difference in parentheses. * .10 ** .05 *** .01.

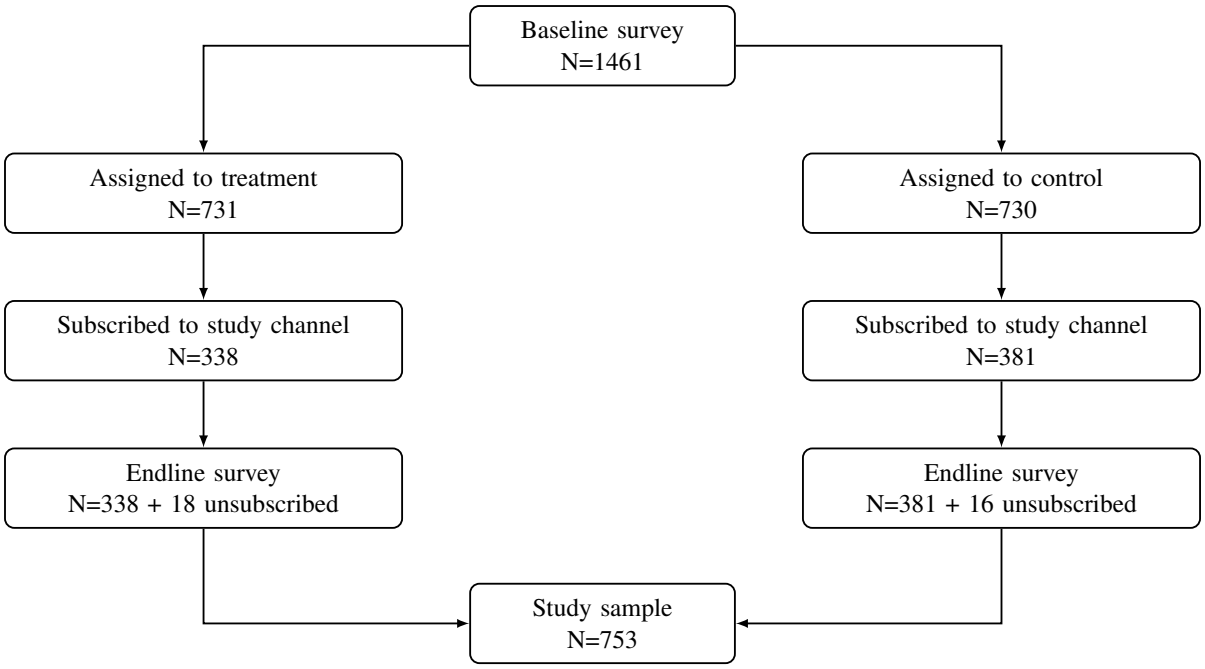
[1] These outcomes are on Lickert 1-3 scale.

Table 1 presents a balance check of covariates, primary outcomes, and other relevant vari-

ables for participants who completed the endline survey and remained subscribed to the treatment channels (stayers). Two political information variables show statistically significant differences between the control and treatment groups, likely due to chance. Specifically, the control group reports higher consumption of political information primarily through the Internet, which supports the argument that the treatment group was not exposed to more information overall. Conversely, the treatment group reports lower consumption of political information via print media and radio; however, since consumption through other sources is not higher, this difference is less concerning. These variables are included in robustness checks to address any potential imbalance in the outcomes of interest.

Additionally, Table A.2 shows no statistical difference between treatment and control among those who subscribed to the study channels as assigned. The compliance rate is 96% among those who subscribed to the channels.

Figure 4: Study design



3.3 Data

Baseline and Endline Survey Data Sets: Baseline and endline surveys are used to measure political attitudes and beliefs. The following data are collected: i) media and political information consumption, ii) knowledge (quiz on political issues and events), iii) beliefs (if democratically governed), iv) attitudes (about protesters and information control), v) values (importance of honest elections and change in government), vi) self-reported behaviours (search and reading of information, requests and encouraging voting).

Additionally, the following covariate variables are collected and considered in the analy-

sis: i) stratum fixed effects: language of communication (Kazakh or Russian), and ii) control variables: urbanity, age, gender, and education.

Experimenter Demand: Participants were informed that they were enrolled in a study evaluating the effect of fact-checked information provided by MediaNet to improve its service. However, they were not explicitly told that the study focused on their political attitudes and beliefs, although they could likely infer this from the survey questions and treatment messages. To assess this potential concern, I included a direct survey question asking respondents to guess the study's hypothesis in an incentivized choice. As shown in Figure A.1, only 11% of respondents correctly identified the study's hypothesis.

The endline survey assesses experimenter demand using the methodology outlined in Quidt et al. (2018). The participants are presented with a strong demand treatment: "The researcher expects that those who were exposed to the Factcheck information would be more critical of the government, be more interested in politics and undertake the actions listed below. You will do the researcher a huge favor if you confirm this in your answers below."

3.4 Outcomes

3.4.1 Primary Outcomes

Demand for Truthful Information: The first outcome of interest is the demand for truthful information. After the study concludes, participants are offered a subscription to the main MediaNet Telegram channel, ensuring that they have the opportunity to continue accessing independent political information. In the endline survey, both treatment and control groups are asked whether they subscribed to this main Telegram channel.

Political Beliefs and Attitudes: To measure participants' beliefs about the state of democracy in the country, the survey asks: "How democratically is this country currently governed" Responses are recorded on a 5-point Likert scale. For greater analytical precision, the outcome is transformed into a 3-point Likert scale by grouping positive responses ("very democratically" and "democratically") together, grouping negative responses similarly, and keeping the neutral middle response separate.

To measure attitudes, the surveys ask about the participants attitude toward protesters: "What is your general attitude toward protesters?" Given that there is no viable opposition in the country, asking about any particular opposition figure may not measure the general attitude toward an alternative leader. Therefore, the attitude toward protesters represents that sentiment. Figure A.3 summarizes the responses at baseline.

Political Behaviour: To measure how treatment affects political engagement, I ask if the respondents if they have engaged in any of the following activities: online (search or read what others are writing) and offline (sent requests to the government or encouraged others to vote). I construct a behaviour index to summarize political engagement. The index is GLS-weighted following Anderson 2008. Figure A.4 summarizes the responses at baseline.

Social Desirability Bias:Given the history of heavy repression against previous protests, there is a potential concern that respondents might be reluctant to openly express support for the protesters. To address this concern, I included a list experiment in the surveys using the "Item Count Technique" to indirectly elicit the extent to which there is true support for protesters in this context.

3.4.2 Secondary Outcomes

Autocracies are generally expected to exhibit higher levels of polarization than democracies. To measure polarization, I use two indicators: polarization in attitudes toward government officials (and opposition) and polarization in views on information control. The survey asks participants to respond to the following statements: "How well does this statement describe you: I find it difficult agreeing with the statements of government officials (opposition)" and "How well does this statement describe you: I think that the government should control the information provided to the public (opposition should freely express their ideas)."

3.4.3 Heterogeneity Variables

I use two variables to assess any heterogeneity in demand: trust of the informaiton source (MediaNet) and value of independent media. I ask the following questions in the baseline survey: "To what extent do you trust or distrust the information provided by MediaNet?" and "How important is it to you to have independent media in the country?"

I anticipate heterogeneity in political beliefs, attitudes, and behavior based on individuals' support for democratic norms. To capture this, I construct a GLS-weighted index using two variables that reflect key democratic principles: support for fair elections and peaceful transitions of power. I use the following survey questions to elicit how much democracy is valued: "How important is it to you to have fair elections?" and "How important is it to you to see changes in the government and politics?" After constructing the index, I divide the sample into two subgroups: those with positive support for democratic norms and those with negative support. Additionally, I examine whether ideology might explain heterogeneity, drawing on literature from democratic countries where political disinformation is often resistant to correction due to ideological biases or motivated reasoning (Barrera et al., 2020). However, in this context, I find that ideology does not play a significant role.

3.5 Empirical Strategy

Given the random assignment of treatment, the the following ITT regression estimates the causal average treatment effect:

$$Y_{is} = \alpha + \beta * Treatment_{is} + \gamma_s + \mathbf{X}'_{is}\delta + \zeta Y_{isbaseline} + \epsilon_{is}, \quad (1)$$

where Y_{is} is an outcome of interest for respondent *i* in stratum *s*, $Treatment_{is}$ is the randomized treatment variable (=1 if the respondent received independent and truthful political information in addition to the news received by the control group), the coefficient β gives the average treatment effect of the treatment, γ_s is language strata fixed effects, $\mathbf{X}_{is}$ is the vector of covariates (urbanity, age, gender, and education), $Y_{isbaseline}$ is the outcome of interest at baseline if available, and ϵ_{is} is the error term. The regressions use heteroskedasticity-robust standard errors.

To test heterogeneous treatment effects, the above equation is modified to include the baseline covariate associated with a heterogeneous treatment effect and its interaction with treatment:

$$Y_{is} = \alpha + \beta * Treatment_{is} + \eta * H_{is} + \rho * Treatment_{is} * H_{is} + \gamma_s + \mathbf{X}'_{is}\delta + \epsilon_{is}, \qquad (2)$$

where H_{is} is the heterogeneity variable measured at baseline.

3.6 Theoretical Framework

The theoretical framework is based on the basic job market signaling model (Spence, 1973) and signaling model of accountability (Guriev and Treisman, 2020, Besley, 2006). The purpose of the theoretical framework is to understand the role the new political information plays and how its exogenous variation is expected to affect political outcomes. It is also useful in disentangling the sign of the observed treatment effect on political action.

Setup

Players: The players in the model include a leader and a continuum of citizens of two types: informed elite and uninformed citizens. The leader is either competent ($\theta = 1$) or incompetent ($\theta = 0$). The leader and the elite observe the leader's type, but the citizens remain uninformed. The leader engages in propaganda and sends a signal $l = \{0, 1\}$ to manipulate the citizens for their support using the controlled state media. The elite send a signal $e = \{0, 1\}$ to the citizens informing them of the leader's type: $e = 1$ (the leader is competent) or $e = 0$ (the leader is incompetent). The signal is uncensored and can reach the citizens.

The uninformed citizens observe the propaganda and elite signals. Using information from both sources they update their beliefs about the leader's type θ. They subsequently decide on the political action, $a = \{0, 1\}$, where $a = 1$ if citizens support the regime and $a = 0$ if citizens protest the regime.

Payoffs: The elite's payoff is given by the following utility function: $u(a, e, \theta) = a - c(e, \theta)$ for a political action a and some cost function c (the basic cost function e/θ can be assumed here). This guarantees that for a given level of political action, the elite signal e is increasing in

θ.

The citizens' payoffs are given by: $\theta - a_i$ for citizen i. The actions of all citizens aggregate to the political action that enters the leader and elite's payoff.

Strategies and Beliefs: The leader's pure strategy is to send $l \rightarrow \{0, 1\}$. Since there is no cost to sending a propaganda signal, the competent leader sends $l = 1$ because otherwise, if $l = 0$, the citizens know that the leader is incompetent.

The elite's pure strategy is to send $e \rightarrow \{0, 1\}$. The elites are assumed to be independent in this model and cannot be bribed to alter their true signal.

The citizens' pure strategy is political action after they have observe signal e from the elites: $a_i(e) \rightarrow \{0, 1\}$.

The common beliefs are denoted by $p(e)$, which is the probability that the leader is competent ($\theta = 1$) when e is observed.

Since the citizens do not know the true leader's type, they form expectations about the leader: $E[\theta|e] = p(\theta|e)\theta_1 + (1 - p(\theta|e))\theta_0$.

The zero-profit condition implies that $a(e) = p(\theta|e)\theta_1 + (1 - p(\theta|e))\theta_0$.

Assumptions: The model assumes that there is a competitive market where citizens hire the leader with the assistance of the elites. Although the setting is autocratic and de facto voting does not exist, the assumption of competition is still relevant, as the citizens can protest the leader and consider hiring an alternative one.

The second assumption is that citizens' actions depend only on the updated beliefs about the leader's type. The first set of results tests that citizens update their beliefs about the democracy in the country. The exclusion condition is not testable and is assumed.

Equilibrium

Elites: The relevant equilibrium is the separating equilibrium ($e(\theta_1) \neq e(\theta_0)$) since the elite send messages that address state manipulations and provide independent information otherwise unavailable through the state media.[1]

Citizens: The players are rational and Bayesian. The Bayesian consistency implies that the citizens update their beliefs in accordance with the elite signal: $p(\theta_1|e(\theta_1)) = 1$ if the elite signal is that the leader is competent and $p(\theta_1|e(\theta_0)) = 0$ if the elite signal is that the leader is incompetent.

Once the elite signal is received and the conclusion about the leader's type is formed, citizens choose a political action. The citizens' actions stem from the zero-profit condition, $a(e) = p(\theta|e)\theta_1 + (1 - p(\theta|e))\theta_0$. If the signal is that the leader is competent ($p(e) = 1$), then they choose $a = 1$ (support the regime). If the signal is that the leader is incompetent ($p(e) = 0$), then the action chosen is $a = 0$ (protest the regime).

[1] Proof is in Appendix B.

Heterogeneity

Citizens can be of two different types: those with high and low support for democratic norms. They differ in their beliefs about the possibility of an alternative leader. Let d be the share of citizens who show high support for democratic norms and believe there is a viable alternative leader with the payoff of 1. Citizens with high support for democratic norms form beliefs about the alternative and engage in political action in response. Citizens with low support for democratic norms believe that the current autocrat is the only viable option, and therefore, they do not form beliefs or engage in political action outside of the model presented.

With this extension, the observed political action for citizens with high support for democratic norms is a combination of political action for and against the current leader, as well as an alternative leader. Let $q(e)$ be the common beliefs about the alternative leader, $\theta = [0, 1]$ be the alternative leader type, and $a(e)$ – the citizens' political action towards the alternative leader.

The zero-profit conditions then are:

$a(e) = (1-d)[p(\theta|e)\theta_1 + (1-p(\theta|e))\theta_0] + d[p(\theta|e)\theta_1 + (1-p(\theta|e))\theta_0]$

$a(e) = (1-d)0 + d[q(\theta|e)\theta_1 + (1-q(\theta|e))\theta_0]$

The total political action of citizens observed empirically for citizens with high support for democratic norms is $A(e) = a(e) + a(e)$.

Comparative Statics and Empirical Implications

Based on the model, the exogenous variation is introduced in the elite signal e. In the absence of the elite signal, the uninformed citizens are only exposed to the leader's propaganda signal l, which is that the leader is competent, l=1. There is some variation in how convincing the propaganda is, which produces an equilibrium level of political action in support of the leader before the treatment. The level of political action observed at baseline is the equilibrium before the uninformed citizens receive the elite signal e.

After the exogenous variation is introduced, the elites send signal e, which in the case of the messages sent by MediaNet to the treatment group is e=0. The control group receives only news and is still subject to the leader's propaganda l=1. This is equivalent to receiving signal e=1 since they are only under the influence of the propaganda.

As the citizens in the treatment group receive e=0, they update their beliefs about the leader's type, $p(\theta_1|e(\theta_0)) = 0$. Using the zero-profit condition, their chosen political action is to withdraw the support for the current leader, $a(e) = p(\theta_1|e(\theta_0))\theta_1 + (1-p(\theta_1|e(\theta_0)))\theta_0 = 0 * \theta_1 + (1-0) * \theta_0 = 0$. The control group is under the influence of the propaganda, which implies that for them $p(\theta_1|e(\theta_1)) = 1$, and therefore the chosen action is to support the leader: $a(e) = p(\theta_1|e(\theta_1))\theta_1 + (1-p(\theta_1|e(\theta_1)))\theta_0 = 1 * \theta_1 + (1-1) * \theta_0 = 1$. Therefore, the treatment group withdraws the support for the current leader, while the control group remains supportive. The empirical implication is that as the treatment group learns of the leader's type and updates their beliefs, they decrease political action as compared to the control group who

remain supportive.

The extension to include heterogeneity implies that for the treated citizens with democratic preferences, the observed political action is the sum of $a(e) = 0$ and $a`(e) = 1$. They withdraw the support for the current leader and increase political action with respect to the alternative leader. Therefore, the observed political action may not be decreasing, $A(e) = a(e)+a(e) \neq 0$. For the treated citizens with autocratic preferences, the observed political action is only decreasing with the elite signal $e = 0$ since they do not consider an alternative leader. Therefore, $a(e) = 0$. Thus, the observed political action for those with autocratic preferences is less than for those with democratic preferences after the exposure to the elite signal, $a(e) < A(e)$. This is tested empirically.

A testable implication is that for those with democratic preferences, the political action that supports the current leader is decreasing in the elite messages, while the political action that seeks an alternative is increasing in the elite messages. Thus, it is expected to observe $a(e(\theta_0)) < a(e(\theta_0))$. This is tested using different political actions, such as reading what others publish online on political topics, which would be weakly supportive of a new leader, versus encouraging others to vote, which would be in support of the current leader.

4 Results

4.1 Demand

Table 2 shows no significant difference in demand for truthful information compared to news in the control group after participants experienced receiving truthful information during the study. While receiving fact-checked political information directly through Telegram lowers the search cost of acquiring information, the absence of an effect suggests that search cost is not the primary barrier to demand for truthful information.

Table 2: Difference in Wishing to Continue to Receive MediaNet Information in Treated and Control Groups

	(1)	(2)
	Continue Subscription	Continue Subscription
Assigned to treatment	0.037	0.037
	(0.033)	(0.034)
Demographic controls	No	Yes
Strata FE	Yes	Yes
N	719	719
Control group mean (Subscription to news)	0.69	0.69

Notes: The dependent variable is binary. Robust standard errors in parentheses.

Using two heterogeneity variables—trust in the information source and the value placed on independent media—I test for heterogeneous treatment effects on demand. Both variables have a strong positive effect on demand overall, but as shown in Table A.3, there is no differential treatment effect on demand across these groups.

One potential concern is whether respondents have difficulties using Telegram as the platform for the intervention. First, Telegram is a popular platform in this context and is generally perceived as more trustworthy than WhatsApp or other messaging apps, partly due to its widely advertised policy of not sharing user data in response to government requests. Second, I test for differences in demand for truthful political information between participants who reported using Telegram as a primary news source at baseline and those who rely on other platforms. Table A.4 shows no statistically significant differences in demand between these groups.

4.2 Primary Outcomes

Beliefs and Attitudes: Exposure to MediaNet's political information leads participants to revise their views on the low level of democracy in the country. On the 3-point Likert scale, receiving truthful political information reduces the average democracy assessment by 0.11 points from the control group mean of 1.83 — a 6% decrease (Figure 5 and Table A.5). Similarly, support for protesters increases by 0.08 points on the same scale, representing a 5% increase compared to the control group mean of 1.7. The effect sizes are comparable when using the 5-point Likert scale, though estimates are less precise due to the smaller sample size. Table A.6 shows that these effects remain robust after controlling for baseline covariates with imbalances.

The list experiment, designed to address potential fear in truthfully reporting support for protesters, estimates that 5.7% of respondents support the protesters, compared to 7.6% who openly express support when asked directly in the same subsample.

The experimenter demand treatment with a strong positive ask does not produce an effect on the beliefs about the level of democracy, but slightly increases the support for the protesters to 0.1 points on Lickert 1-3 scale. This corresponds to 0.04 standard deviations. For reference, Quidt et al 2018 find the average experimenter demand bounds to be modest at about 0.13 standard deviations. The experimenter demand in this context is considerably smaller.

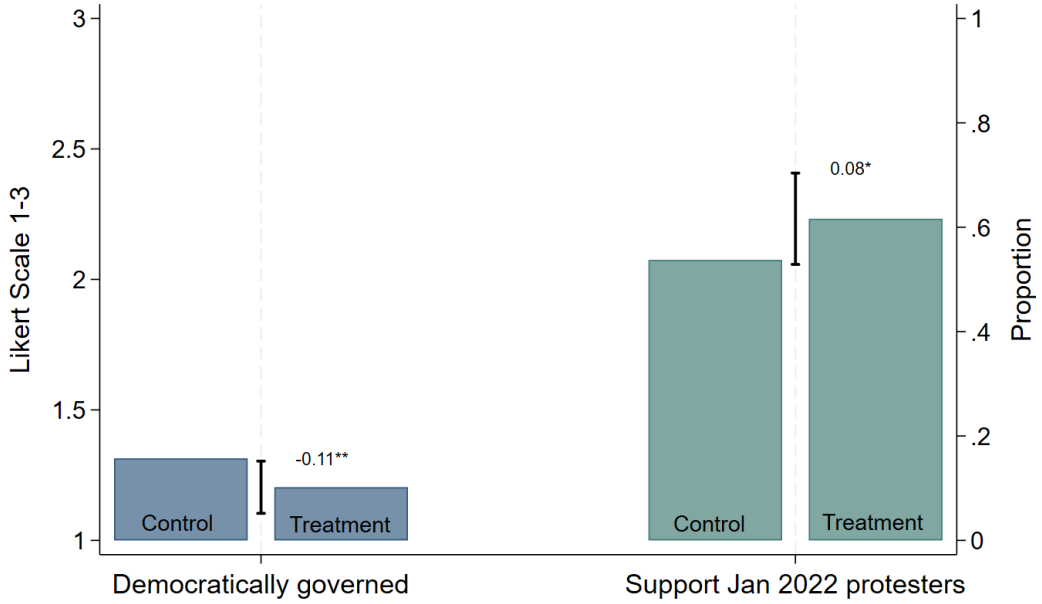


Figure 5: Treatment Effects on Political Beliefs and Attitudes

The bars represent mean values of the outcomes for treatment and control groups. The confidence intervals are for the estimates of the differences in means between treatment and control groups obtained using the regressions in Table A.5.

Political Behaviour: Regarding treatment effects on self-reported political behavior, I find a notable increase in political disengagement. Using a GLS-weighted Anderson index composed of two online and two offline behaviors, exposure to truthful information reduces political engagement by 15%, or 0.15 standard deviations (Figure 6). Table A.7 reports the treatment effect on the overall index and details the four individual behaviors it includes. The observed treatment effect primarily stems from two behaviors: searching for political information online and sending requests to government agencies. Both show similar disengagement magnitudes, decreasing by about 5 percentage points. While the reduction in searching for political information online may partly reflect a mechanical effect of receiving information through the treatment, the concurrent decline in engagement with government agencies via request submissions signals a broader trend of political disengagement. These effects remain robust after controlling for baseline imbalanced covariates (see Table A.8).

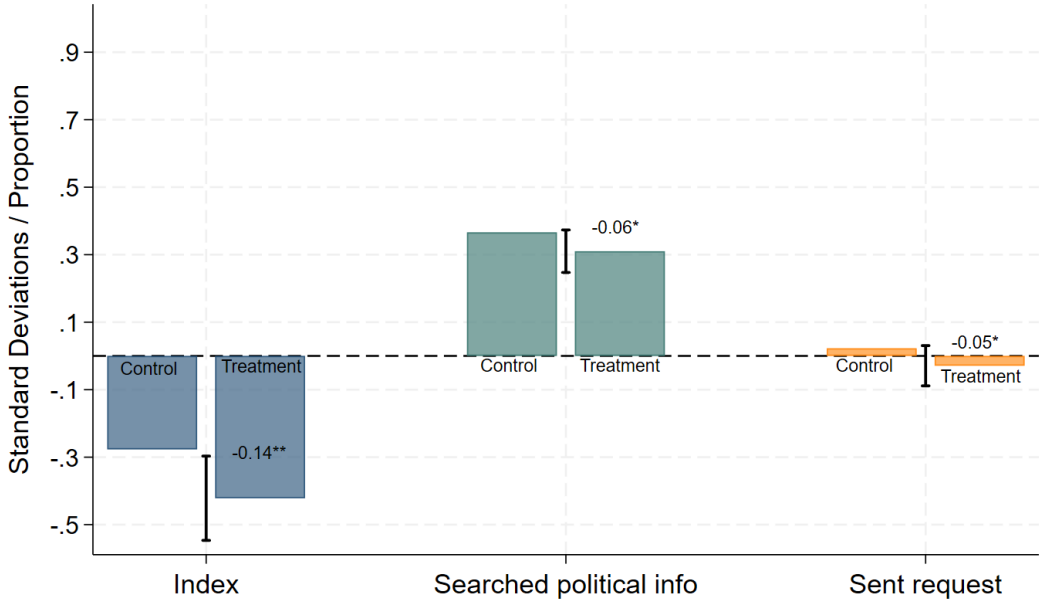


Figure 6: Treatment Effects on Political Behaviour

The bars represent mean values of the outcomes for treatment and control groups. The confidence intervals are for the estimates of the differences in means between treatment and control groups obtained using the regressions in Table A.7.

Donation to MediaNet: In addition to the primary outcomes of interest, I also measured in an incentivised way if the respondents wanted to donate to MediaNet. They were told that whatever amount they indicated, it would be subtracted from their final pay subject to the amount remaining to be paid to them. Table 3 shows that on average there are no treatment effects, but this masks an important heterogeneity revisited later.

Table 3: Treatment Effect on Donation to MediaNet

	(1)	(2)
	Donated	Donated
Assigned to treatment	0.051	0.055
	(0.036)	(0.036)
Demographic controls	Yes	Yes
Strata FE	Yes	Yes
Imbalanced variables	No	Yes
N	753	753
Control group mean	0.40	0.40

Notes: The outcome is binary. Robust standard errors in parentheses.

In summary, the demand for truthful political information remains low—even after individuals are exposed to it and can access it easily via Telegram—mirroring findings from Chen and Yang (2019). While exposure to accurate information leads people to update their beliefs about the lack of democratic governance and increases their support for protesters, it also results in greater political disengagement, both online and offline. There are no average differential effects on donation to MediaNet.

Using the theoretical framework, I interpret political engagement as a signal of support for the current leadership. Therefore, a decline in political engagement indicates reduced public support for the autocratic regime. In this context, the provision of truthful information leads citizens to update their beliefs and attitudes accurately, which in turn translates into diminished support for the autocratic government. However, it also implies that there are limited prospects for political change if citizens become disengaged.

4.3 Secondary Outcomes

Following Levy (2021), I measure polarization in perspective as the difference between each participant's reported difficulty agreeing with government officials and difficulty agreeing with the opposition. Since the treatment messages emphasize government manipulation, I expect participants in the treatment group to report increased difficulty agreeing with the government and decreased difficulty agreeing with the opposition. Consequently, polarization in perspective is expected to increase in the treatment group.

Similarly, polarization in ideas is measured as the difference for each participant between agreement with government information control and agreement with freedom of expression for the opposition. Since the treatment provides political information that is unavailable in mainstream media and helps fill the information gap, I expect agreement with government information control to decrease and agreement with freedom of expression to increase among the treatment group. Therefore, polarization in ideas is expected to decrease in the treatment group.

The results on polarization of perspectives and ideas further reinforce the primary findings. As expected, Figure 7 shows that exposure to truthful political information increases polarization in perspectives toward the state and opposition, as citizens update their attitudes based on new information about state manipulations. At the same time, polarization of ideas regarding information control decreases, reflecting a growing appreciation for independent information unavailable through mainstream media.

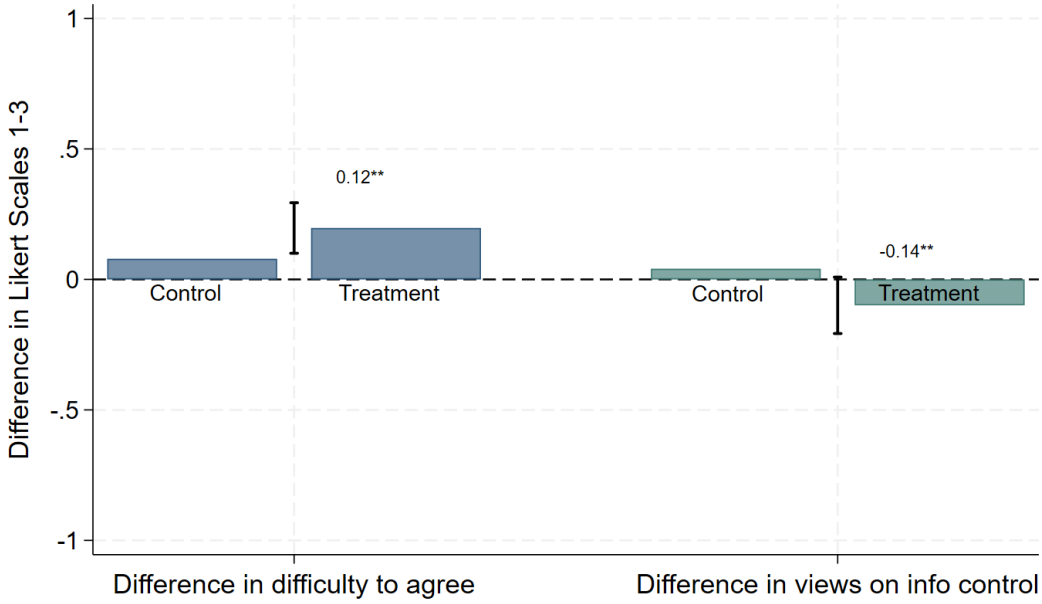


Figure 7: Treatment Effects on Polarization

The bars represent mean values of the outcomes for treatment and control groups. The confidence intervals are for the estimates of the differences in means between treatment and control groups obtained using the regressions in Table A.9.

4.4 Heterogeneous Effects

Beliefs, Attitudes and Behaviour: I use an index to divide the sample into two groups: individuals with high support for democratic norms and those with low support. The results show that belief updating and increased support for protesters are primarily driven by those with strong democratic values. As shown in Table 4, the treatment effect for perceived democratic governance among this group is a 0.15-point decrease on the 3-point Likert scale.

Table 4: Heterogeneous Treatment Effects on Political Beliefs and Attitudes

	(1) Democratically governed	(2) Support January 2022 protesters
Assigned to treatment	-0.058	0.052
	(0.083)	(0.069)
Treat X High support of democratic norms	-0.089	0.044
	(0.105)	(0.090)
Treat + Interaction	-0.147**	0.096
	(0.06)	(0.06)
Demographic controls	Yes	Yes
Strata FE and Baseline analog	Yes	Yes
Main effects	Yes	Yes
N	753	753
Control group mean	1.83	1.70

Notes: The dependent variables are on Lickert 1-3 scale. Robust standard errors in parentheses.

The observed average effect of political disengagement is driven primarily by individuals with low support for democratic norms. This group does not significantly update its beliefs or attitudes in response to truthful political information but instead withdraws from political engagement upon exposure. In contrast, individuals who strongly value democratic norms do revise their beliefs and attitudes—recognizing the country's low level of democracy and increasing support for protesters—but maintain their level of political engagement, remaining as active as those in the control group.

Table 5: Heterogeneous Treatment Effects on Political Beliefs and Attitudes

	(1) Index	(2) Searched pol info	(3) Read others	(4) Sent request	(5) Encourage voting
Assigned to treatment	-0.168*	-0.050	-0.042	-0.097**	-0.016
	(0.095)	(0.048)	(0.046)	(0.047)	(0.048)
Treat X High support of democratic norms	0.038	-0.012	0.048	0.081	-0.049
	(0.128)	(0.065)	(0.064)	(0.062)	(0.065)
Treat + Interaction	-0.130	-0.063	0.006	-0.017	-0.065
	(0.09)	(0.04)	(0.04)	(0.04)	(0.04)
Demographic controls and trustsource	Yes	Yes	Yes	Yes	Yes
Strata FE and Baseline analog	Yes	Yes	Yes	Yes	Yes
Main effects	Yes	Yes	Yes	Yes	Yes
N	753	753	753	753	753
Control group mean	0	0.52	0.42	0.35	0.30

Notes: The subgroup index variable is GLS-weighted. The outcome index variable is GLS-weighted and standardized over the control group. Other dependent variables are binary. Robust standard errors in parentheses.

Table 5 shows that the Anderson index of political engagement decreases by 0.17 standard deviations, or 16%, among individuals with low support for democratic norms. This effect is primarily driven by a decline in the behavior of sending requests to the government, which includes activities such as participating in polls, submitting complaints about corruption, and reporting on the performance of quasi-government agencies. These findings suggest that exposure to truthful information reduces engagement with the autocratic state—engagement that could otherwise contribute to improving the functioning of democratic institutions.

Using the theoretical framework, these findings align with the idea that individuals with low support for democratic norms exhibit lower support for the current leader but lack belief in viable alternatives. As a result, they disengage politically when exposed to truthful information. In contrast, individuals with high support for democratic norms remain politically engaged because they see the possibility of change. Their political activity reflects not only a withdrawal of support for the current leader but also an active search for democratic alternatives.

Polarization: I find that polarization is higher among individuals who show strong support for democratic norms (Table 6). Further analysis in Table A.10 reveals that this group, in particular, expresses a greater desire for reduced information control.

Table 6: Heterogeneous Treatment Effects on Polarization

	(1) Difference in difficulty to agree	(2) Difference in views on information control
Assigned to treatment	0.057	-0.050
	(0.061)	(0.078)
Treat X High support of democratic norms	0.099	-0.153
	(0.094)	(0.110)
Treat + Interaction	0.157**	-0.203***
	(0.07)	(0.08)
Main effects	Yes	Yes
Demographic controls	Yes	Yes
Strata FE	Yes	Yes
N	753	753
Control group mean	-0.05	-0.05

Notes: The dependent variable is on Lickert 1-3 scale.
Robust standard errors in parentheses.

Donation to MediaNet: Table 7 shows that the null effect observed on average hides that those who show high support for both democratic norms are more likely to donate to MediaNet by 8 and 10 percentage points, respectively.

Table 7: Treatment Effect on Donation to MediaNet

	(1) Donated	(2) Donated
Assigned to treatment	-0.041	-0.040
	(0.056)	(0.056)
Treat X High support of democratic norms	0.162**	0.168**
	(0.073)	(0.073)
Treat + Interaction	0.121**	0.128***
	(0.05)	(0.05)
Demographic controls	Yes	Yes
Strata FE	Yes	Yes
Main effects	Yes	Yes
Imbalanced variables	No	Yes
N	753	753
Control group mean	0.40	0.40

Notes: The outcome is binary. Robust standard errors in parentheses.

To summarize the heterogeneous treatment effects: individuals with high support for democratic norms update their beliefs about the country's low level of democracy and maintain their level of political engagement. They are more willing to donate to MediaNet than the control group. They also become more polarized in perspectives about the state and opposition and want to see less information control. Conversely, those with low support for democratic norms do not update their beliefs and instead disengage politically, both online and offline. This disengagement aligns with their decreased support for the current leader.

5 Conclusion

To conclude, this research finds that even when truthful information is readily available, people do not necessarily demand it. Individuals with high support for democratic norms update their beliefs negatively about the state of democracy and increase their support for protesters. They maintain their political engagement, reflecting their belief in an alternative to the regime. As expected, their views on the state and opposition become more polarized, and they express a stronger desire for reduced censorship. In contrast, those with low support for democratic norms do not update their beliefs about democracy or increase support for protesters but instead disengage politically. Interestingly, in this and many contexts, the share of individuals with low support for democratic norms exceeds that of those with strong democratic values. This means that the average effect of exposure to truthful information is one of political disen-

gagement. While such disengagement may signal declining support for the regime, it does not necessarily translate into a pathway for political change. Given the limited opportunities for political participation in authoritarian contexts, these conditions may increase the likelihood of future protests. However, further empirical research is needed to substantiate this possibility. Similarly, more work is required to understand how to foster political engagement in authoritarian settings to support peaceful and sustainable democratic transitions.

References

D. Acemoglu, C. G. Aksoy, C. Baysan, C. Molina, and G. Zeki. Misperceptions and demand for democracy under authoritarianism. Technical report, National Bureau of Economic Research, 2024.

O. Barrera, S. Guriev, E. Henry, and E. Zhuravskaya. Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of public economics*, 182:104123, 2020.

C. Baysan. Persistent polarizing effects of persuasion: Experimental evidence from turkey. *American Economic Review*, 112(11):3528–3546, 2022.

J. Bowles, K. Croke, H. Larreguy, S. Liu, and J. Marshall. Sustaining exposure to fact-checks: Misinformation discernment, media consumption, and its political implications. *American Political Science Review*, pages 1–24, 2023.

Y. Chen and D. Y. Yang. The impact of media censorship: 1984 or brave new world? *American Economic Review*, 109(6):2294–2332, 2019.

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

R. Enikolopov, M. Petrova, and E. Zhuravskaya. Media and political persuasion: Evidence from russia. *American economic review*, 101(7):3253–3285, 2011.

R. Enikolopov, M. Rochlitz, K. J. Schoors, and N. Zakharov. The effect of independent online media in an autocracy. *Available at SSRN 4131355*, 2022.

S. Guriev and D. Treisman. Informational autocrats. *Journal of economic perspectives*, 33(4): 100–127, 2019.

E. Henry, E. Zhuravskaya, and S. Guriev. Checking and sharing alt-facts. *American Economic Journal: Economic Policy*, 14(3):55–86, 2022.

G. King, J. Pan, and M. E. Roberts. How censorship in china allows government criticism but silences collective expression. *American political science Review*, 107(2):326–343, 2013.

R. Levy. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review*, 111(3):831–870, 2021.

M. Manacorda and A. Tesei. Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567, 2020.

E. Van Duyn and J. Collier. Priming and fake news: The effects of elite discourse on evaluations of news media. *Mass communication and society*, 22(1):29–48, 2019.

E. Zhuravskaya, M. Petrova, and R. Enikolopov. Political effects of the internet and social media. *Annual review of economics*, 12(1):415–438, 2020.

Appendix A

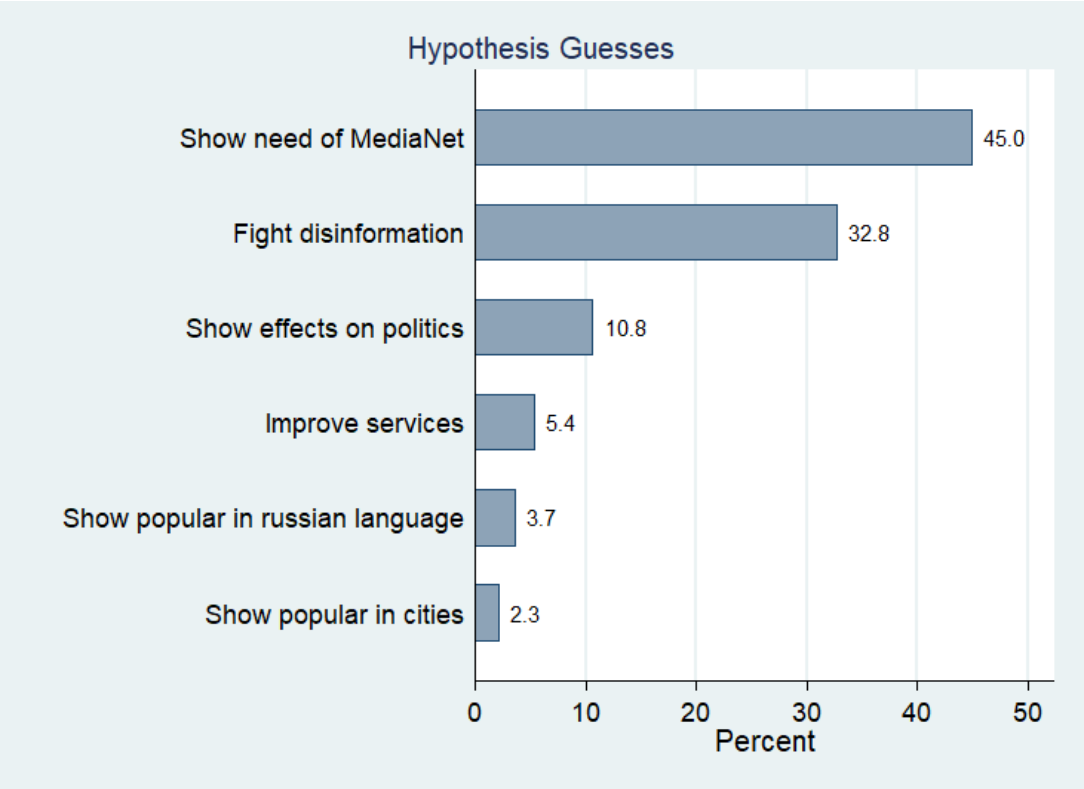


Figure A.1: Hypothesis guesses

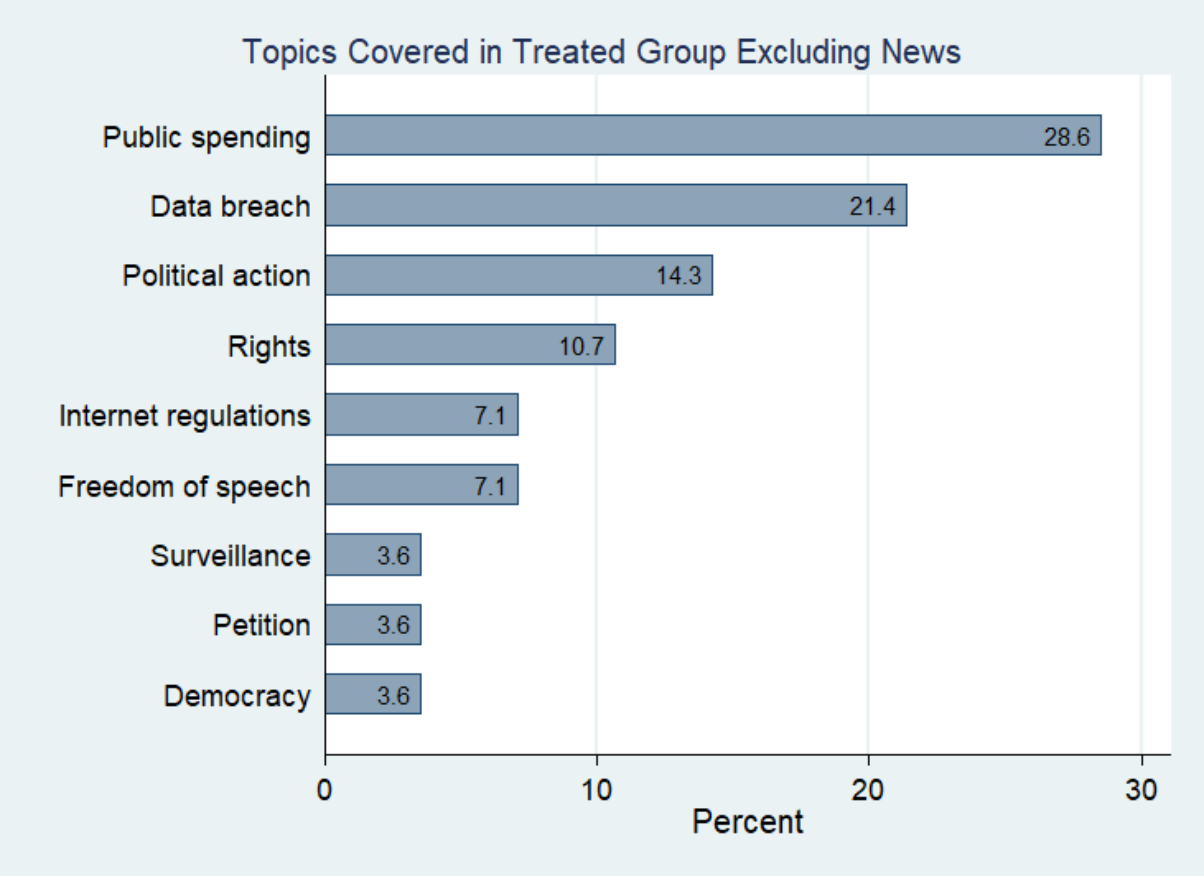


Figure A.2: Topics Covered in Treatment

Table A.1: Political knowledge at endline

	(1) Control group	(2) Treatment group	(3) Difference
Index of being correct on political facts	0.58	0.51	0.07
	(1.05)	(1.02)	(0.08)
Central Asian summit	0.75	0.68	0.07**
	(0.43)	(0.47)	(0.03)
Cyber attack investigations	0.18	0.15	0.03
	(0.38)	(0.35)	(0.03)
Reporters without borders	0.29	0.31	-0.01
	(0.46)	(0.46)	(0.03)
ProTenge	0.28	0.30	-0.02
	(0.45)	(0.46)	(0.03)
Businesses data privacy	0.36	0.33	0.03
	(0.48)	(0.47)	(0.03)
Know political figures	0.90	0.89	0.00
	(0.30)	(0.31)	(0.02)
Index of believing political facts	0.75	0.72	0.03
	(0.78)	(0.79)	(0.06)
Surveil gatherings	0.67	0.63	0.04
	(0.47)	(0.48)	(0.03)
Surveil residents	0.54	0.53	0.01
	(0.50)	(0.50)	(0.04)
Summit agreement	0.58	0.61	-0.03
	(0.49)	(0.49)	(0.04)
Pardon	0.71	0.67	0.03
	(0.46)	(0.47)	(0.03)
N	397	356	753

Notes: * .10 ** .05 *** .01. Index variables are GLS-weighted.

Table A.2: Compliance with Subscription to Study Channels in Treated and Control Groups

	(1) Subscribed	(2) Subscribed
Assigned to treatment	-0.010	-0.011
	(0.015)	(0.015)
Demographic controls	No	Yes
Strata FE	Yes	Yes
N	753	753
Control group mean	0.96	0.96

Notes: The dependent variable is binary. Robust standard errors in parentheses.

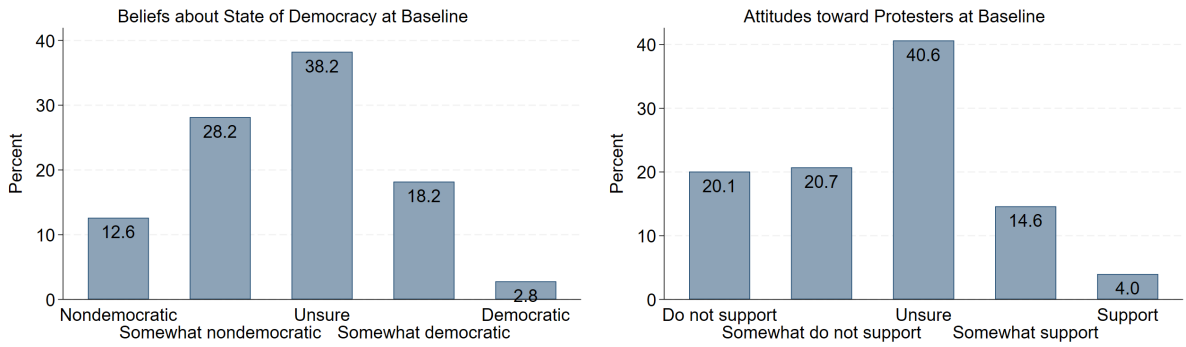


Figure A.3: Summary Statistics of Beliefs and Attitudes

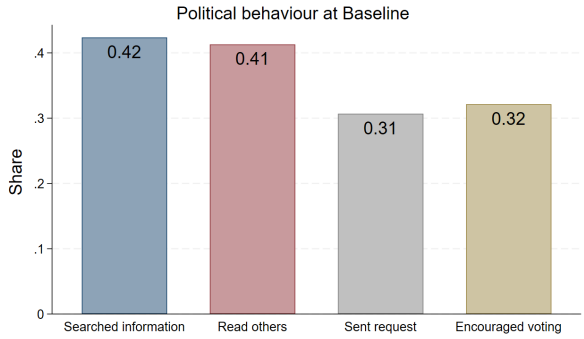


Figure A.4: Summary Statistics of Behaviour

Table A.3: Heterogeneous Difference in Subscription to MediaNet by Trust of Source and Need of Independent Media

	(1) Continue Subscription	(2) Continue Subscription	(3) Continue Subscription	(4) Continue Subscription
Assigned to treatment	0.062	0.061	0.080	0.078
	(0.079)	(0.079)	(0.059)	(0.058)
Trust source	0.360***	0.359***		
	(0.056)	(0.056)		
Treat X Trust Factcheck	-0.051	-0.051		
	(0.086)	(0.086)		
Prefer independent media			0.123**	0.127***
			(0.049)	(0.049)
Treat X Prefer Indep Media			-0.069	-0.069
			(0.071)	(0.071)
Treat + Interaction	0.011	0.010	0.011	0.009
	(0.03)	(0.03)	(0.04)	(0.04)
Demographic controls	No	Yes	Yes	Yes
Stratum FE and Baseline analog	Yes	Yes	Yes	Yes
N	753	753	753	753
Control group mean	0.69	0.69	0.69	0.69

*Notes:*The dependent variable is binary. Robust standard errors and joint standard errors in parentheses.

Table A.4: Check Demand Not Affected by Telegram Choice

	(1) Continue Subscription	(2) Continue Subscription	(3) Continue Subscription	(4) Continue Subscription
Assigned to treatment	0.022	0.024	-0.032	-0.028
	(0.041)	(0.041)	(0.066)	(0.067)
Top news source: Telegram channels	-0.022	-0.010		
	(0.050)	(0.050)		
Treat X Source Telegram	0.047	0.039		
	(0.070)	(0.070)		
Top news source: Other social media			-0.033	-0.040
			(0.052)	(0.052)
Treat X Source Other Platforms			0.094	0.087
			(0.077)	(0.077)
Treat + Interaction	0.069	0.063	0.063	0.060
	(0.06)	(0.06)	(0.04)	(0.04)
Demographic controls	Yes	Yes	Yes	Yes
Strata FE and Baseline analog	Yes	Yes	Yes	Yes
N	753	753	753	753
Control group mean	0.69	0.69	0.69	0.69

Notes: The dependent variable is binary. Robust standard errors and joint standard errors in parentheses.

Table A.5: Treatment Effects on Political Beliefs and Attitudes

	(1) Democratically governed	(2) Support January 2022 protesters
Assigned to treatment	-0.111**	0.079*
	(0.051)	(0.045)
Demographic controls	Yes	Yes
Strata FE and Baseline analog	Yes	Yes
N	753	753
Control group mean	1.83	1.70

Notes: The dependent variables are on Lickert 1-3 scale. Robust standard errors in parentheses.

Table A.6: Robustness for Treatment Effects on Political Beliefs and Attitudes

	(1) Democratically governed	(2) Support January 2022 protesters
Assigned to treatment	-0.108**	0.082*
	(0.052)	(0.045)
Demographic controls	Yes	Yes
Strata FE and Baseline analog	Yes	Yes
Imbalanced var	Yes	Yes
N	753	753
Control group mean	1.83	1.70

Notes: The outcomes are on Lickert 1-3 scale. Robust standard errors in parentheses. The regressions control for the imbalanced variables at baseline.

Table A.7: Treatment Effects on Political Behaviour

	(1) Index	(2) Index	(3) Searched pol info	(4) Read others	(5) Sent request	(6) Encourage voting
Assigned to treatment	-0.145**	-0.158**	-0.056*	-0.012	-0.052*	-0.043
	(0.064)	(0.063)	(0.032)	(0.032)	(0.030)	(0.032)
Double/debiased ML controls	No	Yes	No	No	No	No
N	753	753	753	753	753	753
Control group mean	0.00	0.00	0.52	0.42	0.35	0.30

Notes: The index variable is GLS-weighted and standardized over the control group. Other dependent variables are binary. Robust standard errors in parentheses. All regressions include demographic controls trust source Strata FE and baseline analogs.

Table A.8: Robustness for Treatment Effects on Political Behaviour

	(1) Index	(2) Searched pol info	(3) Read others	(4) Sent request	(5) Encourage voting
Assigned to treatment	-0.160**	-0.055*	-0.018	-0.058*	-0.048
	(0.064)	(0.032)	(0.033)	(0.030)	(0.033)
Demographic controls and trust of source	Yes	Yes	Yes	Yes	Yes
Strata FE and Baseline analog	Yes	Yes	Yes	Yes	Yes
Imbalanced var	Yes	Yes	Yes	Yes	Yes
N	753	753	753	753	753
Control group mean	0.00	0.52	0.42	0.35	0.30

Notes: The index variable is GLS-weighted and standardized over the control group. Other dependent variables are binary. Robust standard errors in parentheses.

Table A.9: Treatment Effect on Polarization

	(1) Difference in difficulty to agree Government - Opposition	(2) Difference in views on information control Control - Free
Assigned to treatment	0.118**	-0.140**
	(0.049)	(0.055)
Demographic controls	Yes	Yes
Strata FE	Yes	Yes
N	753	753
Control group mean	0.05	-0.05

Notes: The dependent variable is on Lickert 1-3 scale. Robust standard errors in parentheses.

Table A.10: Heterogeneous Treatment Effect on Polarization

	(1) Difficulty to agree with state	(2) Difficulty to agree with opposition	(3) Agree state should control info	(4) Agree opposition should express freely
Assigned to treatment	0.041	-0.017	-0.091	-0.041
	(0.066)	(0.070)	(0.076)	(0.069)
Treat X High support of democratic norms	0.031	-0.068	-0.132	0.021
	(0.093)	(0.095)	(0.105)	(0.090)
Treat + Interaction	0.072	-0.085	-0.223***	-0.021
	(0.06)	(0.06)	(0.07)	(0.06)
Main effects	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
N	753	753	753	753
Control group mean	2.32	2.27	2.53	2.58

Notes: The dependent variable is on Lickert 1-3 scale. Robust standard errors in parentheses.

Appendix B

Proof: To prove the separating equilibrium, it is necessary to show that $e(\theta_0)) = 0$ and $e(\theta_1)) = 1$. The elite signal is truthful. Suppose the leader is incompetent, but the elite signal states otherwise, $e(\theta_0)) > 0$. Then the deviation of the elite to the truthful signal, $e(\theta_0)) = 0$, is profitable because it weakly increases their expected payoff:

$a(0) = p(0)\theta_1 + (1 - p(\theta|e))\theta_0 \geq \theta_0$

The utility of the elite with the deviation is also higher:

$a(0)–c(0, \theta) \geq \theta_0 - c(0, \theta_0) > \theta_0 - c(e(\theta_0), \theta_0)$

Therefore, the elite will deviate to $e(\theta_0)) = 0$.

To show that $e(\theta_1)) = 1$, the individual rationality and incentive compatibility constraints need to be satisfied. The individual rationality constraint requires that if the leader is incompetent, then the elite choose to signal θ_0 over θ_1. This holds because the cost to the elite of reporting the truth is smaller than the cost of misreporting.

$\theta_1–c(e(\theta_1), \theta_0) \leq \theta_0–c(0, \theta_0) = \theta_0$

The incentive compatibility constraint requires that it is beneficial for the elite to report truthfully. The payoffs are such that if the leader is competent, then the benefit is higher, and the cost is lower:

$\theta_1–c(e(\theta_1), \theta_1) \geq \theta_0–c(0, \theta_1) = \theta_0$

Therefore, it is weakly more beneficial to report the truth.