The digital punishment puzzle: The prevention and escalation effect of digital repression

Medek, Paulina

The digital punishment puzzle:

The prevention and escalation effect of digital repression

Bachelor Thesis submitted in partial fulfilment of the requirements for the degree of in

International Relations and Organizations (Political Science), Bachelor of Science

by Paulina Medek

Faculty of Social and Behavioral Sciences

Institute of Political Science

BAP 17: The Politics of Authoritarian Rule

Supervisor: Dr. K. Köhler

Universiteit Leiden

Word count: 7995

Date of submission: 30 May 2022

ABSTRACT

The digitalisation of political expression impacted social movement organisation. The global proliferation of information and communication technology (ICT) digitalised social movements towards networked collective action. This fuelled the optimism about ICTs as liberalisation tool. Yet, authoritarian regimes showed increasing sophistication in digital repression which urges to investigate the link between digital repression (DR) and mobilisation. I use complementary insights from *connective action theory* and *disconnective action theory* to address the research question: *What is the effect of DR on mobilisation?* By arguing DR is multi-dimensional, the goal of this analysis is to establish the relationship between DR and mobilisation in authoritarian regimes. To this end, I conduct a systematic, global analysis. With an OLS regression, my large-N study analyses authoritarian regimes from 2000 until 2021. The findings are qualitatively supported with two case studies to address endogeneity concerns. On the one hand, I expect a prevention effect on mobilisation if DR is performed as long-term *information manipulation.* On the other hand, I hypothesise an escalation effect on mobilisation if DR is executed in form of a short-term *information vacuum*. Though the results are mixed, I conclude that there is initial support for both prevention and escalation effect of DR. I derive that DR is multi-dimensional and can prevent, but also escalate mobilisation.

Keywords: digital repression, digitalisation, connective action, mobilisation, punishment puzzle, repression mobilisation nexus

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

> 'We were very hopeful that now we had the [digital] tools to change the world. We were telling ourselves that you cannot really suppress people [who have the] internet, I think we were a bit naïve.' (Guest, 2022, para. 26)

The global proliferation of information and communication technology (ICT) digitalised social movements towards networked collective action (Bennett & Segerberg, 2014b). This interplay fuelled the optimism of activists – hoping for digital liberation from authoritarian repression. However, the negative course of the Arab Spring in Algeria in 2011 illustrates that authoritarian regimes can dominate the digital battleground – crushing previous hopes with digital repression (DR). The statement by Egyptian activist Abdelrahman Ayyash reflects this shift from optimism about digital liberalisation to pessimism about DR. Did authoritarian regimes win the digital battle? Did digitalisation introduce new power-asymmetries into the regime-opposition relationship?

Empirically, the evidence seems to be contradictory. As another example, the expansive implementation of internet filters by the regime repressed impactful protest in Kazakhstan since 2011. Yet, the Kazakh Bloody January (BJ) in 2022 showed a huge civic mobilisation wave amid the use of internet shutdown as additional form of DR. Why was DR successful since 2011, but ineffective in 2022? Academically, research does not provide a solid explanation for the effects of DR in authoritarian regimes. This is partly due to the tendency to understand DR as one-dimensional, either focused on the infrastructure, network, or application layer of the internet (Keremoğlu & Weidmann, 2020). A multi-dimensional understanding of DR is necessary to assess the effectiveness of DR and thereby its practical relevance for social mobilisation.

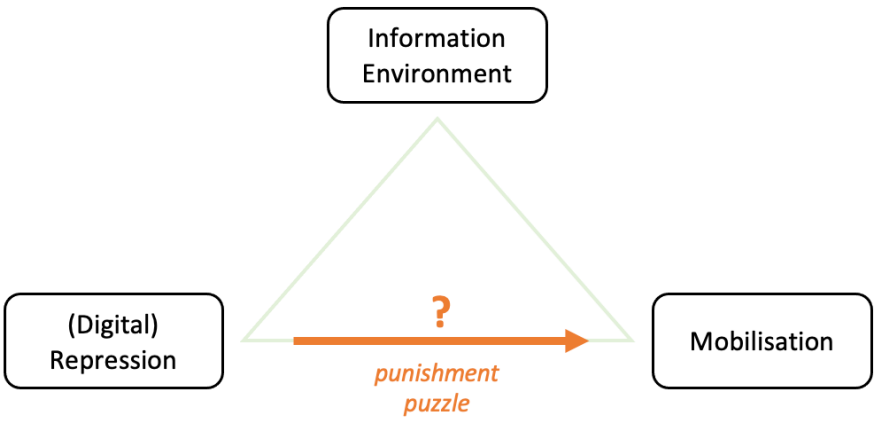
Acknowledging that theories of traditional repression (TR) partly explain the dynamics of DR, I use the strength and weaknesses of these efforts to derive my research question: *What is the effect of DR on mobilisation?* To investigate this puzzle, I use complementary insights from *connective action theory* and *disconnective action theory*. Based on five hypotheses, I expect a prevention and escalation effect of DR on physical mobilisation. The argument is that DR is multi-dimensional and additionally preventive, as opposed to one-dimensional and exclusively reactive. With an OLS regression, my large-N study analyses authoritarian regimes from 2000 until 2021. The findings are qualitatively supported with two case studies to address endogeneity concerns. Though the results are mixed, I conclude that there is initial support for both prevention and escalation effect of DR.

This thesis proceeds as follows. The next section reviews the existing literature in the field of regime repression and mobilisation. Thereafter, the theoretical framework conceptualises DR and mobilisation and establishes five hypotheses. This is followed by an explanation of the methods. The thesis proceeds with presenting and discussing the key findings and finishes with concluding remarks and recommendations for future research.

2 The state of the art: From traditional to digital repression

Most research analysing the relationship between (digital) repression and mobilisation in authoritarian regimes builds on theoretical insights of the *repression-mobilisation nexus* (Lichbach, 1987; Moore, 2000). Taking a rational choice perspective, scholars of this theory argue that there is strategic interaction between two unified actors, namely the regime and the opposition. The argument is that repression and mobilisation interact in a sequential action-reaction model. Accordingly, each actor uses information of the other actor's previous behaviour (action) for decision-making on its own behaviour (reaction). Each actor changes the information environment for the other actor. As such, the theoretical model is retrospective and assumes reactive decision-making (Moore, 2000). The derived hypotheses are two-fold: mobilisation influences repression, and repression influences mobilisation. Though there is agreement that mobilisation triggers repression (Rydzak, 2018, p. 31), scholars debate how repression influences mobilisation (Pierskalla, 2010). The inconclusive results are conceptualised by Davenport (2007a) as the *punishment puzzle* (Dickson, 2007). Applying this to DR, I summarised these theoretical insights in Figure 1.

Figure 1 *Relationship between digital repression, information environment and mobilisation*



I argue that the phenomenon of technological globalisation changes the very nature of mobilisation and, due to their interdependence, also of repression. Due to the interaction of digitalisation and

globalisation, anti-regime sentiments are no longer territorially bound (Dalmasso et al., 2018). In this vein, transnational protests highlight the significance of information control over physical control for authoritarian regimes (Michaelsen, 2017). This results in the *dictator's digital dilemma* defined as a struggle between political control and technological innovation (Howard et al., 2011). There is consensus that digitalisation changes the regime-opposition relationship (Dragu & Lupu, 2021; Feldstein, 2019; Michaelsen, 2018). This resonates with the changing strategies of government repression towards the use of ICTs as digital tools. The *digital repression gap* (Feldstein, 2021, p. 78) indicates that this applies especially to authoritarian regimes which show higher levels of DR than democracies. Yet, efforts to analyse DR as conceptually independent from traditional, physical repression are rare.

The few approaches analysing the effectiveness of ICTs as political tool independently are tainted by digital optimism (Diamond, 2010; Lynch, 2011). Advocating ICTs as *liberalisation* tool, these optimists focus on the benefit for one actor, the opposition group. Though this approach explains how ICTs enable collective action by unifying opposition groups, it overlooks the strategic benefit the digital age offers to authoritarian regimes. Therefore, more pessimistic scholars argue that ICTs can have a dual effect: liberalising and repressing the opposition (Dalmasso et al., 2018; Dragu & Lupu, 2021; Michaelsen, 2018; Moss, 2018). Understanding ICTs as *repression* tool, DR provides a novel form of regime power (King et al., 2013; Michaelsen, 2017; Rød & Weidmann, 2015). Interestingly, DR scholars continue to disagree on the effect direction of DR despite the theoretical foundation of the *repression-mobilisation nexus* (Earl et al., 2022; Keremoğlu & Weidmann, 2020).

There are several explanations for these inconclusive findings. To begin with, scholars do not distinguish between the different dimensions of DR. Additionally, the *repression-mobilisation nexus* advocates two inadequate assumptions: the unified actor and the reactive decision-making assumption. These shortcomings translate to three unaddressed aspects. Firstly, DR needs to be established as multi-dimensional concept to receive the necessary academic attention in distinction from TR. Secondly, collective action precedes the unification of opposition groups. As such, I argue that any actor group initially represents an ununified collective. Thirdly, strategic interaction can be reactive, but also preventive. In combination with the fact that previous scholarly efforts analysed DR and mobilisation in single case studies (Gohdes, 2015; Hassanpour, 2014), future research needs to account for these aspects in a systematic global analysis to establish generalisable trends. Accounting for these theoretical and methodological challenges, I pose the following research question:

What is the effect of digital repression (DR) on mobilisation?

3 Theoretical framework

3.1 Conceptualisations: Digital repression and mobilisation

Adopting Feldstein's (2021) definition, I define DR as "the use of information and communications technology (ICT) [...] to coerce, or to manipulate individuals or groups to deter specific activities or beliefs that challenge the state" (p. 26). Here, DR varies between information control (denial or reduction) and information channelling (obfuscation) (Earl et al., 2022; Rodzvilla, 2019). In addition, Keremoğlu and Weidmann (2020) state that DR covers three technical layers of the internet: infrastructure, network, and application.

Bringing these findings together, there are three main ways by which authoritarian regimes influence the digital space. Shutdowns influence the whole infrastructure by denial of access. Both internet and social media shutdown are intended to disrupt the network communication platform and as such, the overall internet infrastructure. Internet filtering demolishes specific networks to reduce content. Misinformation campaigns are staged on social media as application platforms to obfuscate the information environment. Whereas shutdowns are overt forms of repression, meaning they are attributable and provocative actions by the government, internet filtering and misinformation campaigns are covert forms. All DR forms are often employed simultaneously (Earl et al., 2022). Since this typology accounts for varying technical layers, this three-dimensional structure sets the foundation for a comprehensive conceptualisation of DR.

Mobilisation is defined as violent and/or non-violent collective actions and methods of political struggle undertaken by groups of non-state actors. Since scholars theorise that DR influences both violent and non-violent forms of repression, distinguishing between these forms of mobilisation is not necessary (Postmes & Brunsting, 2002; Stephan & Chenoweth, 2008). Mobilisation is inherently dangerous for authoritarian regime stability (Gerschewski, 2013). This includes any form of mobilisation – albeit politically, economically, and/or culturally motivated. I therefore define mobilisation neutrally as any form of civil unrest.

As collective action, mobilisation depends on the effective coordination of individuals with diverse identities. Scholars emphasise that psychological and structural factors play a role when conducting collective action (Pearlman, 2012; Polletta & Jasper, 2001). I argue that neither psychological, nor structural factors alone are sufficient to explain the success of *sustained* mobilisation. Yet, I underline that digitally enabled collective action eases *initial* mobilisation. By enabling the coordination of large

groups, the threshold for individual participation to collective action decreases (Hassanpour, 2014). In effect, the incentive to participate for a common cause is higher. In addition, the personalised communication in digital space replaces the necessity to form a social identity before mobilisation (Bennett & Segerberg, 2014a). In short, structural factors outweigh psychological factors to coordinate opposition groups for initial mobilisation.

3.2 Resource mobilisation theory (RMT) in a digital environment

The importance of these structural factors for social mobilisation is described by *resource mobilisation theory* (RMT). The argument is that mobilisation depends on a combination of different forms of structural resources which can be categorised as organisational, financial, and informational (Chenoweth & Ulfelder, 2017; McCarthy & Zald, 1977). Since this thesis analyses the dynamics of DR, I apply RMT to a digital environment.

Organisationally, ICT-enabled networks replace political organisations (Bennett & Segerberg, 2014b). Communication networks establish channels and thereby coordinate the interaction between a large number of activists. Financially, the use of communication networks is low-cost. Considering the global increase of internet penetration, people participate with low financial effort at low risk (Lupia & Sin, 2003). Online activity on social media enables noticeability across various interest groups in an anonymised manner. Hence, the use of ICTs delivers the organisational and financial, but not the informational resources. In short, RMT underlines the centrality of informational resources for digitally enabled mobilisation. I use this insight to address the two inadequate assumptions of the *repression-mobilisation nexus.* To recap, I argue that the understanding of unified and reactive decision-making is too narrow. As such, the following section understands opposition groups in their collective nature to deduce the preventive effect of DR.

3.3 Digital-physical mediation: The organisation dynamics of digitally incentivised protest

3.3.1 Connective action theory

According to *connective action theory,* ICTs facilitate collective action and thereby serve as liberalisation tool. The argument is that communication networks provide an information and coordination channel (Diamond, 2010; Enikolopov et al., 2020; Lynch, 2011). In effect, information cascades can be established which activate attentive participants. Ultimately, the theory explains the unification of diverse interest groups via communication networks. This gives valuable insights on the effect of external events, i.e., DR.

It is important to underline the strategic shift from defensive to offensive means of DR within authoritarian regimes (Moss, 2018). To reduce regime-threatening content, non-democracies use filtering to prevent the emergence of potential information and coordination channels. This indirectly affects mobilisation. China's *Great Firewall* is an illustrative example. By providing limited access to the internet, China designs a highly censored intranet. Thereby, the demand for politically sensitive information decreases in the long-term (Y. Chen & Yang, 2019; Rodzvilla, 2019). By changing the information environment for all Chinese citizens, mobilising groups struggle to convince others to support their cause. And even if successful, these groups cannot coordinate and thus mobilise across the communication network. Based on this, I deduce the following hypothesis:

> *Hypothesis 1*: Internet filtering decreases mobilisation. This is the *prevention effect* of DR.

Another way to indirectly impede the information consumption of citizens is captured by the *active engagement strategy* of authoritarian regimes (Munger et al., 2019; Sanovich et al., 2018). Instead of redefining the scope of internet accessibility as via internet filtering, regimes actively create content on established communication networks. They manipulate the information environment via state-linked or official state accounts. This strategy averts illegitimacy claims by the public, since the overall access to internet platforms is preserved. In analogy to internet filtering, I establish the following hypothesis:

> *Hypothesis 2a*: Government disinformation campaigns decrease mobilisation. This is the *prevention effect* of DR.

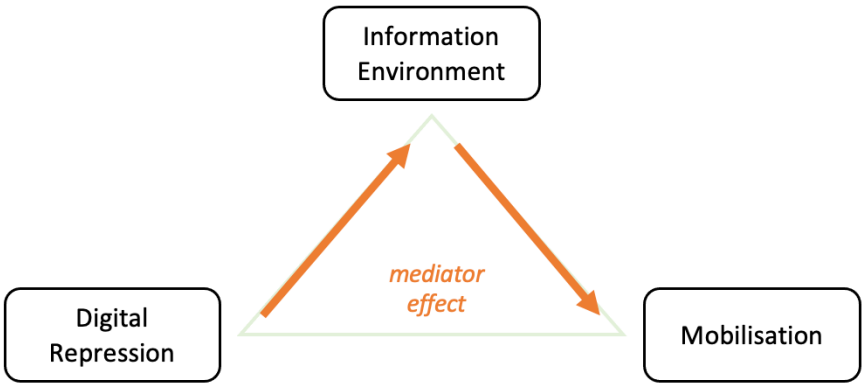
Yet, this *active engagement strategy* is conditional. If citizens' social media penetration rates are higher, they are more susceptible to online disinformation campaigns intended to influence their decision to mobilise (Enikolopov et al., 2020; Feldstein, 2021). In this case, disinformation campaigns are expected to be more successful meaning decreasing mobilisation. The opposite applies to low social media penetration. As such, the following hypothesis is deduced:

> *Hypothesis 2b*: The success of government disinformation campaigns is conditional on people's social media penetration. This is the *interaction effect* of disinformation campaigns.

The predicted prevention effects are long-term processes. Internet filtering and misinformation campaigns by autocracies are aimed at reducing or obfuscating the information environment of

citizens to change their belief system (Y. Chen & Yang, 2019). These manipulation strategies weaken information and coordination channels of mobilisation groups (Steinert-Threlkeld et al., 2015; Tufekci & Wilson, 2012; Valenzuela, 2014). As visualised in Figure 1.1, this *procedural* understanding predicts an *indirect* effect of DR by impacting the process of organisation. This indirect, mediator effect represents a long-term control mechanism, and thus preventive repression tool.

Figure 1.1 *Indirect effect of internet filtering and online disinformation (prevention effect)*



3.3.2 Disconnective action theory

Disconnective action theory explains short-term effects within an information vacuum (Rydzak et al., 2020). Following this line of argument, the sudden disappearance of previously established information and coordination channels has different effects on mobilisation. In the moment of *just-in-time shutdowns* (Keremoğlu & Weidmann, 2020, p. 1642), the movement is used to and has already benefitted from the organisational and financial resources of communication networks. Therefore, shutdowns create civil grievances due to the loss of established connectivity (Keremoğlu & Weidmann, 2020). Furthermore, a dispersion effect can be witnessed (Hassanpour, 2017). To explain, the sudden disconnect promotes local, physical mobilisation. The disappearance of previous resources triggers two related processes: provoking citizens and triggering a spontaneous change in strategies. As such, the socio-economic resources of ICTs are replaced with human resources. In combination with a social-spill-over to people who are detached from the cause (Y. Chen & Yang, 2019), this dispersion acts as a catalyst for mobilisation (Hassanpour, 2014; Hobbs & Roberts, 2018). In line with this argument, I formulate the hypothesis:

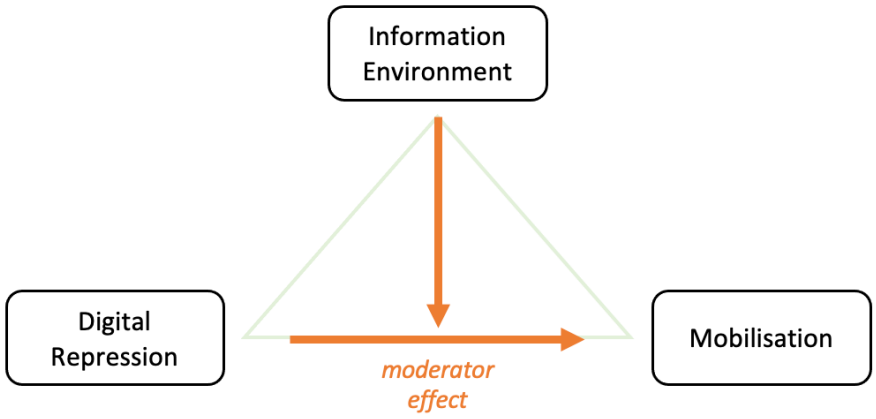
Hypothesis 3: Internet shutdown increases mobilisation. This is the *escalation effect* of DR.

Shutdowns vary in connection quality (Rydzak et al., 2020). Whereas an internet shutdown is complete, social media shutdowns are partial. Since social media platforms are the main tool to organise ICT-enabled mobilisation (Howard, 2010), the effect of social media shutdowns is expected to be smaller, but similar. To explain, partial and full *just-in-time shutdowns* (Keremoğlu & Weidmann, 2020, p. 1692) strategically destroy communication channels in the short-term.

> *Hypothesis 4*: Social media shutdown increases mobilisation. This is the *escalation effect* of DR.

To summarise, under certain conditions the use of ICT as DR has an additional, *substantive* dimension. This is addressed by scholars arguing for the dual effect of ICTs which states that ICT represses, but also triggers mobilisation (Dalmasso et al., 2018; Dragu & Lupu, 2021; Michaelsen, 2018; Moss, 2018). To explain the latter, sudden information denial via internet shutdowns impacts the organisation of protest. In this case, ICT has a *direct* effect on the relationship between DR and mobilisation as visualised in Figure 1.2. This form of DR sends a provocative signal which triggers physical dispersion (Rydzak, 2018; Hassanpour, 2014). This signal creates protest momentum and thereby reverses the indirect, negative relationship between long-term DR and mobilisation on a short-term.

Figure 1.2 *Direct effect of internet and social media shutdown (escalation effect)*



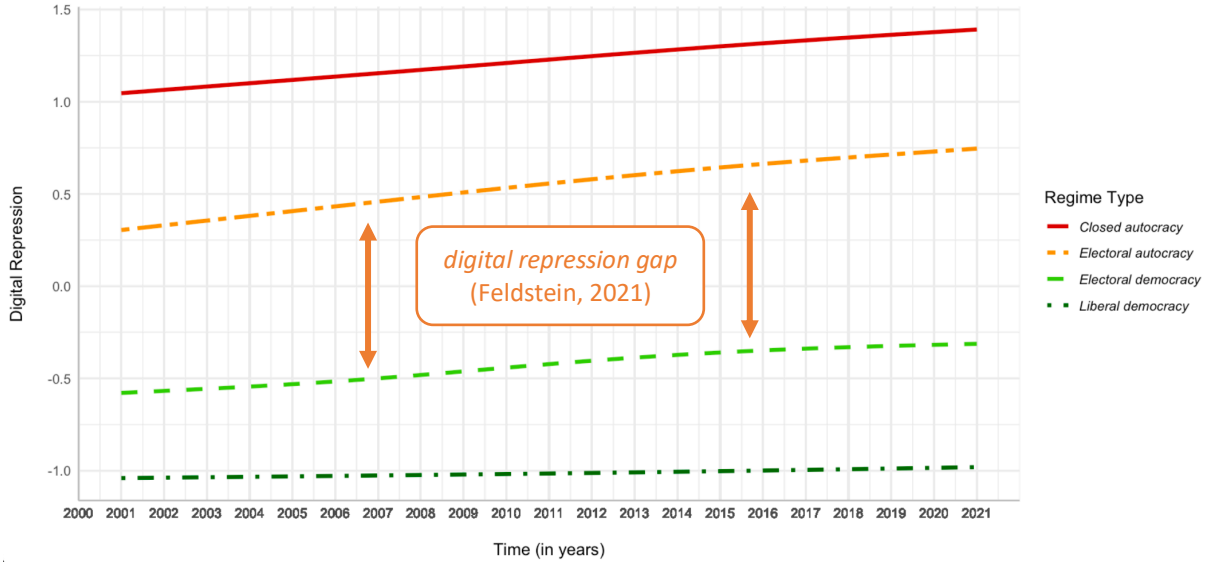
4 Methodology

4.1 Level of analysis and sample selection

The subset of authoritarian countries provides the units of analysis to analyse the link between DR and mobilisation. The level of analysis is domestic state structures, i.e., the DR apparatus. The justification is that high levels of protest can be related to high levels of DR. Whereas authoritarian states tend to supress opposition groups (Gerschewski, 2013), democracies are less inclined to repress

their citizens (Davenport, 2007b). Figure 2 uses the aggregated DR index by Feldstein (2022) to visualises this *digital repression gap* (Feldstein, 2021, p. 79). It shows that democracies have on average lower DR scores compared to non-democracies. In addition, protest numbers are generally higher in democracies due to the relatively high freedom of expression (Goldstone, 2004). The relationship between repression and mobilisation in democracies risks to be confounded.

Figure 2 *Digital repression for regimes of the world (RoW) from 2001 till 2021*



The link between DR and mobilisation is expected to be detectable in non-democracies, meaning closed and electoral autocracies. These two regime types provide the necessary variation of DR to analyse the relationship. The sub-selection is based on the *Regimes of the World* (RoW) variable by the *Varieties of Democracy (V-Dem) Project* (v12) (0: *Closed Autocracy*, 1: *Electoral Autocracy*, 2: *Electoral Democracy*, 3: *Liberal Democracy*) (Coppedge et al., 2022b). I include cases which categorise as *Closed Autocracy* or *Electoral Autocracy*.

4.2 Dependent variable: Mobilisation

I use the *Mobilization for Democracy* variable by V-Dem to operationalise mobilisation. V-Dem's more general alternative, *Mass Mobilization*, covers state-orchestrated events as well. Thus, I decide for the more conservative alternative to focus on pro-democratic, and as such exclusively political, mobilisation. Following the V-Dem data collection scheme, this indicator for pro-democratic mass mobilisation is derived from expert survey data. It includes explicit democratic support by advancing democracy and implicit democratic support by supporting civil liberties. This mobilisation can take place as demonstration, strike, or sit-in. The ordinal variable is measured on a scale from 0 (*There have been virtually no events*) to 4 (*There have been many large-scale and small-scale events*). After being

standardised by the total mean for all country-years in the V-Dem sample covering a time range from 1900 till 2021, the mobilisation variable ranges from -5 to 5 (Coppedge et al., 2022b). The V-Dem indicator increases the measurement validity of mobilisation by accounting for frequency and size of mobilisation.

4.3 Independent Variable: Digital repression

To operationalise DR and account for its multi-dimensional nature, four ordinal variables have been used to indicate the frequency and scope of internet shutdown, social media shutdown, internet filters and domestic government disinformation campaigns. The *Digital Society Survey* (DSS) data set by the *Digital Society Project* follows the data collection scheme of V-Dem and accounts for the three technical layers of DR. Covering each layer, I overcome the one-dimensional approach of previous research and compare each dimension with each other. All four variables are measured based on expert survey data. The scale ranges from 0 (*Extremely often*) to 4 (*Never, or almost never*). After being standardised by the total mean of the DSS sample, the variable ranges from -5 to 5. For the ease of interpretation, I reversed the coding. As such, high values indicate a high frequency of DR events within a particular country and year. The dataset covers the time range from 2000 to 2021 (Mechkova, et al., 2022).

Regarding the infrastructure layer, I use *Government Internet shut down in practice* and *Government social media shut down in practice*. For regimes' intrusion on the network layer, internet filters are commonly used indicators (Earl et al., 2022). I use the variable *Government Internet filtering in practice*. To operationalise the application layer and account for the active engagement strategy by authoritarian regimes, disinformation campaigns are indicated by the variable *Government dissemination of false information domestic*.

4.3.1 Interaction effect: Disinformation campaigns and social media penetration

Social media penetration influences the incidence and size of protests (Enikolopov et al., 2020). In addition, I expect that the effect of domestic disinformation campaigns depends on social media penetration (Feldstein, 2021, p. 68). This results in an interaction effect. Social media penetration is defined as the *Average people's use of social media to organize offline action*. The V-Dem data set offers a 5-point scale indicator from 0 (*Never or almost never*) to 4 (*Regularly*) (Coppedge et al., 2022). For ease of interpretation, 0 describes the lowest value in the sample. To clarify, the analogous argument of an interaction effect with social media shutdown cannot be made. As elaborated in the theoretical framework (section 3.3.2), the sudden digital disruption acts as trigger for the

intensification of physical mobilisation (dispersion effect). The presence of this physical mobilisation trigger is independent of peoples' *average* social media activity, since what matters is that there are *few* people who are susceptible to the disruption to initiate the protest mobilisation. The descriptive statistics of the main variables have been summarised in Table 1.

Table 1 *Descriptive statistics of main variables*

Variables	*N*	*Min*	*Max*	*M*	*SD*	*Variance*
Dependent variable						
Mobilisation	1835	-3.12	4.11	-0.08	1.48	2.18
Independent variables						
Internet Filtering (lag)	1747	-2.15	3.91	0.64	1.36	1.85
Disinformation (lag)	1747	-1.60	3.48	0.77	1.01	1.02
Internet Shutdown	1747	-1.93	4.15	0.25	1.23	1.51
Social Media Shutdown	1747	-1.88	4.01	0.22	1.30	1.68
Social Media Penetration (lag)	1747	0.00	7.28	3.12	1.29	1.67

4.4 Control variables

Two sets of control variables have been included: (1) political controls potentially confounding the analysed relationships, and (2) basic digital, economic, and demographic controls. Politically, there are five controls. The political stability index measures perceptions of the likelihood that the government will be destabilised by unconstitutional or violent means (World Bank, 2022c). Being reversed, the index indicates political instability indicating the presence of politically motivated violence and terrorism. It is included as proxy for the regime's perception of illegal and/or violent threats (Earl et al., 2022). According to the *law of coercive responsiveness*, threats against the regime trigger repressive actions (Davenport, 2007a). This also applies to the political polarisation among citizens which cover the legal and/or non-violent tension within the political system (Coppedge et al., 2022). Both political instability and polarisation increase the overall level of mobilisation (Hollyer et al., 2015; Kleiner, 2018).

By including the reversed *Polity5 score* as indicator for the institutional degree of authoritarianism, I acknowledge that more authoritarian regimes use more repression and discourage mobilisation in its organisational roots (Deibert et al., 2010). As socio-political norm, civil liberties have been included to account for the normative degree of authoritarianism. The influence of civil liberties on DR depends on the form of repression. High levels of civil rights increase the use of covert forms of DR over overt

forms of DR. In addition, more civil liberties encourage low-intensity, non-violent mobilisation but thereby discourage high-intensity, violent mobilisation on the streets (Chenoweth & Ulfelder, 2017; Earl et al., 2022).

Physical repression is an indicator for traditional repression (TR). TR and DR are related; DR substitutes, reinforces, or complements TR. In all cases, the use of TR incentivises the use of DR (Keremoğlu & Weidmann, 2020). In addition, TR has an independent effect on the level of mobilisation (Lichbach, 1987; Pierskalla, 2010). The expected effects of the political confounds have been summarised in Table 2. The internet penetration rate, GDP per capita as proxy for economic development, population size, and population density control for digital, economic, and demographic differences between the authoritarian states (World Bank, 2022). All data for GDP per capita, population size, and population density has been log transformed to treat highly skewed values.

Table 2 *Expected effects of political confounds on digital repression and mobilisation*

Variable	*Theoretical justification*	*Effect on DR*	*Effect on mobilisation*	*Data set*
Political Instability	Davenport (2007a); Hollyer et al. (2015)	+	+	World Bank (2022c)
Polarisation	Davenport (2007a); Kleiner (2018)	+	+	V-Dem Coppedge et al., 2022a)
Polity5 (reversed)	Davenport (2007b); Deibert et al. (2010)	+	-	Centre for Systemic Peace (2020)
Civil Liberties	Earl et al. (2022); Chenoweth & Ulfelder (2017)	+/-	+/-	Freedom House (2022)
Physical Repression	Keremoğlu & Weidmann (2020), Pierskalla (2010) & Lichbach (1987)	+	-	Political Terror Scale (Gibney et al., 2021)

4.5 Statistical model

The goal of this analysis is to establish the relationship between DR and mobilisation in authoritarian regimes. To generalise the results, a systematic, global analysis is conducted. For this reason, a multiple linear regression is well suited because it is sensitive to small effects and enables comparative conclusions for the different dimensions of DR. All variables are treated as interval-ratio. After

controlling with robust standard errors for autocorrelation and heteroskedasticity, all assumptions are met by the statistical model. For further explanation, see *Appendix*. To account for endogeneity concerns, all predictor variables, except internet and social media shutdowns, are lagged by one year (Xu, 2020). The endogeneity problem for the shutdown predictors with short-term effects is addressed in a qualitative vignette illustrating the causal mechanism. For comparability across indicators, the predictors are z-standardised which means that zero values describe the average of the sample. In exception, the zero values of social media penetration indicate the lowest value. The final statistical model (including all control variables) can be mathematically described as follows:

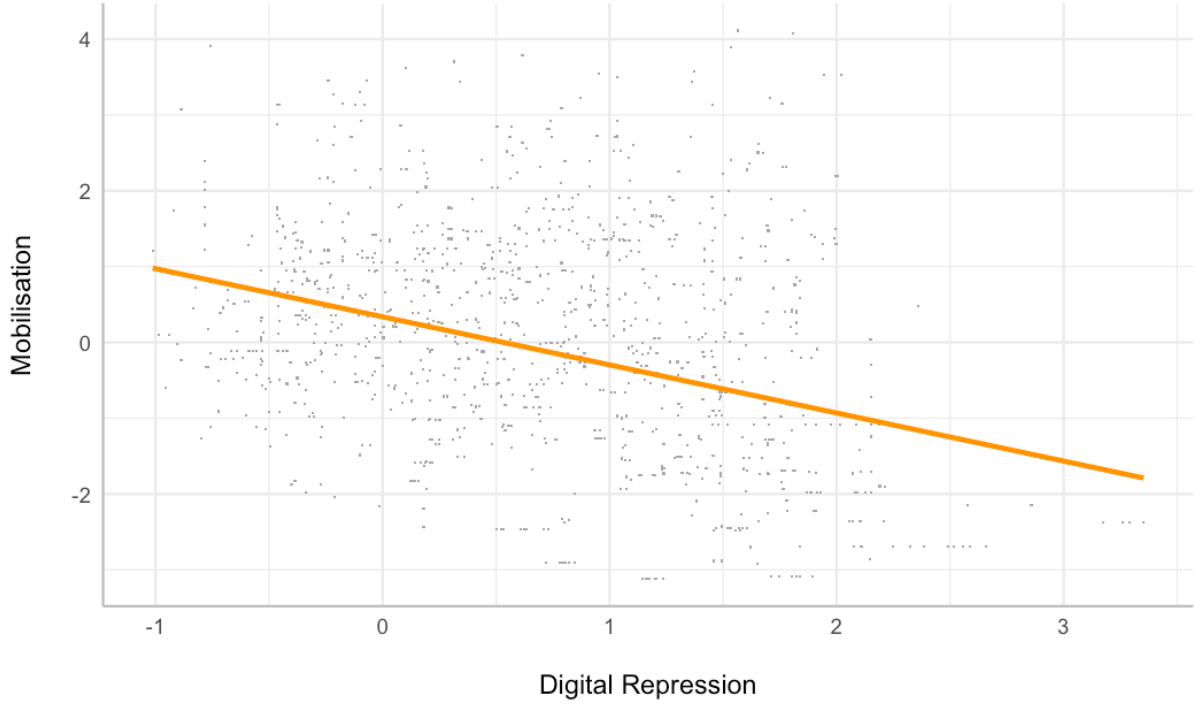
$$\begin{aligned} Mobilisation_i = {} & b_0 + b_1 * InternetFiltering_i + b_2 * Disinformation_i + b_3 * InternetShutdown_i + \\ & b_4 * SocialMediaShutdown_i + b_5 * SocialMediaPenetration_i + \\ & b_6 SocialMediaPenetration * Disinformation_i + b_7 * PoliticalStability_i + \\ & b_8 * Polarisation_i + b_8 * Polity5_i + b_8 * CivilLiberties_i + b_8 * PhysicalRepression_i + \\ & b_9 * InternetPenetration_i + b_{10} * log(GDP)_i + b_{11} * log(PopulationSize)_i + \\ & b_{12} * log(PopulationDensity)_i + \varepsilon_i \end{aligned}$$

5 Results

5.1 Quantitative analysis of digital repression

The multiple linear regression analysis is conducted hierarchically in four steps. Before conducting the regression, the relationship between DR and mobilisation is visualised. Figure 3 presents the scatterplot of mobilisation against the aggregated index of DR as established by Feldstein (2021). Without accounting for different dimensions, DR shows a negative relationship with mobilisation. However, as elaborated in the theoretical framework, different forms of DR are employed simultaneously by authoritarian regimes and are expected to have different effects. Using an aggregated index in Figure 3, the different effect directions are not visible. Bringing these findings together, I argue that the visualisation provides initial support for the hypotheses, but simultaneously highlights the need to disentangle the different dimensions of DR.

Figure 3 *Scatterplot of mobilisation against digital repression*



In the following analysis, Model 1 accounts for the reciprocal effects of the four forms of DR. I hereby test the effect of DR on mobilisation isolated from other confounding effects. To fully explore the expected relationship, Model 2 includes the interaction effect. Models 3 and 4 account for potential confounding effects. Model 3 controls for political factors. By including digital, economic, and demographic factors to the analysis, Model 4 is the final model. The results are presented in Table 3.

Table 3 *Multiple linear regression model of mobilisation with standardised predictors*

	Model 1 (Baseline)	**Model 2 (Interaction)**	**Model 3 (Controls 1)**	**Model 4 (Controls 2)**
(Constant)	-0.001	-1.135***	-0.926***	-0.858***
	(0.073)	(0.124)	(0.138)	(0.149)
Internet Filtering (lag)	-0.650***	-0.571***	-0.310***	-0.312***
	(0.113)	(0.085)	(0.085)	(0.098)
Disinformation (lag)	0.226*	-0.234+	-0.198	-0.128
	(0.100)	(0.138)	(0.140)	(0.127)
Internet Shutdown	0.354**	0.344***	0.352***	0.354***
	(0.128)	(0.095)	(0.101)	(0.100)
Social Media Shutdown	-0.237+	-0.100	-0.217*	-0.262**
	(0.132)	(0.100)	(0.105)	(0.101)
Social Media Penetration (lag)		1.224***	0.996***	0.922***
		(0.116)	(0.133)	(0.154)
Disinformation x Social Media Penetration		0.347**	0.227+	0.185
		(0.106)	(0.122)	(0.114)
Political Instability (lag)			0.167*	0.206**
			(0.069)	(0.077)
Polarisation (lag)			0.151*	0.230***
			(0.068)	(0.065)
Regime Type (Polity5) (lag)			-0.133+	-0.118+
			(0.068)	(0.065)
Civil Liberties (FH) (lag)			-0.048	-0.045
			(0.067)	(0.066)
Physical Repression (PTS) (lag)			0.061	-0.029
			(0.059)	(0.060)
Internet Penetration (lag)				-0.021
				(0.058)
GDP per capita$_{log}$ (lag)				0.013
				(0.075)
Population Size$_{log}$ (lag)				0.086
				(0.072)
Population Density$_{log}$ (lag)				-0.094+
				(0.055)
R^2	0.256	0.494	0.585	0.597
Adj. R^2	0.255	0.492	0.582	0.592
N	1747	1747	1389	1296

Note: The OLS regression coefficients with robust standard errors between brackets.
***p < 0.001, **p < 0.01, *p < 0.05, +p < 0.1

5.1.1 Baseline model: The dimensions of digital repression

Controlling for the reciprocal effects between the forms of DR, Model 1 provides mixed results. The baseline model (Model 1) shows that there is a significant negative effect of internet filtering, $b_1 = -0.650$, $p < 0.001$, on mobilisation and a significant positive effect of internet shutdown, $b_3 = 0.354$, $p < 0.01$. This means that the more often the regime reduces political content on the internet by filtering, the lower the intensity, meaning frequency and size, of mobilisation. With an increase by one standard deviation on internet filtering, mobilisation decreases on average by 0.650 standard deviations. The opposite applies to internet shutdown. The more often the authoritarian government restricts domestic access to the internet, the more frequent and bigger mobilisation events are. With an increase by one standard deviation of internet shutdown, mobilisation on average increases by 0.354 standard deviations. The effects are in line with the hypothesised effects (hypothesis 1 and hypothesis 3).

In contrast, disinformation campaigns, $b_2 = 0.226$, $p < 0.05$, have an unexpected positive effect on mobilisation and social media shutdown, $b_4 = -0.237$, $p = 0.073$, is statistically insignificant at the 5% significance level. Interestingly, these two effects have an opposite direction from what I hypothesised (hypothesis 2a and hypothesis 4). Compared to the effects of internet filtering and internet shutdown, the effects are small, but together the four indicators of DR account for 25.5% of the total variance of mobilisation (Adj. $R^2_{\text{Model 1}} = 0.255$). The explanatory power of the baseline model (Model 1) is small which indicates omitted variable bias. The following analysis further investigates the counter-intuitive effect of disinformation campaigns and social media shutdown on mobilisation.

5.1.2 Interaction effect: Disinformation and social media penetration

Model 2 includes the interaction between disinformation campaigns and social media penetration. In partial support with the expectation, the effect of disinformation campaigns is conditional on the social media penetration within a country. But although significant, the interaction between disinformation campaigns and social media penetration is unexpectedly positive, $b_6 = 0.347$, $p < 0.01$. In addition, the main effect of disinformation campaigns is negative, but only significant at the 10% significance level, $b_2 = -0.234$, $p = 0.091$. As such, the support for hypotheses 2a and 2b is insufficient. At the lowest level of social media penetration (indicated by zero), an average level of disinformation campaigns (indicated by zero) has a negative effect on mobilisation in non-democratic regimes. But due to the interaction, this effect decreases if social media use increases. If social media penetration increases by one standard deviation, the effect of disinformation on mobilisation is even reversed,

ultimately, increasing levels of mobilisation. This contradicts the expectations (hypothesis 2a and hypothesis 2b).

With average levels on all forms of DR, social media penetration has an independent effect on mobilisation which is significant and positive, $b_5 = 1.224$, $p < 0.001$. The more people on average use social media to organise offline protest, the higher the level of mobilisation within a country. An increase of social media penetration by one standard deviation increases mobilisation by 1.224 standard deviations. This effect increases with increased levels of disinformation. Adding social media penetration and its interaction with disinformation campaigns to the baseline model (Model 1) has three contributions. First, the relative influence decreases particularly for internet filtering. Second, the effect of disinformation campaigns is reversed and as predicted negative, though strictly speaking insignificant. Third, adding social media penetration to the regression analysis significantly improves the explanatory power by 23.7%, $F(2, 1740) = 408.66$, $p < 0.001$. Thus, the interaction model ultimately explains 49.2% of the total variance, Adj. $R^2_{\text{Model 2}} = 0.492$. As such, accounting for the dynamics of people's social media use provides a better explanation for mobilisation.

5.1.3 Confounding effects: Political, digital, economic, and demographic influences

To rule out omitted variable bias, the following section explores the changes in the main effects after adding controls: political, digital, economic, and demographic. Focusing on political confounds, Model 3 describes that only political instability, $b_7 = 0.167$, $p < 0.05$, and polarisation, $b_8 = 0.151$, $p < 0.05$ have a significant, positive impact on mobilisation. The more likely the presence of violence and terrorism and the more politically divided a country, the more mobilisation is predicted on average by the model. An authoritarian regime's level of institutional authoritarianism (indicated by Polity5), of physical repression (indicated by the PTS score), and of civil liberties have no significant effect on mobilisation.

Furthermore, the main effects are supported. Internet filtering, $b_1 = -0.310$, $p < 0.001$, and internet shutdown, $b_3 = 0.352$, $p < 0.001$, significantly influence the level of mobilisation as already established in Model 1. But compared to Model 2, the relative effect strength of internet filtering decreases. This also applies to the effect of disinformation campaigns. However, the effect of disinformation, $b_2 = -0.198$, $p = 0.158$, is insignificant and the negative effect direction of social media shutdown, $b_4 = -0.217$, $p < 0.05$, is counterintuitive. Additionally, the interaction effect between disinformation and social media penetration becomes significant at the 10% significance level only, $b_6 = 0.227$, $p = 0.064$, which further weakens support for hypothesis 2b.

To finalise the multiple linear regression analysis, I added digital, economic, and demographic sets of controls (Model 4). All effects of these controls are insignificant. The effect directions of the political factors align with the expectations as summarised in Table 2. To sum up, the two main effects are maintained across all models. In the final model (Model 4), internet filtering, $b_1 = -0.312$, $p < 0.01$, and internet shutdown, $b_3 = 0.354$, $p < 0.001$, significantly influence mobilisation. Respectively, an increase by one standard deviation leads to a decrease by 0.312 standard deviations and an increase by 0.354 standard deviations on mobilisation. The implications are two-fold: the relative effect strength and the support for hypothesis 1 and hypothesis 3 persist across models. Yet, it should be highlighted that social media use for offline activity has the comparatively strongest effect on mobilisation, $b_5 = 0.922$, $p < 0.001$.

Against the expectations and accounting for internet penetration, GDP per capita, population size and density, the final model (Model 4) supports the significant negative effect for social media shutdown, $b_4 = -0.262$, $p < 0.05$. Holding all other factors constant, this means that the more often the regime shuts down social media platforms, the less often and smaller mobilisation events on average are. An increase by one standard deviation in social media shutdown leads to a decrease by 0.262 standard deviations in mobilisation. Considering hypothesis 4, this opposing effect direction is puzzling. Overall, Model 4 explains 59.2% of the total variance. This is 1.0% more than Model 3, but 10.0% more than Model 2. In sum, whereas the inclusion of political controls (Model 3) increased the explanatory power relevantly, the inclusion of digital, economic, and demographic controls (Model 4) only minimally contributes to explaining mobilisation.

5.1.4 Robustness checks

Two robustness tests are conducted: the inclusion of country-fixed effects and the inclusion of the lagged dependent variable as independent variable. The former controls for all unobserved, time-invariant country-specific factors. The latter rules out any remaining autocorrelation concerns not addressed by using robust standard errors. Despite the presence of country-fixed effects, the negative effect of internet filtering and the positive effect of internet shutdown remain. The latter effect also remains when including mobilisation as lagged explanatory factor to the model with robust standard errors. Therefore, I conclude that the effects of internet filtering and internet shutdown are robust (see *Appendix*).

5.2 Vignette: Qualitative analysis of digital repression

The quantitative analysis could not resolve the endogeneity problem for short-term DR. This vignette complements the quantitative results with qualitative evidence. I argue that internet filtering and internet shutdown simultaneously influence mobilisation. Statistically, I deduced that aggregated over a year, more filtering leads to more mobilisation, and more shutdowns to less mobilisation. Theoretically, I explained that filtering has a long-term (section 3.3.1), and shutdown a short-term effect (section 3.3.2). To illustrate the interplay of these two dimensions, I outline the mobilisation process of two cases: *Hirak Movement* (HM) in Algeria, 2019 and *Bloody January* (BJ) in Kazakhstan, 2022. I use a combination of journalistic and academic articles to analyse these two movements.

5.2.1 Cases: Hirak Movement in Algeria, 2019 and the Bloody January in Kazakhstan, 2022

I justify my case selection with a comparative introduction. Regarding the similarities, HM and BJ are both cases of social movements for which the authoritarian regimes applied DR (Rydzak et al., 2020; Guest, 2022). More specifically, in both cases internet filtering and internet shutdown are present. This similarity on the independent variable allows to draw conclusions on the short-term effect of shutdown under the presence of internet filtering as long-term mechanism. The two movements differ in what triggered the mobilisation. Whereas HM started as anti-regime movement against the Bouteflika incumbency (Volpi, 2020), BJ began as economic demonstrations against rising gas prices (Toleukhanova, 2022). As theorised before (section 3.1), different forms of dissent, here economic and political, can trigger DR. This differing feature contributes to the generalisability of the argument about the relationship between DR and mobilisation.

5.2.2 Analysis: The influence of internet filtering and shutdown on mobilisation

Regarding the long-term effect of DR, internet filtering became important in Algeria and Kazakhstan after protest movements in 2011. Respectively, the Arab Spring and the Zhanaozen oil worker demonstrations presented digital media as potential threat to autocratic stability (Anceschi, 2015; Benamra, 2020). In response, both autocratic leaders implemented internet filters and thereby conducted repression as prevention. Given that it suppressed comparable mobilisation in the future, this aligns with hypothesis 1. To link these findings to the theoretical framework, I highlight two aspects. First, these cases illustrate that mobilisation triggers repression. Second, in the digital age, repression translates to preventive information control of future regime threats (Michaelsen, 2017). Despite these repressive long-term strategies, the announcement of the Bouteflika candidacy in Algeria and the lifting of the oil price cap in Kazakhstan provoked exceptionally intense waves of demonstrations (Parks, 2019; Putz, 2022). Starting from these protest triggers, Figure 4 summarises

the result of my qualitative analysis on the development of HM and BJ. The escalation in mobilisation highlights that information filtering cannot fully protect the regime from sudden triggers of dissent.

Figure 4 *Results of qualitative analysis on mobilisation escalation under shutdown*

Figure 4.1 *Timeline of HM in Algeria, 2019*

date	*event*	*intensity*
10.02.	Bouteflika announces candidacy	
16.02.	**small-scale protest** *in Kherrata, North-Algeria*	•
17.02. … 20.02.	**small scale protests** *in several cities in Northwest-Algeria*	•
social media shutdown		
21.02.	**small-scale protests** *in more cities in Northeast-Algeria*	•
national internet shutdown		
22.02.	**medium-scale protests** in additional cities, North-Algeria	••
23.02.	**uncoordinated, large-scale protest** *in Algiers, North-Algeria*	•••
24.02.	**coordinated, large-scale protest** *in Algiers, North-Algeria*	•••
local internet shutdowns		
25.02. … 01.03.	**coordinated, sustained wave of protests** *across North-Algeria*	••

Figure 4.2 *Timeline of BJ in Kazakhstan, 2022*

date	*event*	*intensity*
01.01.	lift of price cap on gas prices	
local internet shutdowns		
02.01.	**small-scale protest** *in Zhanaozen, West-Kazakhstan*	•
03.01.	**small-scale protests** *in several cities in South and West-Kazakhstan*	•
	medium-scale protest *in Zhanaozen, West-Kazakhstan*	••
04.01.	**small-scale protests** *in more cities in South and West-Kazakhstan*	•
	medium-scale protest *in Almaty, Southeast-Kazakhstan*	••
national internet shutdown		
05.01.	**small-scale, violent protests** *in South-Kazakhstan*	•
	large-scale, violent protest *in Almaty, Southeast-Kazakhstan*	•••
06.01.	**large-scale, peaceful protest** *in Zhanaozen, West-Kazakhstan*	•••
07.01. … 10.01.	**large-scale, peaceful protests** *in several cities*	•••

Sources: The results are based on international newspapers and existing case-based research (see Bibliography).
Note: The intensity indicates the relative mobilisation size: • **small-scale,** •• **medium-scale,** ••• **large-scale**
The connection quality ranges from high to low: high, medium, low

Focusing on the short-term effect of internet shutdown, I deduce three insights from my analysis (see Figure 4.1 and Figure 4.2). First, for both movements mobilisation took place the same and the following days of shutdown. Second, mobilisation intensified in size at key locations (as indicated by the intensity) while the internet was locally or nationally shut down. To illustrate, the protest in Zhanaozen, Kazakhstan developed from small to large-scale within four days between the 2nd and the 6th of January 2022. A similar development applies to the protest period between the 22nd and 24th of February 2019 in Algiers, Algeria. Third, both movements intensified in frequency as shown by the geographical dispersion. The local shutdown in Zhanaozen, Kazakhstan on the 2nd of January 2022 (Figure 4.2), and the partial shutdown on the 21st of February 2019 and full shutdown on the 22nd of February 2019 in Algeria (Figure 4.1) fuelled the expansion of protests into other regions in less than two weeks (Krapiva et al., 2022; NetBlocks, 2019; Rydzak et al., 2020).

To explain this mobilisation process with *dispersion theory* (Hassanpour, 2014), digital media is used in both cases as online coordination tool (Marat & Tutumlu, 2022; Volpi, 2020). The inclusion of other interest groups via word-of-mouth reinforces this mobilisation process offline in instances of shutdown. HM exemplifies that internet shutdown initiates this procedural shift from online to offline mobilisation (Volpi, 2020). Additionally, BJ exemplifies the substantive dimension of local shutdown. Here, criticism against the regime on Twitter for limiting the freedom of speech further amplified dissent (Krapiva et al., 2022). Both cases support hypothesis 3. To conclude, DR procedurally and substantively influenced mobilisation, not the other way around. I deduce that internet and social media shutdowns lead to the intensification of protest. Overall, this vignette focuses on sub-national changes and analysis the days after internet shutdown. This approach traces the effect of DR on mobilisation. It thereby contributes to resolving the statistical endogeneity problem and the theoretical *punishment puzzle* (Davenport, 2007a).

5.3 Discussion

Before discussing the findings of the quantitative and qualitative analysis, I recap the theoretical framework of this research. This thesis started with the *punishment puzzle* meaning whether and what the effect of DR is on mobilisation. Advocating for the multi-dimensional nature of DR, I used RMT to accentuate the centrality of information for the mobilisation of protest. Using *connective action theory* and *disconnective action theory*, I established two groups of hypotheses. On the one hand, I expected a prevention effect on mobilisation if DR is performed as long-term information manipulation on the network and application layer. By obfuscating the information environment of opposition groups, the connection quality between people is reduced. Hence, the ability for mobilisation is lowered. On the other hand, I hypothesised an escalation effect on mobilisation if DR is executed in form of a short-term information vacuum on the infrastructure layer. The argument is that full or partial shutdown of the internet is an information signal itself which provokes people to mobilise.

Overall and as shown in Table 3, the empirical results are mixed but deliver important theoretical insights. For both hypothesised effects on mobilisation, there is partial support. Regarding the prevention effect, there is a significant, negative relationship of internet filtering. The analysis shows that limited access to politically sensitive information indeed decreases the average level of mobilisation within authoritarian states. In support of hypothesis 1, this preventive form of DR on the network layer can be concluded to be an effective repression tool. The implications are that this changes the form of power regimes hold; internet filtering dictates to people what to think arguably

before they have a reason to question their civil liberties restriction (Y. Chen & Yang, 2019; King et al., 2013).

Contrastingly, hypotheses 2a and 2b were not supported. The main effect of disinformation and its interaction with social media penetration are not consistently significant at conventional levels. After controlling for other factors, the active disinformation engagement of governments does not influence mobilisation (hypothesis 2a). This holds even after addressing the interaction with the use of social media to organise offline activities (hypothesis 2b). Though statistically insignificant, the opposite effect directions of main and interaction effect remain puzzling. A potential explanation is the quality of disinformation campaigns. Accordingly, disinformation which is seen as uncredible creates civic grievance against the regime (J. Chen & Xu, 2017; Honari, 2018). Here, especially societies with high social media use have more access to alternative, credible sources (Feldstein, 2021). Becoming resilient to disinformation, the resulting grievance encourages mobilisation which turns disinformation counter effective. The presence of both credible and non-credible forms of disinformation in the analysis weakens the overall effect. Further analysis needs to scrutinise these predictions by including credibility aspects of campaigns.

Regarding the escalation effect, there is support for hypothesis 3. Internet shutdown has a significant, positive effect on mobilisation. Due to the lack of statistical means to address endogeneity, this short-term dynamic is illustrated by qualitative case studies of HM and BJ. Both quantitative and qualitative evidence suggest that the sudden creation of a short information vacuum triggers mobilisation. In effect, this escalatory form of DR on the infrastructure level is an ineffective repression tactic. People are incentivised to replace the socio-economic resource of digital connectivity with human resources of physical organisation. This indicates that *just-in-time shutdowns* (Keremoğlu & Weidmann, 2020, p. 1692) are used as the last resort for authoritarian regimes. The implication is that short-term DR is a sign of weakness rather than a sign of strength of autocratic durability.

Contrastingly, there is a positive, significant effect of social media shutdowns on mobilisation and thus no support for hypothesis 4. The effect direction is counter-intuitively positive. Assuming a conceptual overlap, there are two possible explanations. Statistically, the proportion of the negative effect of social media shutdowns is covered by internet shutdown turning it positive. This is supported by conducting an analysis excluding internet shutdown (see *Appendix*). Theoretically, the shutdown of social media arguably describes a more strategic form of DR. The expectation is that authoritarian regimes intentionally push citizens into a controlled digital environment via partial denial of access. In

this new information environment, disinformation and filtering are more impactful on people's decision to protest. Incentivising the use of surveilled alternative communication platforms (Moss, 2018), this explanation also highlights the distinct dynamics for social media shutdown and justifies its separate inclusion.

What is the effect of DR on mobilisation? DR can prevent, but also escalate mobilisation. The support of hypothesis 1 and hypothesis 3 is robust across all models There is a relationship between DR and mobilisation. The results persist even after accounting for autocorrelation by including the lagged dependent variable, and contextuality by accounting for country-fixed effects. Next to these supporting findings, a theoretical explanation has been presented for the counter-intuitive statistical results. DR needs to be credible in the case of disinformation campaigns and targeted in the case of social media shutdown. The implication of the results is that the quality of DR matters for its effectiveness.

6 Conclusion

Based on the preceding discussion I want to conclude the findings with a conceptual and a substantive argument. First, the analysis has shown that conceptual differentiation is necessary to understand contradicting effects. This is illustrated by the opposing effect directions of internet shutdown and internet filtering. As such, I established DR as multi-dimensional concept. Second, the effectiveness of DR is not obvious but complex. To avoid mobilisation, DR is overall more successful if established over a long-term (as internet filtering). DR is potentially also effective if targeted to change communication platforms without triggering the citizens' grievance (as indicated by disinformation campaigns and social media shutdowns). Both aspects emphasise the highly strategic nature of DR. As such, autocratic regime durability remains a long-term process. Shutdowns might rather be evaluated as tactic of the last resort and thus signal that a regime's autocratic durability is endangered.

This paper is addressed to scholars, students, academics, and activists devoted to the disciplines of digital repression, (digital) mobilisation, social movement organisation, or politics of authoritarian regimes. Before recommending pathways for further research, it is important to address the two main limitations. Both are related to the data structure. First, endogeneity concerns are not fully resolved. Due to the aggregated data structure, the derived conclusions are indirect. To explain, the quantitative analysis relied on country-year data due to data availability constraints. Here, it is important to highlight two analytically relevant dimensions for inference. Geographically, DR by authoritarian regimes can be distinguished on two levels: regional and national (Gohdes, 2015). Temporally,

different dimensions of DR have different effects: short- and long-term. The interaction of the two dimensions remains unaddressed. Second, the data reliability of expert surveys is questionable. Especially during shutdown, the reporting of unrest decreases simultaneously (Rydzak, 2018). The resulting high dependency on expert knowledge impedes valid and meaningful inferences of the relative effect between the dimensions of DR. Though a comparison is statistically possible (by comparing beta coefficients), the precision of indicators based on expert surveys is limited. The established relationship cannot be translated to practical terms.

Bringing this thesis' implications and limitations together, there are interesting possibilities for future research. By making a conceptual argument, this thesis provides initial support for the relevance and diversity of the effects of DR. It conceptualises DR as multi-dimensional and concludes that it influences mobilisation by means of a systematic global analysis. Generally, authoritarian DR tactics become more sophisticated (Deibert, 2019; Xu, 2020). As such, a profound understanding of sub-national, local dynamics over time calls for the use of direct, technical indicators to track the impact of DR over smaller time units, including hours and days. Clarifying spatial and temporal dynamics via process tracing might resolve limitations of the present research. This would also complement the broader academic debate since regional studies are rare (Earl et al., 2022). Overall, we may have been naïve about the long-term, preventive calculus of authoritarian regimes. However, there is reason to remain optimistic about the people's online and offline mobilisation which continues to oppose state repression today. The debate on the *digital punishment puzzle* has only begun.

BIBLIOGRAPHY

Bennett, W. L., & Segerberg, A. (2014a). Personalized communication in protest networks. In *The Logic of Connective Action* (pp. 55–86). Cambridge University Press. https://doi.org/10.1017/cbo9781139198752.003

Bennett, W. L., & Segerberg, A. (2014b). The logic of connective action. In *The logic of connective action* (pp. 19–54). Cambridge University Press. https://doi.org/10.1017/cbo9781139198752.002

Center for Systemic Peace (2020). *Political regime characteristics and transitions, 1800-2018 (v5)* [Data set]. Retrieved from https://www.systemicpeace.org/inscrdata.html

Chen, J., & Xu, Y. (2017). Information manipulation and reform in authoritarian regimes. *Political Science Research and Methods*, *5*(1), 163–178. https://doi.org/10.1017/psrm.2015.21

Chen, Y., & Yang, D. Y. (2019). The impact of media censorship: 1984 or brave new world? *American Economic Review*, *109*(6), 2294–2332. https://doi.org/10.1257/aer.20171765

Chenoweth, E., & Ulfelder, J. (2017). Can structural conditions explain the onset of nonviolent uprisings? *Journal of Conflict Resolution*, *61*(2), 298–324. https://doi.org/10.1177/0022002715576574

Coppedge, M., Gerring, J., Knutsen, C. K., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Grahn, S., Hicken, A., Hindle, G., Ilchenko, N., Kinzelbach, K., Krusell, J., Marquardt, K. L., McMann, K., Mechkova, V., Medzihorsky, J., Paxton, P., Pemstein, D., Pernes, J., Ryďen, J., von Romer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Sund-ström, A., Tzelgov, E., Wang, Y., Wig, T., Wilson S., & Ziblatt, D. (2022a). *V-Dem [Country–Year] dataset (v12): Varieties of democracy (V-Dem) project [Data set].* Retrieved from https://www.v-dem.net/vdemds.html

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Grahn, S., Hicken, A., Kinzelbach, K., Marquardt, K. L., McMann, K., Mechkova, V., Paxton, P., Pemstein, D., von Römer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Tzelgov, E., Uberti, L., Wang, Y., Wig, T., & Ziblatt, D. (2022b). *V-Dem codebook (v12): Varieties of democracy (V-Dem) project* [Codebook]. Retrieved from https://www.v-dem.net/vdemds.html

Dalmasso, E., del Sordi, A., Glasius, M., Hirt, N., Michaelsen, M., Mohammad, A. S., & Moss, D. (2018). Intervention: Extraterritorial authoritarian power. *Political Geography*, *64*, 95–104. https://doi.org/10.1016/j.polgeo.2017.07.003

Davenport, C. (2007a). State repression and political order. *Annual Review of Political Science*, *10*, 1–23. https://doi.org/10.1146/annurev.polisci.10.101405.143216

Davenport, C. (2007b). *State repression and the domestic democratic peace*. Cambridge University Press. https://doi.org/10.1017/CBO9780511510021

Deibert, R. J. (2019). Three painful truths about social media. *Journal of Democracy*, *30*(1), 25–39. https://doi.org/10.1353/jod.2019.0002

Deibert, R., Palfrey, J., Rohozinski, R., & Zittrain, J. (2010). *Access controlled: The shaping of power, rights, and rule in cyberspace* (R. Deibert, J. Palfrey, R. Rohozinski, & J. Zittrain, Eds.). The MIT Press.

Diamond, L. (2010). Liberation technology. *Journal of Democracy*, *21*(3), 69–83. https://doi.org/10.1353/jod.0.0190

Dickson, E. S. (2007). *On the (in)effectiveness of collective punishment: An experimental investigation*.

Dragu, T., & Lupu, Y. (2021). Digital authoritarianism and the future of human rights. *International Organization*, *75*(4), 991–1017. https://doi.org/10.1017/S0020818320000624

Earl, J., Maher, T. v, & Pan, J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science Advances, 10* (8,). 1-15. https://doi.org/ 10.1126/sciadv.abl8198

Enikolopov, R., Makarin, A., & Petrova, M. (2020). Social media and protest participation: Evidence from Russia. *Econometrica*, *88*(4), 1479–1514. https://doi.org/10.3982/ecta14281

Feldstein, S. (2019). How artificial intelligence is reshaping repression. *Journal of Democracy*, *30*(1), 40–52. https://doi.org/10.1353/jod.2019.0003

Feldstein, S. (2021). *The rise of digital repression: How technology is reshaping power, politics, and resistance*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780190057497.001.0001

Feldstein, S. (2022). *Digital repression index 2001-2021: Mendeley data* (v2) [Data set]. Carnegie Endowment for International Peace. Retrieved from https://data.mendeley.com/datasets/rrnz8p6rvw/2

Freedom House (2022). *Country and territory ratings and statuses, 1973-2022* [Data set]. Retrieved from https://freedomhouse.org/report/freedom-world

Gerschewski, J. (2013). The three pillars of stability: Legitimation, repression, and co-optation in autocratic regimes. *Democratization*, *20*(1), 13–38. https://doi.org/10.1080/13510347.2013.738860

Gibney, M., Cornett, L., Wood, R., Haschke, P., Arnon, D., Pisanò, A., Barrett, G., & Park, B. (2021). *The political terror scale 1976-2020* [Data set]. Retrieved from http://www.politicalterrorscale.org

Gohdes, A. R. (2015). Pulling the plug: Network disruptions and violence in civil conflict. *Journal of Peace Research*, *52*(3), 352–367. https://doi.org/10.1177/0022343314551398

Goldstone, J. A. (2004). More social movements or fewer? Beyond political opportunity structures to relational fields. *Theory and Society*, *33*(3/4), 333–365. https://doi.org/10.1023/B:RYSO.0000038611.01350.30

Hassanpour, N. (2014). Media disruption and revolutionary unrest: Evidence from Mubarak's quasi-experiment. *Political Communication*, *31*(1), 1–24. https://doi.org/10.1080/10584609.2012.737439

Hassanpour, N. (2017). Decentralization of revolutionary unrest: Dispersion hypothesis. In N. Hassanpour (Ed.), *Leading from the periphery and network collective action* (pp. 27–68). Cambridge University Press. https://doi.org/DOI: 10.1017/9781316493403.002

Hobbs, W. R., & Roberts, M. E. (2018). How sudden censorship can increase access to information. *American Political Science Review*, *112*(3), 621–636. https://doi.org/10.1017/S0003055418000084

Hollyer, J. R., Rosendorff, B. P., & Vreeland, J. R. (2015). Transparency, protest, and autocratic instability. *The American Political Science Review*, *109*(4), 764–784. https://doi.org/10.1017/S0003055415000428

Honari, A. (2018). From 'the effect of repression' toward 'the response to repression' . *Current Sociology*, *66*(6), 950–973. https://doi.org/10.1177/0011392118787585

Howard, P. N. (2010). *The digital origins of dictatorship and democracy: Information technology and political Islam*. Oxford University Press.

Howard, P. N., Agarwal, S. D., & Hussain, M. M. (2011). The dictators' digital dilemma: When do states disconnect their digital networks? *Center for Technology Innovation at Brookings*, *13*, 1–11. http://ssrn.com/abstract=2568619

Keremoğlu, E., & Weidmann, N. B. (2020). How dictators control the internet: A review essay. *Comparative Political Studies*, *53*(10–11), 1690–1703. https://doi.org/10.1177/0010414020912278

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *The American Political Science Review*, *107*(2), 326–343.

Kleiner, T. (2018). Public opinion polarisation and protest behaviour. *European Journal of Political Research*, *57*(4), 941–962. https://doi.org/10.1111/1475-6765.12260

Lichbach, M. I. (1987). Deterrence or escalation? The puzzle of aggregate studies of repression and dissent. *The Journal of Conflict Resolution*, *31*(2), 266–297. https://about.jstor.org/terms

Lupia, A., & Sin, G. (2003). Which public goods are endangered? How evolving communication technologies affect the logic of collective action. In *Public Choice* (Vol. 117).

Lynch, M. (2011). After Egypt: The limits and promise of online challenges to the authoritarian Arab state. *Perspectives on Politics*, *9*(2), 301–310. https://about.jstor.org/terms

McCarthy, J. D., & Zald, M. N. (1977). Resource mobilization and social movements: A partial theory, *American Journal of Sociology*, 82(6). 1212-1241. https://doi-org.ezproxy.leidenuniv.nl/10.1086/226464

Mechkova, V., Pemstein, D., Seim, B., & Wilson, S. (2022). *Digital society project (DSP): Dataset (v4)* [Data set and codebook]. Retrieved from http://digitalsocietyproject.org/data/

Michaelsen, M. (2017). Far away, so close: Transnational activism, digital surveillance and authoritarian control in Iran. *Surveillance & Society*, *15*(4), 465–470. https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/6635/6436http://library.queensu.ca/ojs/index.php/surveillance-and-society/index|

Michaelsen, M. (2018). Exit and voice in a digital age: Iran's exiled activists and the authoritarian state. *Globalizations*, *15*(2), 248–264. https://doi.org/10.1080/14747731.2016.1263078

Moore, W. H. (2000). The repression of sissent: A substitution model of government coercion. *The Journal of Conflict Resolution*, *44*(1), 107–127.

Moss, D. M. (2018). The ties that bind: Internet communication technologies, networked authoritarianism, and 'voice' in the Syrian diaspora. *Globalizations*, *15*(2), 265–282. https://doi.org/10.1080/14747731.2016.1263079

Munger, K., Bonneau, R., Nagler, J., & Tucker, J. A. (2019). Elites tweet to get feet off the streets: Measuring regime social media strategies during protest. *Political Science Research and Methods*, *7*(4), 815–834. https://doi.org/10.1017/psrm.2018.3

Pearlman, W. (2012). Precluding nonviolence, propelling violence: The effect of internal fragmentation on movement protest. *Studies in Comparative International Development*, *47*(1), 23–46. https://doi.org/10.1007/s12116-012-9099-2

Pierskalla, J. H. (2010). Protest, deterrence, and escalation: The strategic calculus of government repression. *Journal of Conflict Resolution*, *54*(1), 117–145. https://doi.org/10.1177/0022002709352462

Polletta, F., & Jasper, J. (2001). Collective identity and social movements. *Annual Review of Sociology*, *27*(1), 283–305. https://doi.org/10.1146/annurev.soc.27.1.283

Postmes, T., & Brunsting, S. (2002). Collective action in the age of the internet mass communication and online mobilization. *Social Science Computer Review, 20*(3), 290-301. https://doi-org.ezproxy.leidenuniv.nl/10.1177/08939302020003006

Rød, E. G., & Weidmann, N. B. (2015). Empowering activists or autocrats? The internet in authoritarian regimes. *Journal of Peace Research*, *52*(3), 338–351. https://doi.org/10.1177/0022343314555782

Rodzvilla, J. (2019). Margaret E. Roberts: Censored: Distraction and diversion inside China's great firewall. *Publishing Research Quarterly*, *35*(1), 185–186. https://doi.org/10.1007/s12109-019-09635-x

Rydzak, J. (2018). *A total eclipse of the net: The dynamics of network shutdowns and collective action responses* [Graduate College]. University of Arizona.

Rydzak, J., Karanja, M., & Opiyo, N. (2020). Dissent does not die in darkness: Network shutdowns and collective action in African countries. *International Journal of Communication*, *14*, 4264–4287. https://ijoc.org/index.php/ijoc/article/view/12770/3185

Sanovich, S., Stukal, D., & Tucker, J. A. (2018). Turning the virtual tables: Government strategies for addressing online opposition with an application to Russia. *Comparative Politics*, 435–454.

Steinert-Threlkeld, Z. C., Mocanu, D., Vespignani, A., & Fowler, J. (2015). Online social networks and offline protest. *EPJ Data Science*, *4*(1), 1–9. https://doi.org/10.1140/epjds/s13688-015-0056-y

Stephan, M. J., & Chenoweth, E. (2008). Why civil resistance works: The strategic logic of nonviolent Conflict. *International Security*, *33*(1), 7–44. http://direct.mit.edu/isec/article-pdf/33/1/7/693187/isec.2008.33.1.7.pdf

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir square. *Journal of Communication*, *62*(2), 363–379. https://doi.org/10.1111/j.1460-2466.2012.01629.x

Valenzuela, S. (2014). Facebook, Twitter, and youth engagement: A quasi-experimental study of social media use and protest behavior using propensity score matching. *International Journal of Communication*, 8, 2046-2070. https://ijoc.org/index.php/ijoc/article/view/2022/1189

Volpi, F. (2020). Algeria: When elections hurt democracy. *Journal of Democracy*, *31*(2), 152–165. https://doi.org/10.1353/jod.2020.0028

World Bank. (2022a). *World development indicators. GDP per capita (current US$)* [Data set and codebook]. Retrieved from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

World Bank. (2022b). *World development indicators. Individuals using the Internet (% of population)* [Data set and codebook]. Retrieved from https://data.worldbank.org/indicator/IT.NET.USER.ZS

World Bank. (2022c). *World development indicators: Political stability and absence of violence/terrorism* [Data set and codebook]. Retrieved from https://data.worldbank.org/indicator/PV.EST

World Bank. (2022d). *World development indicators: Population density (people per sq. km of land area)* [Data set and codebook]. https://data.worldbank.org/indicator/SP.POP.DNST

World Bank. (2022e). *World development indicators: Population, total* [Data set and codebook]. The World Bank. https://data.worldbank.org/indicator/SP.POP.TOTL

Xu, X. (2020). To repress or to co-opt? Authoritarian control in the age of digital surveillance. *American Journal of Political Science*, *65*(2), 309–325.

SOURCES FOR VIGNETTE

Ahmed, H. O., & Chikhi, L. (2019, February 26). *Tens of thousands of Algerian students extend protests against Bouteflika*. Reuters. Retrieved May 26, 2022, from https://www.reuters.com/article/usalgeria-protests-idUSKCN1QF1SP

Al Jazeera. (2019, February 22). *Algerians protest against Bouteflika's bid for fifth term*. Retrieved May 26, 2022, from https://www.aljazeera.com/news/2019/2/22/algerians-protest-against-bouteflikas-bid-for-fifth-term

Anceschi, L. (2015). The persistence of media control under consolidated authoritarianism: Containing Kazakhstan's digital media. Demokratizatsiya: he Journal of Post-Soviet Democratization, 23(3), 277–295.

Baala, H. (2019, February 24). *Police fire tear gas against Algeria's anti-Bouteflika protesters*. Al Jazeera. Retrieved May 26, 2022, from https://www.aljazeera.com/news/2019/2/24/police-fire-tear-gas-against-algerias-anti-bouteflika-protesters

Baidildayeva, D. (18–03-26). *Internet censorship in Kazakhstan: More pervasive than you may think*. openDemocracy. Retrieved May 26, 2022, from https://www.opendemocracy.net/en/odr/internet-censorship-in-kazakhstan/

Benamra, B. A. (2020). New and social media in Algeria [E-book]. In N. Miladi & N. Mellor (Eds.), *Routledge Handbook on Arab Media* (pp. 22–33). Taylor & Francis. https://www.taylorfrancis.com/chapters/edit/10.4324/9780429427084-4/new-social-media-algeria-belkacem-amin-benamra

Business & Human Rights Resource Centre. (2022, January 9). *Kazakhstan: At least 164 killed and thousands injured in mass protests over rising prices, low wages and corruption*. Retrieved May 26, 2022, from https://www.business-humanrights.org/en/latest-news/kazakhstan-at-

least-164-killed-and-thousands-injured-in-mass-protests-over-rising-prices-low-wages-and-corruption/

CABAR.asia. (2022, January 28). *Analysis of protests in Kazakhstan 2019–2022*. Retrieved May 26, 2022, from https://cabar.asia/en/analysis-of-protests-in-kazakhstan-2019-2022

Daaji, S., & Menia, R. (2020, June). *The shrinking of the digital space during the COVID-19 pandemic: Movement building and internet governance in North Africa*. Association for Progressive Communication (APC). https://africaninternetrights.org/sites/default/files/Sodfa_Daaji.pdf

franceinfo Afrique. (2019, February 21). *Algérie: La rue ne peut plus "encadrer" la candidature de Bouteflika à un 5e mandat*. Retrieved May 26, 2022, from https://www.francetvinfo.fr/monde/afrique/algerie/election-presidentielle-en-algerie/algerie-la-rue-ne-peut-plus-encadrer-la-candidature-de-bouteflika-a-un-5e-mandat_3200807.html

Guest, P. (2022, April 26). *In the dark*. Rest of World. Retrieved May 26, 2022, from https://restofworld.org/2022/blackouts/#/kill-switch

Keegan, K. (2022, January 10). *Information chaos in Kazakhstan*. The Diplomat. Retrieved May 26, 2022, from https://thediplomat.com/2022/01/information-chaos-in-kazakhstan/

Krapiva, N., Zhyrmont, A., & Skok, A. (2022, January 18). *Kazakhstan internet shutdowns and protests: Timeline*. Access Now. Retrieved May 26, 2022, from https://www.accessnow.org/kazakhstan-internet-shutdowns-protests-almaty-timeline-whats-happening/

Lillis, J. (2011, December 17). *Kazakhstan: State of emergency introduced in troubled oil town*. Eurasianet. Retrieved May 26, 2022, from https://eurasianet.org/kazakhstan-state-of-emergency-introduced-in-troubled-oil-town

Marat, E., & Tutumlu, A. (2022, January 11). *Why Kazakhstan's protests aren't like those in Ukraine and Belarus*. Foreign Policy. Retrieved May 26, 2022, from https://foreignpolicy.com/2022/01/11/kazakhstans-protests-arent-a-color-revolution/

Meddi, A. (2019a, February 22). *Algérie: Retour sur une journée de mobilisation inédite contre la candidature de Bouteflika*. Le Point. Retrieved May 26, 2022, from https://www.lepoint.fr/afrique/algerie-retour-sur-une-journee-de-mobilisation-inedite-contre-la-candidature-de-bouteflika-22-02-2019-2295605_3826.php

Meddi, A. (2019b, February 24). *Manifestations du 22 février: Pourquoi les Algériens sont en colère*. Le Point. Retrieved May 26, 2022, from

https://www.lepoint.fr/afrique/manifestations-du-22-fevrier-pourquoi-les-algeriens-sont-en-colere-24-02-2019-2295841_3826.php

Naylor, A. (2022, January 6). *What really happened during Kazakhstan's internet blackout?* Gizmodo. Retrieved May 26, 2022, from https://gizmodo.com/what-really-happened-during-kazakhstan-s-internet-black-1848315477

NetBlocks. (2019, March 1). *Internet outages in Algeria during new wave of re-election protests*. Retrieved May 26, 2022, from https://netblocks.org/reports/algeria-internet-disruptions-during-re-election-protests-eQyKwOyD

NEWS WIRES. (2019, February 24). *'No to Bouteflika,' chant Algerians protesting his bid for fifth term*. France 24. Retrieved May 26, 2022, from https://www.france24.com/en/20190223-algeria-protest-bouteflika-election-bid

Osman, I. (2021, September 2). *Algeria: The Arab Spring's late bloomer?* Tahrir Institute for Middle East Policy. Retrieved May 26, 2022, from https://timep.org/commentary/analysis/algeria-the-arab-springs-late-bloomer/

Parks, R. P. (2019). From Protesta to Hirak to Algeria's new revolutionary moment. *Middle East Research and Information Project*, *49*(3/4), 10–14. https://merip.org/magazine/292-3/

Putz, C. (2022, January 5). Unprecedented protests rock Kazakhstan as government clings to familiar script. The Diplomat. Retrieved May 26, 2022, from https://thediplomat.com/2022/01/unprecedented-protests-rock-kazakhstan-as-government-clings-to-familiar-script/

Radio Free Europe/Radio Liberty. (2022, January 10). *Kazakhstan: Nearly half of all protests 2018–2021 tied to environmental, labour and livelihood issues*. Business & Human Rights Resource Centre. Retrieved May 26, 2022, from https://www.business-humanrights.org/en/latest-news/kazakhstan-nearly-half-of-all-protests-2018-2021-tied-to-environmental-labour-and-livelihood-issues/

Toleukhanova, A. (2022, February 18). *What comes after Kazakhstan's 'Bloody January'?* Asia Times. Retrieved May 26, 2022, from https://asiatimes.com/2022/02/what-comes-after-kazakhstans-bloody-january/

Zerrouky, M. (2019, June 8). *A Kherrata, aux sources du soulèvement algérien*. Le Monde. Retrieved May 26, 2022, from https://www.lemonde.fr/afrique/article/2019/06/08/a-kherrata-aux-sources-du-soulevement-algerien_5473351_3212.html

APPENDIX

The Online Appendix can be accessed via the following link:

https://www.dropbox.com/scl/fo/2ioeh73t2d37dw86set9n/h?dl=0&rlkey=127v62357ko61099l74pk4b95

It is structured as follows:

Appendix A: Replication data

Appendix B: Assumption tests

Appendix C: Robustness checks

Appendix D: Extra model (section 5.3)