

Online Protest and Repression in Authoritarian Settings

Aytalina Kulichkina ¹, Annie Waldherr ¹, Nanda Wijermans ²

¹University of Vienna, Department of Communication, Kolingasse 14-16, 1090 Vienna, Austria

²Stockholm Resilience Centre, Stockholm University, Albanovägen 28, 10691 Stockholm, Sweden

Correspondence should be addressed to aytalina.kulichkina@univie.ac.at

Journal of Artificial Societies and Social Simulation 29(2) 4, 2026

Doi: 10.18564/jasss.5960 Url: <http://jasss.soc.surrey.ac.uk/29/2/4.html>

Received: 26-03-2025

Accepted: 17-02-2026

Published: 31-03-2026

Abstract: Digital spaces increasingly shape the dynamics of contemporary activism in authoritarian contexts, often complementing or even substituting physical protests. Social media have been particularly empowering by providing relatively safer means of expressing and amplifying grievances in networked environments. However, they have also enabled regimes and their supporters to engage in digital repression. Existing research highlights the need for exploring online protest-repression dynamics, which we address through an agent-based modeling approach. In this study, we integrate theoretical and empirical insights into online protest and repression in authoritarian settings to model social media interactions that can result in enduring campaigns or suppressed dissent. Our findings suggest that active protest expression on social media can positively contribute to ordinary users' latent propensity to protest, even in highly authoritarian contexts and despite prevailing silence and repression. Conversely, when online protest expression is less visible, digital repression becomes more effective, especially in low-authoritarian settings, as it heightens perceptions of successful suppression among ordinary users. We also find that repressors mobilize a small yet stable subset of ordinary users, which helps them reinforce silence across most scenarios, while the impact of voiced dissent quietly accumulates within those who are potentially ready to protest. Our model implementation provides a baseline for studying online protest and repression dynamics, contributing to the theoretical understanding of digital activism in authoritarian settings and offers insights for policymakers, activists, and social media stakeholders seeking to navigate, counter, and prevent digital repression.

Keywords: Protest, Repression, Social Media, Authoritarianism, Agent-Based Model

● Introduction

- 1.1 Over the last two decades, online and social media have been altering how individuals perceive reality and communicate with each other. Social scientists have paid particular attention to *digital public spheres*—online spaces “where participation is open and freely available to everybody who is interested, where matters of common concern can be discussed, and where proceedings are visible to all” (Schäfer 2015, p. 322). For many citizens, digital public spheres have enabled unprecedented access to information and opportunities for political discussion in large social networks, which have positively affected their participation in civic and political life (Boulianne 2015). Consequently, online political participation has become an important research focus in political communication and related fields.
- 1.2 Online participation has been studied from various perspectives and in diverse contexts. Initially, skeptics questioned its democratic potential by coining the term *slacktivism* to emphasize the passivity and ineffectiveness of online actions in achieving political goals (Morozov 2009). Nevertheless, a rapidly growing body of literature has consistently demonstrated the importance of online forms of engagement and proactivity, rather than passivity, associated with information-seeking, online discussion, and online political expression (Ruess et al.

2021). These activities are especially relevant in authoritarian contexts, where offline institutional participation is constrained, traditional media channels are controlled, and political demonstrations are heavily restricted.

- 1.3 Online participation through social media has been particularly empowering, opening opportunities for citizens in repressive regimes to exchange information, express their grievances, and mobilize support for protests. The Arab Spring, the Euromaidan protests in Ukraine, the Gezi Park protests in Turkey, the Umbrella Movement in Hong Kong, and the COVID-19 protests in China are just a few notable examples of resistance movements that relied heavily on social media as a key form of communication. Numerous articles and books have praised the empowering function of digital and social media for such collective and *connective action* (Earl & Kimport 2011; Bennett & Segerberg 2013).
- 1.4 At the same time, another strand of literature has been preoccupied with how autocrats and their supporters employ digital and social media to suppress activism and nascent dissent (e.g., Rød & Weidmann 2015; Feldstein 2021). Various digital repression techniques targeting digital protesters have been identified and scrutinized, including surveillance, censorship, disinformation, and distraction tactics employed by state and private actors (for an overview, see Earl et al. 2022). Both branches of literature indicate that contestation between protesters and repressors unfolds on social media within authoritarian contexts, a dynamic warranting closer attention and joint investigation (Kulichkina et al. 2025a). Of particular interest to this study is how online protest and repression mobilize public support among ordinary social media users, thereby shaping the visibility and power of resistance in digital public spheres.
- 1.5 Studying these dynamics using conventional empirical designs, however, remains challenging due to the delicate subject and data access constraints. On the one hand, social desirability bias and *preference falsification* (Kuran 1989, 1995) complicate empirical data collection in authoritarian societies, which often results in distorted or missing responses about protest-related attitudes and behaviors (Kalinin & de Vogel 2016). On the other hand, social media data alone cannot reveal patterns of passive participation and attitude changes among silent users, since only the public traces of active users are recorded. Moreover, access to such data has become increasingly restricted for academic research by social media platforms. Agent-based modeling (ABM) allows us to bypass these difficulties and explore research questions beyond the scope of traditional empirical approaches.
- 1.6 Existing agent-based models have primarily explored the dynamics between physical protest and repression, including the civil violence model (Epstein 2002), the simulation of the Arab Spring uprisings (Epstein 2014), the emotion-based model of non-violent dissent (Dornschneider-Elkink & Edmonds 2024), models of street protest repression (Akhremenko & Petrov 2020; Petrov et al. 2023), and the policing of protests and counter-protests (Lee 2018). A few conceptual (Waldherr & Wijermans 2017) and implemented agent-based models (Makowsky & Rubin 2013; Hu et al. 2014; Chueca Del Cerro 2023) have accounted for the complementary role of social media in offline collective action. However, the online dimension of protest-repression dynamics in authoritarian contexts remains underexplored.
- 1.7 In this study, we employ ABM to simulate online communication between protesters, repressors, and ordinary citizens in a networked environment under varying degrees of authoritarianism. To this end, we extended and implemented our conceptual model of online protest and repression in authoritarian settings (Kulichkina et al. 2025b). The model is described in detail following the ODD (Overview, Design Concepts, Details) protocol (Grimm et al. 2020); the full ODD is available on CoMSES Net (<https://www.comses.net/codebases/e346d6d4-11d1-4bfa-b818-c2d00941a068/releases/1.0.0/>), and Appendix A offers a condensed summary. The following section provides an overview of the key theoretical and empirical insights, along with relevant concepts, that guide the design and assumptions of our baseline model.

● The OPRAS Model

- 2.1 The Online Protest and Repression in Authoritarian Settings (OPRAS) model aims to explore theoretical understandings of the protest-repression dynamics on social media in authoritarian contexts through an ABM approach. The simulation approximates a 10-day period of online contestation in a networked environment, during which protest and repression unfold, given a specified level of authoritarianism. The initial configuration represents the everyday online status quo prior to any connective action. The agents, representing social media accounts, can display behaviors of repression, protest, or silence. While some committed agents constantly engage in repression or protest (committed users), the majority consists of ordinary users who, depending on their changing perceptions of the social environment and their propensity toward protest or repression, choose

either to remain silent or express their political stance. Figure 1 illustrates the OPRAS overall model, and the following section outlines its theoretical and empirical foundations.

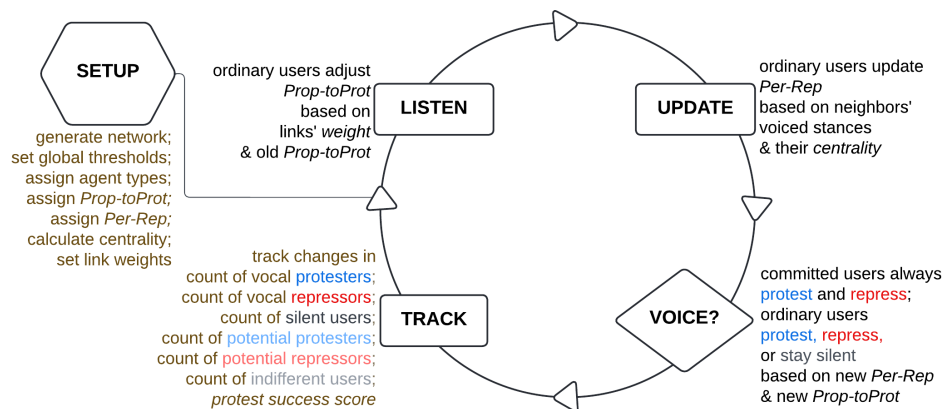


Figure 1: Process diagram of the OPRAS model. The hexagon shows the setup procedures; the circular loop illustrates actions within a single time step. Model entities, variables, and scales are detailed in Appendix A.

● Theoretical and Empirical Foundations of the Model

- 3.1 Current research on online participation and related phenomena, such as echo chambers, filter bubbles, and polarization on social media, predominantly focuses on democratic contexts and relies on theories developed in democratic societies to understand processes concerning citizens in democracies (e.g., Terren & Borge 2021; Ross Arguedas et al. 2022). Social context, however, shapes the experience of online political discussions (Allamong et al. 2024). Authoritarian societies differ significantly in this regard, as unsanctioned political activities are often discouraged and tightly controlled in these contexts. Notably, public political expression is often suppressed due to social pressures, particularly when individuals holding dissenting views are perceived to be in the minority, as explained by the *spiral of silence* theory (Noelle-Neumann 1974). We incorporate this concept into our model by distinguishing between *silent* social media users and those who actively *voice* their protest or their endorsement of the regime.
- 3.2 In non-democratic contexts, opportunities for protest can still arise, depending on a multitude of factors that constitute *political opportunity structure* (Tarrow 1996). In this study, we specifically focus on the micro-level perceptions of individuals, rather than macro-level opportunities, as ordinary citizens in authoritarian contexts often lack complete information needed to accurately assess external opportunities. According to Kim and colleagues' game-theoretic model (Kim et al. 2015), in autocracies, the *perceived probability of successful repression*—repression that effectively deters or suppresses collective dissent—plays a key role in predicting citizen protest. Moreover, uncertainty about the effectiveness of repression is shared not only by ordinary citizens, but also by active protesters and repressors (Lebas & Young 2023). In other words, it is usually unclear how likely repression is to succeed in quelling mobilization—individuals form their own perceptions of its potential success based on the available information. Consequently, the best opportunity for protest is created when the probability of successful repression is perceived to be low (Kim et al. 2015). In our model, we incorporate this perceived probability as an agent attribute that ranges from low (0) to high (1). It changes over time based on the prevalence of either protest or repression and silence in agents' immediate network. The initial values of this variable vary across agents and follow a normal distribution, where the mean value depends on the *level of authoritarianism*.
- 3.3 Authoritarianism has been extensively studied in political science and psychology, leading to various measurement approaches developed over the past 70 years. In this study, we adopt the political science approach, relying on the V-Dem index for regimes of the world (Lührmann et al. 2018), the leading provider of quantitative democracy data that uses the most elaborate and granular democracy indices (Hegedüs 2020). We rely on their classification of regimes, ranging from closed autocracy (e.g., North Korea) to liberal democracy (e.g., Denmark), which is ranked based on the competitiveness of access to power, as well as liberal principles (Lührmann et al. 2018). Informed by this classification, our model can be set to low, moderate, and high *levels of authoritarianism*, where “low” refers to the lower bound of electoral autocracy, “moderate” to electoral autocracy, and

“high” to the upper bound of electoral autocracy. The difference between these levels refers to the degree of electoral manipulation and political freedoms within the regime. Based on (Lührmann et al. 2018), a low level involves limited political freedoms with manipulated elections, a moderate level features some electoral competition but heavier control, and a high level represents regimes with fully controlled elections and no meaningful opposition. In the model, the *level of authoritarianism* further determines the thresholds of *propensity to protest* at which agents publicly express protest or publicly suppress it.

- 3.4 Existing research in democratic contexts offers various approaches to operationalizing and measuring individuals’ *propensity to protest*. This often includes behaviors indicative of non-institutional political participation, such as signing petitions, joining boycotts, attending demonstrations, joining strikes, and other acts of protest (Chang et al. 2021). These activities are typically averaged or summed to create an index reflecting individuals’ propensity to protest ¹. In authoritarian contexts, political protests are actively discouraged and suppressed by governments. Consequently, the measurement of propensity to protest using existing approaches in such contexts is complicated by social desirability bias and *preference falsification* (Kuran 1989, 1995), as expressing support for or compliance with the regime is often perceived as necessary (Kalinin & de Vogel 2016). Therefore, the operationalization of the propensity to protest in authoritarian regimes should differ from that in liberal societies.
- 3.5 In this study, we treat public expressions of criticism against the government and its actions in authoritarian contexts as acts of *protest*. Accordingly, we conceptualize the *propensity to protest* as an individual’s inclination to publicly express criticism towards the regime on social media. This variable ranges from a strong inclination to express support for the regime (0) to a strong inclination to express protest against the regime (1). Since passive acceptance and apathy are characteristics of authoritarian regimes (Linz 1964), we expect the majority of ordinary citizens to have little to no inclination to publicly express support for the regime or protest. Drawing on research on digital repression (for an overview, see Earl et al. 2022), we conceptualize *repression* as expressing stances aligned with government propaganda or posting information aimed at raising the costs for digital social movement activity. Therefore, in our model, low values of propensity to protest signify a tendency toward involvement in repression on social media.
- 3.6 Social media have been pivotal for major social movements and protests of the last two decades. The key feature of social media that both enables and constrains connective action is their networked structure, which allows for a rapid and widespread dissemination of information and inclusive online participation (Howard & Hussain 2013; Bennett & Segerberg 2013; Tufekci 2017). Therefore, we model communication dynamics within a networked environment, with nodes representing *social media accounts* and links representing *connections* between them. Since many protests in authoritarian settings have utilized large public social media platforms like Twitter (Kermani 2025), the model must appropriately represent a power law, where a few accounts have many connections while most have relatively few, enabling both broadcasting and viral diffusion (Goel et al. 2016). In our model, the power law is implemented in a scale-free network layout generated using the Barabási-Albert preferential attachment mechanism (Barabási & Albert 1999). The resulting degree distribution, illustrating the variability in agent connectivity, is shown in Appendix A, Figure 10.
- 3.7 The model includes agents representing different types of *social media accounts*: *ordinary users*, *committed protesters*, and *committed repressors*. Ordinary users reflect the behavior of ordinary citizens on social media: they consume information coming from their connections, are influenced by it, and sometimes choose to actively post themselves. Committed accounts are characterized by perseverance and repeated public participation in discussion (Freelon et al. 2018). They may represent activists, social bots, trolls (i.e., disruptive accounts), or elves (i.e., counter-troll accounts) who continuously promulgate their stances in support of either protest or repression and never go silent.
- 3.8 Since being active or silent on social media depends on the surrounding network structures (Sohn & Choi 2023), in our baseline model, social media accounts influence each other via the undirected links representing *connections* (i.e., mutual access to content). ² The links are defined by their *weight*, which reflects the strength or weakness of the relationships between agents, ranging from 0 (absent) to 1 (strongest). Strong ties are known to facilitate successful persuasion for collective action (Centola & Macy 2007). Thus, in our model, social media accounts are persuaded by linked agents with strongly weighted links (i.e., above 0.5). Persuaded agents are encouraged to publicly post messages, thereby influencing others in the immediate network (Sohn & Leckenby 2007), while uninfluenced agents remain silent. In the model, the influence through public expression in the immediate network affects *propensity to protest* in ordinary users.
- 3.9 Finally, individuals weigh multiple opinions before expressing their own (Granovetter & Soong 1988), a principle that also applies to social reinforcement in online networks (Centola 2010). Thus, before deciding to protest or repress, our agents assess the strength of repression in their immediate networks based on the *probability of*

successful repression and the *centrality* of linked users. The centrality score represents the importance or well-connectedness of an agent in the network, ranging from 0 (lowest) to 1 (highest). Ordinary users assign greater importance to information from well-connected neighbors when assessing and adjusting their perceived probability of successful repression. Figure 2 illustrates the entire decision-making process of ordinary users.

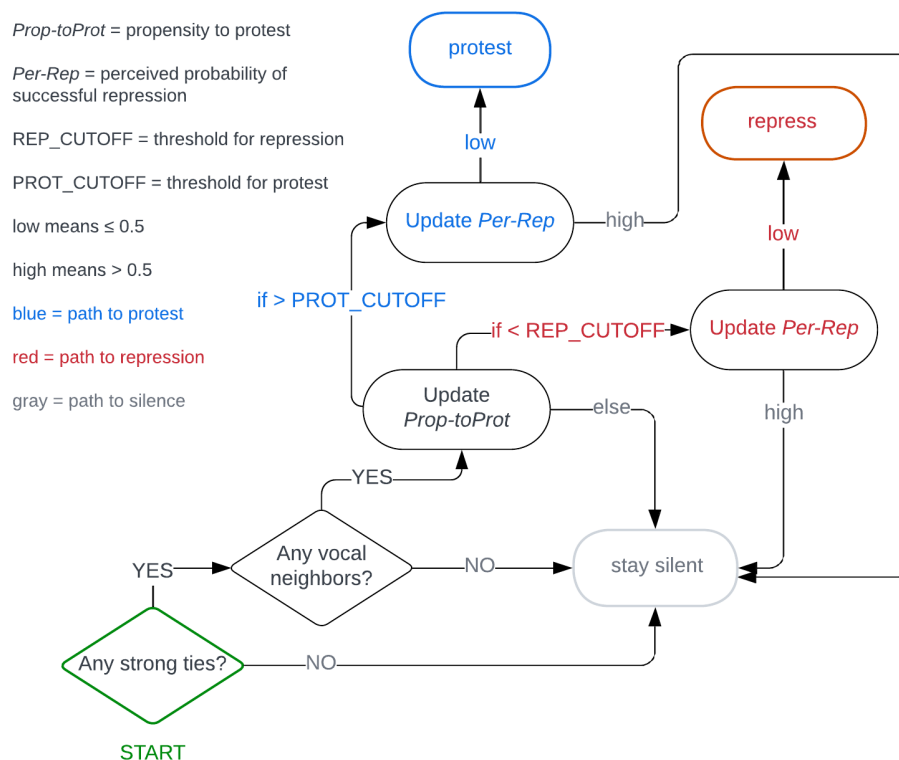


Figure 2: Ordinary user's decision-making process towards staying silent, protest or repress within a single time step. Model variables and submodels are detailed in Table 1 and Appendix A.

Experimental Design

- 4.1 To understand how macro-level protest-repression outcomes in online arenas emerge from micro-level actor interactions—dynamics largely unexamined in existing research—we explore how varying the proportion of highly active committed accounts shapes outcomes in simulated networked environments. The model thus captures how online protest and repression dynamics unfold from differing initial actor compositions toward stable equilibrium states. Building on the theoretical and empirical foundations outlined above, we model interactions among ordinary citizens, committed protesters, and committed repressors under different levels of authoritarianism to address the following research question:
- 4.2 *RQ1: How do online protest and repression dynamics unfold across varying compositions of committed accounts and different levels of authoritarianism?*
- 4.3 To address RQ1, we focus on outcome variables that indicate the manifest dynamics of online protest and repression. Specifically, we track changes in *actor composition*, including the number of accounts voicing protest, publicly supporting repression, and remaining silent. Thus, the analysis is aimed at investigating and comparing changes in *actor composition* across varying initial combinations of *committed accounts* under low, moderate, and high authoritarianism levels.
- 4.4 Additionally, to explore how these manifest system-level dynamics are accompanied by latent micro-level processes—an aspect underexplored in authoritarian contexts due to social desirability bias and preference falsification—we trace individual propensities and perceptions by posing the following research question:
- 4.5 *RQ2: How do ordinary users' internal propensity to protest and their perceived probability of successful repression change during online protest and repression contestation?*

- 4.6** To address RQ2, we track changes in variables that reflect the aggregate latent states of ordinary users, including the *average propensity to protest* and the *average perceived probability of successful repression*, which are calculated by averaging individual values across all agents at each time step. This can reveal how the dynamics of online protest and repression shape the overall internal states of agents, reflecting their potential to engage in protest or repression. These outcome variables are compared across varying initial combinations of *committed accounts* under low, moderate, and high authoritarianism levels.
- 4.7** The variation of the potential independent variables, such as *number of accounts*, *initial propensity to protest*, *initial perceived probability of successful repression*, *strength of ties*, and *ticks* is not of interest for the current research questions, therefore, they remain fixed as control variables in the model. The variable *number of accounts* determines the total number of agents in the model and is set to 1000 for the comparability of the results.
- 4.8** All agents start with a randomly assigned *initial perceived probability of successful repression*, which is normally distributed around a mean determined by the authoritarianism level with a standard deviation of 0.2 to account for variability. This variable always stays between 0 and 1. The same applies to the *initial propensity to protest*, which is normally distributed with a mean of 0.5 and a standard deviation of 0.2. A sensitivity analysis was conducted to assess how variations in the initial values would impact the outcome variables, ensuring that the chosen parameter values accurately reflect plausible scenarios (for details, see Appendix B). All links in the network obtain a different random weight between 0 and 1, which indicates the *strength of ties*. The final control variable *ticks* defines the number of steps per simulation and is set to 250 since the model tends to stabilize around this value. Table 1 presents an overview of the variables of interest. All simulation parameters are listed in Appendix C.³

Independent variables	Control variables	Dependent variables
Actor composition	Number of accounts	Actor composition
Authoritarianism level	Initial propensity to protest	Average propensity to protest
	Initial perceived probability of successful repression	Average perceived probability of successful repression
	Strength of ties	
	Ticks	

Table 1: Classification of variables

- 4.9** For each level of authoritarianism (low, moderate, and high), we vary the percentages of *committed protesters* and *committed repressors* in 2% increments. Since the majority of social media users (approx. 80%) are only occasionally active, and largely idle under everyday circumstances (Williams et al. 2012), we assume that during protests, the percentage of committed accounts may increase but should remain within realistic bounds. Therefore, we test scenarios with a maximum of 20% of committed protesters and 20% committed repressors. Each design point was simulated for 1000 runs to account for stochastic variability. All relevant factors and their factor level ranges are outlined in Appendix D.

● Results

Research question 1

- 5.1** RQ1 inquired about the development of online protest and repression dynamics across varying compositions of committed accounts and different levels of authoritarianism. First, we present the results for the mean number of vocal protesters. Across 1,000 simulation runs, the mean number of protesters varies smoothly with the proportions of committed actors, increasing with the proportion of committed protesters and decreasing with the proportion of committed repressors (Figure 3). The main difference across the three levels of authoritarianism is that the total number of mobilized protesters decreases by roughly 10% as authoritarianism increases. Moreover, under conditions of high authoritarianism and in the absence of committed protesters, almost no protest is voiced. In contrast, low-authoritarian settings enable greater mobilization, especially when committed protesters are present.

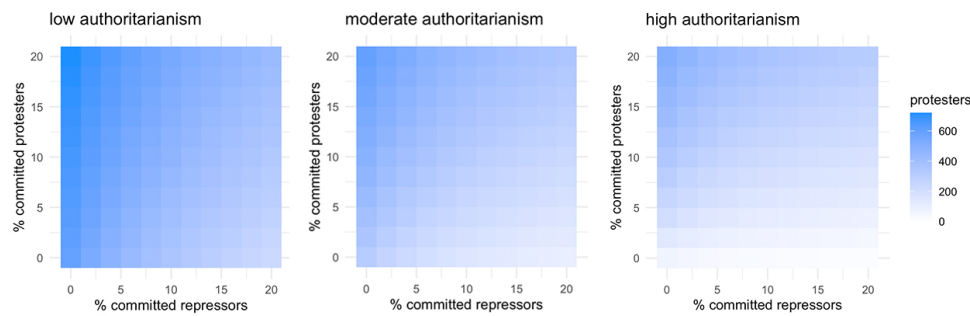


Figure 3: The mean number of protesters at the final tick under varying authoritarianism levels averaged over 1,000 runs.

- 5.2** The corresponding standard deviations are generally low across most compositions of committed accounts, indicating convergence toward stable and predictable outcomes in most scenarios (see Appendix F). However, slightly elevated variability is observed at low to moderate numbers of committed accounts, suggesting more stochastic outcomes. These outcomes are especially pronounced under low authoritarianism or when the number of committed repressors is low in higher authoritarian settings.
- 5.3** Next, we report the results for the mean number of vocal repressors. Across 1,000 simulation runs, the mean number of ordinary citizens mobilized by committed repressors increases only slightly (less than 5%) as the number of committed repressors grows (Figure 4). However, it remains stable across all values of committed protesters across all levels of authoritarianism. This indicates that the presence of committed protesters does not significantly influence the mobilization of vocal repressors, regardless of the authoritarianism level.

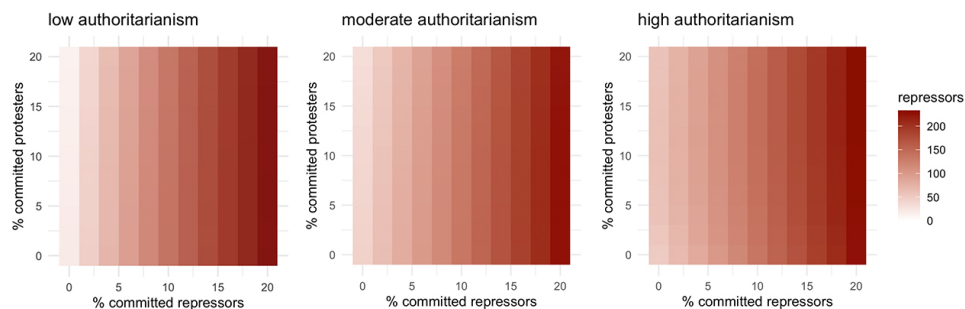


Figure 4: The mean number of repressors at the final tick under varying authoritarianism levels averaged over 1,000 runs.

- 5.4** Additionally, standard deviations (see Appendix F) reveal only minor differences across the studied scenarios, with the highest variability occurring under high authoritarianism and low numbers of committed repressors. Under moderate authoritarianism, the lack of committed accounts on both sides leads to less predictable outcomes. In contrast, under low authoritarianism, greater variability is observed when there are few committed protesters.
- 5.5** Finally, we present the findings for the mean number of silent accounts. Across 1,000 simulation runs, the mean number of silent ordinary citizens steadily increases as the number of committed protesters declines and the number of committed repressors rises, particularly under low authoritarianism (Figure 5). However, under moderate and, even more so, under high authoritarianism, an increase in committed repressors leads to a decrease in the number of silent accounts. When considered alongside the earlier results on the mean number of vocal protesters and repressors, this pattern suggests a higher tendency among individuals to voice support for the repressive agenda as authoritarianism intensifies. Overall, these findings indicate that greater visibility of expressed political stances encourages ordinary users to take a side and engage in online contestation.

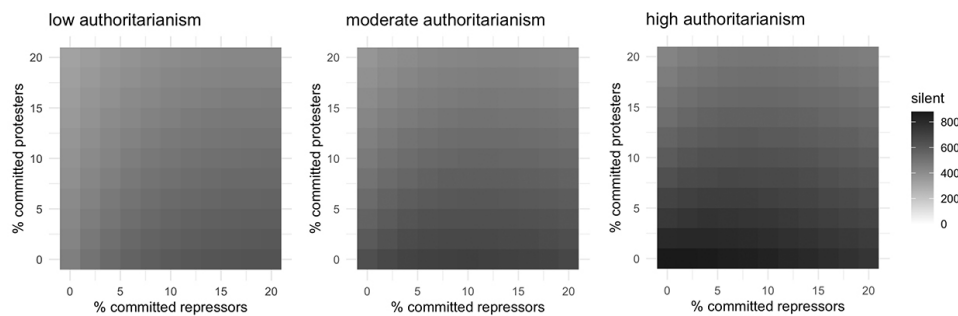


Figure 5: The mean number of silent users at the final tick under varying authoritarianism levels averaged over 1,000 runs.

- 5.6** Standard deviations (see Appendix F) reveal slightly elevated variability at low to moderate numbers of committed accounts under low authoritarianism. In contrast, greater variability is observed at low levels of committed repressors under moderate and high authoritarianism. This suggests that silence during online protests in authoritarian regimes becomes less predictable with fewer committed accounts.
- 5.7** We illustrate the development of online protest-repression dynamics over time using the mid-range configuration: under moderate authoritarianism with 10% committed repressors and varying shares of committed protesters (Figure 6). Compared to scenarios with lower authoritarianism or fewer committed repressors, this configuration marks a threshold in the dynamics, as silent and pro-regime accounts together consistently form the majority for the first time. Notably, silent accounts consistently dominate the online space, including in some scenarios with 20% committed protesters. However, as the proportion of committed protesters increases, the number of active protesters also grows and stabilizes at higher equilibrium levels, indicating that even modest increases in committed participation can influence the overall protest capacity.

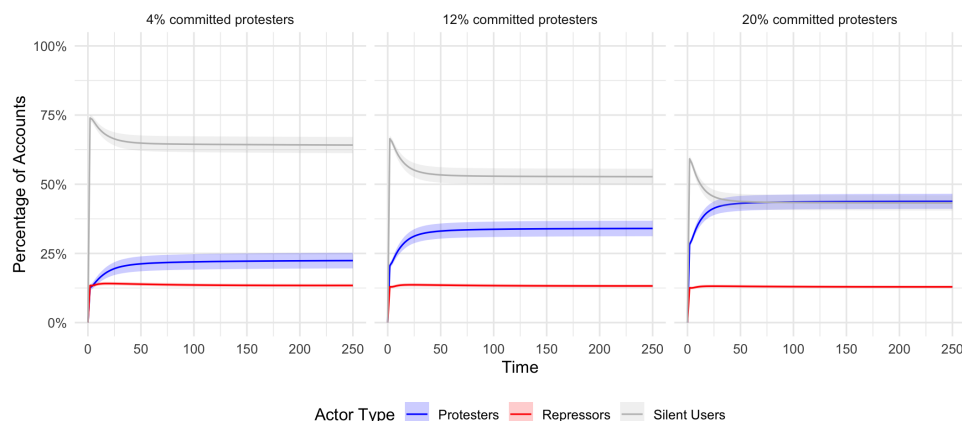


Figure 6: Actor dynamics over time under moderate authoritarianism with 10% committed repressors, averaged over 1,000 runs.

- 5.8** Full over-time dynamics across different combinations of authoritarianism levels and actor compositions are presented in Appendix E.

Research question 2

- 5.9** Having examined the manifest system-level dynamics, we now turn to the latent processes crucial for understanding how individual internal drivers collectively sustain or dampen online mobilization. RQ2 inquired about the ordinary users' internal *propensity to protest* and their *perceived probability of successful repression* during online protest and repression dynamics. The findings appear to align with expectations: the propensity to protest increases as the number of committed protesters grows and the number of committed repressors

declines (Figure 7). What is notable, however, is that these dynamics remain consistent across all levels of authoritarianism. This suggests that active expressions of protest on social media can positively contribute to ordinary users' latent propensity to protest, even in highly authoritarian contexts and despite prevailing silence and regime support.

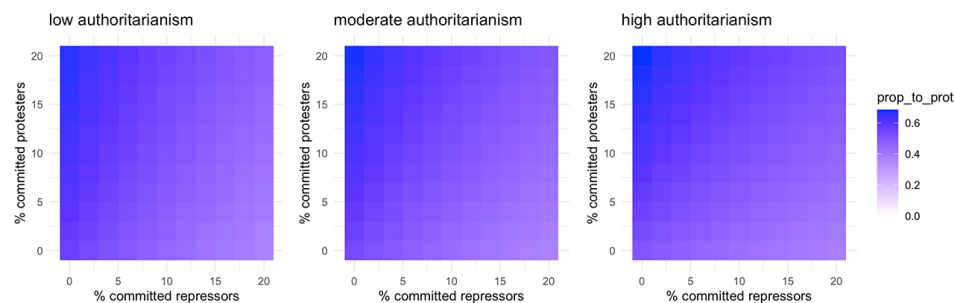


Figure 7: Mean propensity to protest at the final tick under varying authoritarianism levels averaged over 1,000 runs.

- 5.10** The standard deviations of the propensity to protest (see Appendix F) indicate relatively low variability (MAX = 0.019), suggesting that ordinary users' protest intentions are generally stable across runs and do not fluctuate widely. Under conditions of low authoritarianism, variability is minimal and concentrated around balanced numbers of committed protesters and repressors. However, as authoritarianism increases, variability slightly rises, particularly when the number of committed repressors is low.
- 5.11** In contrast, the perceived probability of successful repression varies significantly across contexts, increasing with the level of authoritarianism and the proportion of committed repressors (Figure 8). The lowest values are observed in scenarios where there are no committed repressors under low authoritarianism, as well as when the number of committed protesters increases under moderate and high authoritarian settings. In contrast, the highest values are observed under high authoritarianism, especially in scenarios with less committed protesters and more committed repressors. This leaves only a narrow window of opportunity for protests in highly authoritarian contexts, specifically when committed repressors are largely absent and committed protesters are widely present.

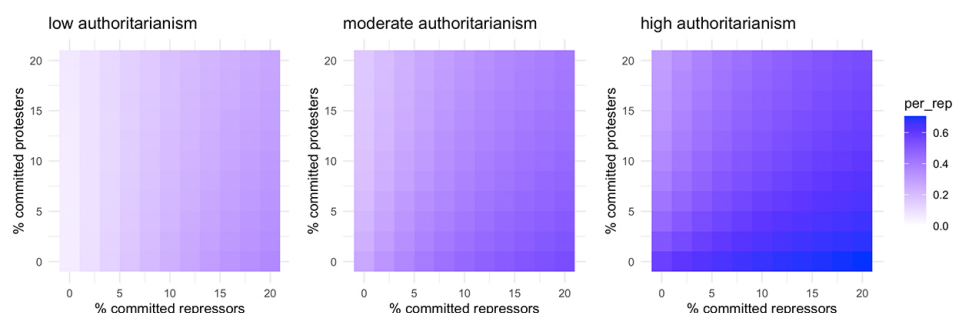


Figure 8: Mean perceived probability of successful repression at the final tick under varying authoritarianism levels averaged over 1,000 runs.

- 5.12** The standard deviations indicate a greater potential for modest variability under low authoritarianism, as well as when the number of repressors is low under moderate and high authoritarian settings (see Appendix F).
- 5.13** We illustrate how the two latent processes evolve over time under moderate authoritarianism across three configurations (Figure 9): with no committed protesters and repressors (0%–0%), with 10% committed protesters and repressors (10%–10%), and with 20% committed accounts (20%–20%). The results show that increasing the proportion of committed actors simultaneously leads to lower average propensities to protest and higher averaged perceived probabilities of successful repression. Notably, although the proportions of committed protesters and repressors are equal in each scenario, greater overall commitment amplifies perceptions of successful repression more than it stimulates the inclination to protest across the three configurations. Nevertheless, a subtle increase in the propensity to protest is observed in individual scenarios, suggesting that committed online participation can enhance the potential for protest mobilization.

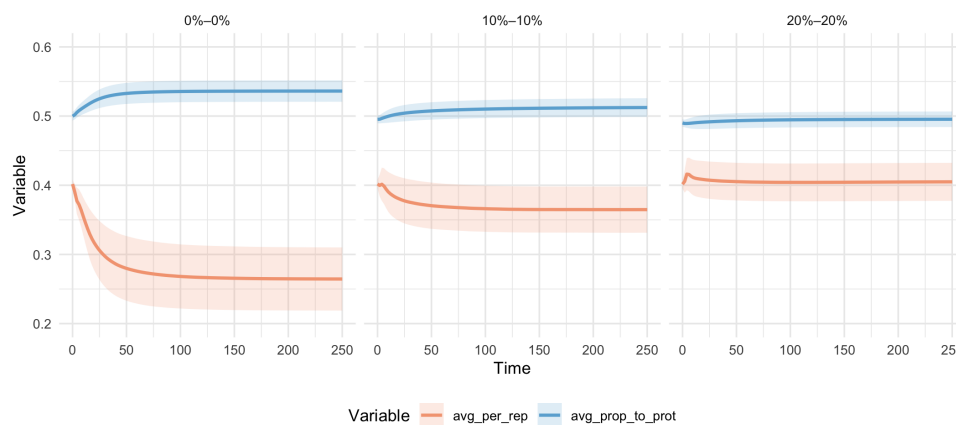


Figure 9: Development of propensity to protest and perceived probability of successful repression over time under moderate authoritarianism averaged over 1,000 runs.

5.14 Full over-time dynamics across different combinations of authoritarianism levels and actor compositions are presented in Appendix E.

● Discussion

- 6.1** Our findings demonstrate that even under high authoritarianism, expressed protest continues to influence ordinary users. While not all individuals become vocal participants, many experience an increase in their internal propensity to protest. This suggests that exposure to dissenting voices, even in repressive environments, can contribute to latent attitudinal shifts. Repressors, on the other hand, tend to mobilize a relatively small but stable proportion of loyal supporters over time. The effects of their activity manifest primarily through the growth of silent users, some of whom also develop an internal propensity to support the regime.
- 6.2** To our knowledge, the only empirical study on the dynamic relationship between online protest and repression in authoritarian settings is a case study of the 2021 Pro-Navalny protests in Russia (Kulichkina et al. 2025a). The authors found that protesting Twitter accounts were significantly more vocal (i.e., posted 17.5 times more tweets) than pro-regime accounts throughout the months of the movement. This aligns with our findings, suggesting that committed repressors mobilize only a small proportion of vocal supporters, who remain less visible than protesters in most scenarios. Since Russia was categorized as an electoral autocracy in 2021 by the V-Dem project (Lührmann et al. 2018), which corresponds to a moderate level of authoritarianism in our model, the existing empirical pattern aligns with scenarios at this level (see Figure 6 and Appendix E), suggesting a low proportion of committed repressors and a high proportion of committed protesters.
- 6.3** A key takeaway from our analysis is that the average *propensity to protest* tends to rise when the number of committed protesters is equal to or exceeds the number of committed repressors, especially in high-authoritarian settings. This is a promising sign, even though, on the surface, silence appears to dominate. The impact of protest expression and exposure to dissent is gradual and may not lead to immediate mobilization. However, the potential for an accumulating effect over time remains plausible, opening possibilities for large-scale mobilization when political opportunities arise. This aligns with prior research on social movements before the popularization of the internet (Kuran 1989) and during the golden times of social media research (Tufekci 2017), suggesting that latent dissent can eventually spill over into action given the right triggers. Future research could further investigate, conceptualize, and categorize various political opportunities in fueling or delaying dissent online. These opportunities can then be integrated into the model as new agents, links, or environmental elements.
- 6.4** Our model serves as a baseline model version featuring a fixed network of 1,000 social media accounts, randomly generated for each model run. While this provides a controlled environment for studying online protest-repression dynamics and their underlying capabilities, future research could expand on this by modeling networks' evolutions over time. This could involve incorporating directed networks and mechanisms such as unfriending, following and creating new accounts, as well as strengthening or weakening virtual relationships. Such model extensions would allow longer simulations spanning months and years. To this end, future research

could focus on gathering empirical data to establish robust foundations for these mechanisms in authoritarian settings.

- 6.5** In its current form, our model has built-in mechanisms that enable the exploration of online contestation also in democratic contexts, such as those between competing coalitions of climate advocates and skeptics (Adam et al. 2019) or social movements and countermovements, such as Black Lives Matter and political conservatives (Freelon et al. 2018). With minimal modifications, the model can be applied to study the role of influential users and issues in various online phenomena, such as hate speech propagation (Kim & Ogawa 2024) or issue-attention cycles (Waldherr 2014) in networked environments. Such a model should rely on empirical and theoretical foundations established in existing literature on democratic societies. The parameters in our model allow for manipulation of different levels of democracy (i.e., from the lower bound of electoral democracy to liberal democracy, according to Lührmann et al. 2018) for research on varying political conditions. Given the recent global democratic backsliding (Knutsen et al. 2024), this remains a particularly relevant area of research.
- 6.6** The increasing role of social media discourse and interactions in shaping perceived realities and political behavior highlights the need for additional empirical inquiries and theory building. The role of platform owners and stakeholders cannot be ignored in this regard, as discontinued data access for academic research prevents proper investigation of relevant online phenomena and reliable calibration and validation of agent-based models. This is especially relevant for modeling repression tactics by a central authority in larger simulation studies. In authoritarian contexts, this could be a central government exercising physical repression towards digital activists or internet blackouts, temporarily reducing the activity of protesters (Feldstein 2021). In democratic contexts, this could be a social media platform, engaging in shadow-banning, account suspensions, and content suppression through algorithmic curation (Earl et al. 2022). Since digital platforms themselves can be viewed as actors in contentious politics (Wijermars & Lokot 2022), future research should conceptually differentiate between corporate-owned, centralized platforms and decentralized alternatives such as Bluesky and Mastodon.
- 6.7** Our study provides a foundation for future simulations of online contestation and underlying capacities of social media interactions. It underscores the importance of looking beyond observable macro behaviors by exploring latent processes forming the perceptions and capabilities of ordinary citizens. Understanding these dynamics is especially important in authoritarian regimes, where traditional methods of data collection produce distorted and often misleading results (Kalinin & de Vogel 2016). Importantly, our study shows that observable online behavior cannot be viewed as a proxy for the underlying capabilities for connective action. This is especially relevant to contexts with moderate or high-level authoritarianism, where silence and repression often dominate over voiced dissent. Such observable dynamics may benefit autocrats and strengthen their regime. Therefore, pro-democracy activists and politicians in democratic societies could engage in committed amplification of anti-authoritarian protests in online spaces beyond their local public spheres. This practice has been found to have beneficial outcomes for social movements in authoritarian regimes (Zeng 2020) and would align with growing translocalization of digital public spheres (Waldherr et al. 2024; Baran & Stoltenberg 2025). Besides, in the face of rising digital repression and democratic backsliding, policymakers and social media platforms in liberal societies should focus on supporting translocal online deliberation and participation. This includes implementing measures that promote online political expression on a global scale, such as stronger encryption and anonymity, detection and mitigation of digital repression, and promotion of censorship-resistant virtual private networks and decentralized hosting. Protecting online participation and protest is essential to safeguarding democracy—locally and globally.

● Model Documentation

The model is implemented in NetLogo 6.4 (Wilensky 1999) and is available on CoMSES Net here: <https://www.comses.net/codebases/e346d6d4-11d1-4bfa-b818-c2d00941a068/releases/1.0.0/>. The model design, initialization, and submodels are detailed in the Appendix A and in the full ODD Protocol at the same repository.

● Acknowledgements

Open access funding provided by the University of Vienna.

We are thankful to the anonymous reviewers, the participants and lecturers of the BEHAVE Summer School 2023, the participants of the 19th annual Social Simulation Conference, and the discussants of the 28th World Congress of Political Science for their valuable feedback and suggestions.

● Appendix A: Model Description

Following the ODD protocol (Grimm et al. 2006, 2020), we include an abridged model description in this Appendix. The full ODD, with all sections and implementation details, is available on CoMSES Net.

Purpose and patterns

The purpose of this model is to understand how online protest and repression dynamics unfold in authoritarian settings and how these dynamics affect ordinary citizens' states and behavior on social media. The model takes into account the most relevant social media features as well as the most central factors shaping online protest and repression in authoritarian contexts. Specifically, the following questions are addressed:

1. How do online protest and repression dynamics unfold under different levels of authoritarianism and varying compositions of committed accounts?
2. How do ordinary users' internal propensity to protest and perceived probability of successful repression change during online protest and repression contestation?

Ultimately, the model develops understanding of how dissenting voices are empowered and suppressed online in authoritarian contexts. Thus, two general macro patterns observable in the real world guide the evaluation of the model:

1. *Enduring protest*: a scenario where online protest dominates, as the number of vocal protesters grows and exceeds the number of vocal repressors, reducing the number of silent users and leading to a stable majority of protesters.
2. *Suppressed protest*: a scenario where online protest is suppressed, as the number of vocal repressors and silent users grows in response to vocal protest, leading to a sustained majority of repressive and silent accounts.

Entities, state variables, and scales

First, the model includes agents representing *social media accounts* that reflect the behavior of different types of social media users or automated accounts. Each agent is randomly assigned an *agent type* reflecting their role in protest-repression dynamics through the content they post: *ordinary user*, *committed protester*, or *committed repressor*. Second, the model includes undirected links representing *connections* between the accounts, which allow users to communicate political content. *Social media accounts* and *connections* form a randomly generated scale-free *network*. The *network* and the *observer* are the system-level entities representing the social media platform as a system and environment.

Entity	Variable	Definition	Value range
Agent	agent-type	the agent's role in the model	{"committed repressor", "ordinary user", "committed protester"}
	Prop-toProt	the agent's propensity to protest	[0,1]
	Per-Rep	the agent's perceived probability of successful repression	[0,1]
	myBehavior	= repress (red) if Prop-toProt < REP_CUTOFF; = protest (blue) if Prop-toProt > PROT_CUTOFF; = silent (grey) otherwise	repress, silent, protest
Links	weight	the strength of the relationship between the linked agents	[0,1]

Table 2: Model's entities, variables, and scales

Agents differ from each other based on several features (see Table 2). First, their agent type (*agent-type*: *committed repressor*, *ordinary user*, *committed protester*) is linked to each agent's propensity to protest (*Prop-toProt* $\in [0, 1]$). Values of *Prop-toProt* in $[0, \text{REP_CUTOFF}]$ indicate a propensity to repress, whereas values in $[\text{PROT_CUTOFF}, 1]$ indicate a propensity to protest. All other values represent silence. Ordinary users

start with Prop-toProt in $(\text{REP_CUTOFF}, \text{PROT_CUTOFF})$, which changes over time due to the influence of linked agents. Committed repressors have a stable Prop-toProt in $[0, \text{REP_CUTOFF}]$, and committed protesters have a stable value in $[\text{PROT_CUTOFF}, 1]$. These values cannot be influenced and do not change over time.

The *perceived probability of successful repression* (Per-Rep $\in [0, 1]$) is another distinct feature of agents, applicable only in authoritarian contexts. Values below 0.5 indicate a low perceived probability of successful repression and a better opportunity for protest, whereas values above 0.5 indicate a high perceived probability of successful repression, a worse opportunity for protest, and less motivation for repression.

Per-Rep is assigned randomly at the beginning based on the level of authoritarianism. Let A denote the authoritarianism level. The mean is

$$\mu = 0.2 (A - 5),$$

and the standard deviation is $\sigma = 0.2$. We assume that repression is generally perceived as less effective in less authoritarian contexts and more effective in highly authoritarian contexts, while also allowing for individual variation in perception to account for uncertainty and diverse experiences.

Per-Rep decreases for an agent if the majority of its linked agents have expressed protest. Likewise, it increases for an agent if the majority of its linked agents have expressed support for the regime.

Links are characterized by their *weight*, which represents the strength of the relationship between the linked agents. The value of *weight* lies in $[0, 1]$, where 0 denotes an absent relationship and 1 the strongest relationship. A value above 0.5 represents a strong tie, while a value below 0.5 indicates a weak tie.

The social media environment is modeled in NetLogo as a non-geographic, two-dimensional patch space of size 40×25 , comprising a scale-free network of social media accounts. The network is generated using the Barabási-Albert preferential attachment mechanism (Barabási & Albert 1999) via the NetLogo *nw* extension, yielding a degree distribution that follows a power law (see Figure 10). Most social media networks are believed to exhibit power-law degree distributions, with a few accounts having disproportionately many connections while the majority have relatively few, enabling both broadcasting and viral diffusion (Goel et al. 2016).

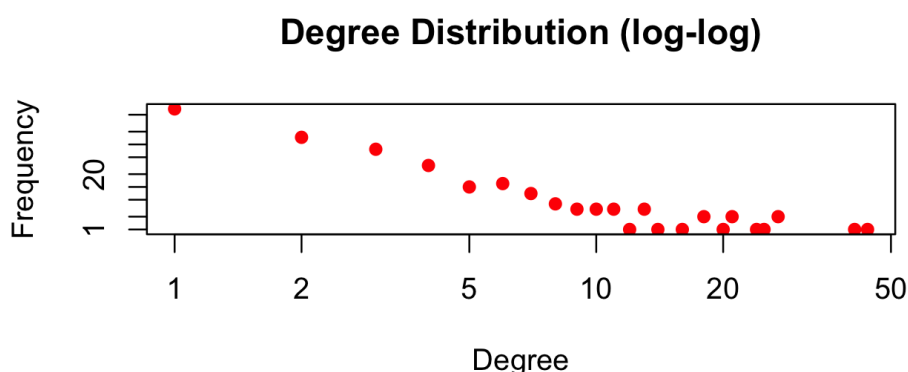


Figure 10: Variability in the number of connections per agent.

The *time scale* represents hours, as communication on social media is typically fast; thus, one tick corresponds to approximately one hour (≈ 1 hour). In general, it is plausible to post or repost on social media within an hour, or to be exposed to a post within the same time window (Howarth 2024).

Process overview and scheduling

The observer executes a setup initialization procedure (detailed below). At each next time step, the “listen”, “update-per-rep”, “voice”, and “update-outcome-var” submodels (detailed below) are executed in this order. At the end of each step, the outcome variables are measured and reported (see Figure 1). All agents are called in a random order to perform their actions.

Design concepts

This model addresses a long-standing problem of the protest-repression nexus within relatively new spaces—digital public spheres (Schäfer 2015). Existing agent-based models focus on conventional street protests and repression (e.g., Carey 2006; Lee 2018; Akhremenko & Petrov 2020; Dornschneider-Elkink & Edmonds 2024), accounting for different forms of repression, levels of authoritarianism, and protest outcomes.

In our model, we incorporate concepts related to the online forms of protest and repression. Specifically, we adopt the concept of *connective action* proposed by Bennett & Segerberg (2013), which emphasizes the essential role of digital media in modern protests, characterized by the diminished role of formal organizations and more horizontal network-building and participation as opposed to the conventional logic of protest organization. We further adopt the concept of *digital repression*, as defined by Earl and colleagues, “as actions directed at a target to raise the target’s costs for digital social movement activity and/or the use of digital or social media to raise the costs for social movement activity, wherever that contestation takes place” (Earl et al. 2022, p. 1).

We further refine our focus to online political participation within *digital public spheres* (Castells 2008). In authoritarian contexts, online protests play a significant role in political life, since offline non-institutional actions are strictly controlled. However, sustaining an online protest can be difficult due to digital repression through propaganda and manipulation, which affects the willingness of the online community to continue active engagement (Feldstein 2021). Silence and inaction on social networks are also meaningful, as they signal the “power” of the regime, which over time can strengthen its influence and perceived legitimacy.

To ground agents’ behavior, we apply theories relevant to authoritarian contexts, such as the *theory of the spiral of silence* (Noelle-Neumann 1974), which suggests that individuals are less likely to publicly express political opinions if they perceive them to be in the minority. This theory explains how minority opinions can be suppressed based on individuals’ perceptions, giving more importance to seemingly dominant opinions. Publicly shared opinions and stances in online spaces have psychological effects and can exert social influence (O’Reilly et al. 2022). We also build on the observation by Kim et al. (2015) that when the *perceived probability of successful government repression* is low, it creates an opportunity for protest. Thus, social media users express protest when they perceive repression as less likely to succeed and believe that others will support their stance. Finally, we draw on Centola (2018) to explain why persuasion for action happens via strong ties as opposed to weak ties in social networks.

Initialization

At the initial state of the model world, the observer initializes key parameters, generates a scale-free social media network, assigns agent attributes, and sets thresholds for continuous variables. The initialization process follows these steps:

1. *Random seed generation*

A new random seed (experiment-seed) is generated and used.

2. *Network formation*

The model generates a scale-free network using the preferential attachment algorithm (Barabási & Albert 1999) with the number of total agents specified in the interface. Agents are randomly positioned within this network in the 2D NetLogo world. The network can be made more visually readable by clicking the `arrange network` button in the interface, which uses the spring layout algorithm 100 times to position nodes based on a force-directed model. Links pull connected nodes together with strength 0.5; the preferred distance between connected nodes is 5; and resistance controlling the dampening of movement is set to 1.

3. *Global thresholds*

A value of 0.5 is fixed as a threshold for the weight (distinguishing between weak and strong relationships) and the perceived probability of successful repression (distinguishing between high and low probability). The threshold for the propensity to protest sets the lowest propensity (`PROT_CUTOFF`) required for agents to voice protest (authoritarianism level/10) and the highest propensity (`REP_CUTOFF`) at which agents engage in repression ($0.5 - (1 - \text{authoritarianism level}/10)$).

4. Agent initialization

Each agent is initially created as an ordinary user. The propensity to protest (Prop-toProt) is drawn from a normal distribution with a mean of 0.5 and a standard deviation of 0.2, ensuring natural variation in protest inclination.

The perceived probability of successful repression (Per-Rep) is drawn from a normal distribution with a mean based on the level of authoritarianism and a standard deviation of 0.2. The mean is calculated as

$$\mu_{\text{Per-Rep}} = 0.2 (\text{authoritarianism level} - 5)$$

A subset of accounts, corresponding to a percentage defined in the interface, is randomly selected as committed protesters. Their propensity to protest (Prop-toProt) is randomly assigned within the range $[\text{PROT_CUTOFF}, 1]$. They are visually represented as blue triangles to distinguish them from ordinary users, as committed protesters can also be bots or elves.

Another subset of users, corresponding to a percentage defined in the interface, is randomly selected as committed repressors. Their propensity to protest (Prop-toProt) is randomly assigned within the range $[0, \text{REP_CUTOFF}]$. They are visually represented as red triangles to distinguish them from ordinary users, as committed repressors can be bots or trolls.

5. Network link weights

Links between users are assigned random weights drawn from a uniform distribution between 0 and 1, representing the strength of relationships.

6. Final setup

Finally, agent sets are defined for *ordinary users*, *committed protesters*, and *committed repressors*, and simulation time is initialized by calling `reset-ticks`.

Submodels

• Listen submodel

The purpose of the “listen” submodel is to update the propensity to protest (Prop-toProt) of *ordinary users* based on the influence of their vocal strong-tie neighbors (i.e., vocal protesters and repressors with strong connections).

1. Each ordinary user identifies its strong-tie neighbors, defined as those with a link weight above 0.5.
2. The agent then identifies vocal neighbors among the strong-tie neighbors with a blue (protester) or red (repressor) color.
3. If there are any vocal strong-tie neighbors, the agent calculates their avg-weight (mean tie weight) and avg-Prop-toProt (mean propensity to protest).
4. The difference (p -difference) between the agent's own Prop-toProt and the avg-Prop-toProt of vocal strong-tie neighbors is calculated.
5. If the difference is positive (i.e., the agent's Prop-toProt is lower than that of its vocal strong-tie neighbors), the agent's Prop-toProt is increased based on the avg-weight and a smoothing factor; if the difference is negative, the Prop-toProt it is decreased.
6. The updated Prop-toProt value is constrained to lie within $[0, 1]$.

• Update Per-Rep submodel

The purpose of the “update Per-Rep” submodel is to update the perceived probability of successful repression of *ordinary users* based on the relative dominance of repressors or protesters among influential accounts in their immediate network.

1. Each ordinary user calculates the sum of the centrality values of its linked vocal neighbors. Weighted protest is the sum of centrality values of blue neighbors (protest influence). Weighted repression is the sum of centrality values of red neighbors (repression influence).
2. Each ordinary user calculates the sum of the centrality values of all linked neighbors (protesters, repressors, and silent users) and stores them as the `total centrality`.

3. The agent then calculates the protest ratio by dividing the weighted protest by the total centrality (proportion of protest influence) and the repression ratio by dividing the weighted repression by the total centrality (proportion of repression influence).
4. The agent then compares repression ratio to protest ratio. If the repression ratio is greater than the protest ratio, it means repression is more influential. In this case, the agent calculates an adjustment (difference between repression ratio and protest ratio). The agent's perceived probability of successful repression is then increased by a fraction of this adjustment, scaled by a learning rate of 0.2. The final value is capped at a maximum value of 1. If the protest ratio is greater than the repression ratio, it means protest is more influential. In this case, the agent calculates an adjustment (difference between protest ratio and repression ratio). The agent's perceived probability of successful repression is then decreased by a fraction of this adjustment, scaled by a learning rate of 0.2. The final value is capped at a minimum value of 0.

- *Voice submodel*

- Condition 1: If the ordinary user's Prop-toProt is low (below the REP_CUTOFF) and their Per-Rep is also low (below 0.5), the user chooses to repress (color = red).
- Condition 2: If the ordinary user's Prop-toProt is high (above the PROT_CUTOFF) and their Per-Rep is also low (below 0.5), the user chooses to protest (color = blue).
- Otherwise: If neither Condition 1 nor Condition 2 is met, the user remains silent (color = grey).

- *Update outcome variables submodel*

The observer counts the number of accounts in each category: protesters (agents with color = blue), repressors (agents with color = red), and silent users (agents with color = grey).

The observer also counts the number of accounts based on their internal propensity to protest: potential protesters (agents with Prop-toProt > PROT_CUTOFF), potential repressors (agents with Prop-toProt < REP_CUTOFF), and indifferent users (agents with Prop-toProt between the REP_CUTOFF and PROT_CUTOFF).

Figure 2 illustrates the entire decision-making process of ordinary users.

● Appendix B: Sensitivity Analysis

The sensitivity analysis addresses the potential sensitivity of simulation runs to stochastic variation in the values of mean-Prop-toProt and SD-Prop-toProt used to initialize the model. These values define the mean and standard deviation of a normal distribution from which *propensity to protest* is randomly drawn (with the resulting values constrained between 0 and 1).

Assuming that *propensity to protest* in authoritarian settings follows a normal distribution—where most of the population falls near the center, while a minority has inherent extreme tendencies towards either active protest or repression (Linz 1964)—we examine how varying the mean and standard deviation (low, moderate, and high) impacts the simulation results. Specifically, we test mean-Prop-toProt values 0.25, 0.5, and 0.75, as well as SD-Prop-toProt values 0.1, 0.2, and 0.3, to compare model dynamics across low, medium, and high authoritarianism levels (see Table 3). We vary one parameter at a time while keeping other parameters constant across 1000 model runs, following the standard one-at-a-time (OAT) sensitivity analysis approach. The number of committed accounts was kept at 10% for both protesters and repressors.

Parameter [range]	Description	Base value (model)	Why this value?
mean-Prop-toProt [0, 1]	Population mean value of propensity to protest	0.5 {0.25, 0.75}	Extreme cases are rare, most individuals may have a moderate propensity
SD-Prop-toProt [0, 1]	Population standard deviation of propensity to protest	0.2 {0.1, 0.3}	Moderate SD, a reasonable middle-ground assumption

Table 3: Model parameters used in sensitivity analysis

The results indicate minimal differences over time in the composition of vocal protesters, vocal repressors, and silent accounts for SD-Prop-toProt values of 0.1, 0.2, and 0.3. The only exception occurs under low authoritarianism, where an SD of 0.1, combined with a mean of 0.25 or 0.5, leads to unrealistic distributions (in relatively

liberal settings). Since SD values of 0.2 and 0.3 produce similar outcomes, we adopt the middle value of 0.2 as a balanced choice for further analysis.

Focusing on the mean-Prop-toProt values of 0.25, 0.5, and 0.75, we observe substantial variations in the outcome variable. A low average propensity to protest (0.25) produces nearly identical patterns of extreme suppression across all levels of authoritarianism, while a high average propensity to protest (0.75) leads to the exaggerated dominance of protesting accounts. Since these scenarios are unrealistic and rather extreme, we conducted additional model runs for mean-Prop-toProt values of 0.4 and 0.6.

The results indicate more realistic outcomes for moderate mean-Prop-toProt values (0.4–0.6). However, a value of 0.4 still leads to unlikely patterns of extreme suppression under low authoritarianism, while a value of 0.6 results in improbably heightened protest activity at the moderate level of authoritarianism. Therefore, we set the middle value of 0.5 as the mean-Prop-toProt for model initialization, with randomness ensuring some variation around this mean. The full details of the sensitivity analysis are available on <https://osf.io/d82tq>.

Notes

¹There are similar concepts such as civil engagement, political participation, and online political participation which intersect with but differ from protest participation. The former includes non-political engagement, such as voluntary work (Ekman & Amnå 2012); the second includes institutionalized participation, such as voting or campaign activity (Verba & Nie 1972), which are less relevant to authoritarian contexts; the latter includes participating in an online Q&A session with a politician or donating money to a campaign using a mobile phone, which are also distinct from protest acts (Gil de Zúñiga et al. 2017). Moreover, Teorell et al. (2007) discern protest activity as one of the dimensions of political participation distinct from others.

²In this model, undirected links do not denote formal relationships such as following or friendships but rather indicate mutual content visibility between agents. Link weight reflects the existence and strength of relationships.

³Full Appendices B-F are available at: <https://osf.io/d82tq>

References

- Adam, S., Häussler, T., Schmid-Petri, H. & Reber, U. (2019). Coalitions and counter-coalitions in online contestation: An analysis of the German and British climate change debate. *New Media & Society*, 21(11–12), 2671–2690
- Akhremenko, A. & Petrov, A. (2020). Modeling the protest-repression nexus. Proceedings of the Conference on Modeling and Analysis of Complex Systems and Processes
- Allamong, B. M., Trexler, A., Alqabandi, F., LaChapelle, T., Bail, C. A., Hillygus, D. S. & Volfovsky, A. (2024). Out-numbered online: The consequences of partisan imbalance in online political discussions. OSF Preprint
- Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512
- Baran, Z. & Stoltenberg, D. (2025). Dynamics of scale shift: Contentious places and hybrid activism on social media. *New Media & Society*, 27(9), 5088–5109
- Bennett, W. L. & Segerberg, A. (2013). *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. Cambridge: Cambridge University Press
- Boulianne, S. (2015). Social media use and participation: A meta-analysis of current research. *Information, Communication & Society*, 18(5), 524–538
- Carey, S. C. (2006). The dynamic relationship between protest and repression. *Political Research Quarterly*, 59(1), 1–11
- Castells, M. (2008). The new public sphere: Global civil society, communication networks, and global governance. *The Annals of the American Academy of Political and Social Science*, 616(1), 78–93

- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329(5996), 1194–1197
- Centola, D. (2018). *How Behavior Spreads: The Science of Complex Contagions*. Princeton, NJ: Princeton University Press
- Centola, D. & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734
- Chang, Y.-B., Weng, D. L.-C. & Wang, C.-H. (2021). Personality traits and the propensity to protest: A cross-national analysis. *Asian Journal of Political Science*, 29(1), 22–41
- Chueca Del Cerro, C. (2023). Polarisation and protest mobilisation around secessionist movements: An agent-based model of online and offline social networks. University of Glasgow, PhD Thesis
- Dornschneider-Elkink, S. & Edmonds, B. (2024). Does non-violent repression have stronger dampening effects than state violence? Insights from an emotion-based model of non-violent dissent. *Government and Opposition*, 59(1), 249–271
- Earl, J. & Kimport, K. (2011). *Digitally Enabled Social Change: Activism in the Internet Age*. Cambridge, MA: MIT Press
- Earl, J., Maher, T. V. & Pan, J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science Advances*, 8(10), eabl8198
- Ekman, J. & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22, 283–300
- Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(suppl_3), 7243–7250
- Epstein, J. M. (2014). *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton, NJ: Princeton University Press
- Feldstein, S. (2021). *The Rise of Digital Repression: How Technology Is Reshaping Power, Politics, and Resistance*. Oxford: Oxford University Press
- Freelon, D., McIlwain, C. & Clark, M. (2018). Quantifying the power and consequences of social media protest. *New Media & Society*, 20(3), 990–1011
- Gil de Zúñiga, H., Barnidge, M. & Scherman, A. (2017). Social media social capital, offline social capital, and citizenship: Exploring asymmetrical social capital effects. *Political Communication*, 34(1), 44–68
- Goel, S., Anderson, A., Hofman, J. & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, 62(1), 180–196
- Granovetter, M. & Soong, R. (1988). Threshold models of diversity: Chinese restaurants, residential segregation, and the spiral of silence. *Sociological Methodology*, 18, 69–104
- Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U. & DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2), 115–126
- Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäders, M.-S., Stillman, R. A., Thiele, J. C. & Ayllón, D. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, 23(2), 7
- Hegedüs, D. (2020). Review of *Varieties of Democracy: Measuring Two Centuries of Political Change*. *Perspectives on Politics*, 18(4), 1258–1260

- Howard, P. N. & Hussain, M. M. (2013). *Democracy's Fourth Wave? Digital Media and the Arab Spring*. Oxford: Oxford University Press
- Howarth, J. (2024). Worldwide daily social media users (new 2024 data). Available at: <https://explodingtopics.com/blog/social-media-usage>
- Hu, H.-h., Cui, W.-t., Lin, J. & Qian, Y.-j. (2014). ICTs, social connectivity, and collective action: A cultural-political perspective. *Journal of Artificial Societies and Social Simulation*, 17(2), 7
- Kalinin, K. & de Vogel, S. (2016). Measuring propensity of individual anti-government protest behavior in autocracies. SSRN Working Paper No. 2767663
- Kermani, H. (2025). Introduction: Politics and Twitter in the post-Twitter era. In *Twitter Activism in Iran: Social Media and Democracy in Authoritarian Regimes*, (pp. 1–11). Cham: Palgrave Macmillan
- Kim, H., Whitten-Woodring, J. & James, P. (2015). The role of media in the repression-protest nexus: A game-theoretic model. *Journal of Conflict Resolution*, 59(6), 1017–1042
- Kim, T. & Ogawa, Y. (2024). The impact of politicians' behaviors on hate speech spread: Hate speech adoption threshold on Twitter in Japan. *Journal of Computational Social Science*, 7, 1161–1186
- Knutsen, C. H., Marquardt, K. L., Seim, B., Coppedge, M., Edgell, A. B., Medzihorsky, J., Pemstein, D., Teorell, J., Gerring, J. & Lindberg, S. I. (2024). Conceptual and measurement issues in assessing democratic backsliding. *PS: Political Science & Politics*, 57(2), 162–177
- Kulichkina, A., Righetti, N. & Waldherr, A. (2025a). Protest and repression on social media: Pro-Navalny and pro-government mobilization dynamics and coordination patterns on Russian Twitter. *New Media & Society*, 27(9), 5433–5454
- Kulichkina, A., Waldherr, A. & Wijermans, N. (2025b). An agent-based model of online protest and repression in authoritarian settings. In M. Czupryna, B. Kamiński & H. Verhagen (Eds.), *Advances in Social Simulation*, (pp. 305–319). Cham: Springer
- Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice*, 61, 41–74
- Kuran, T. (1995). *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press
- Lebas, A. & Young, L. E. (2023). Repression and dissent in moments of uncertainty: Panel data evidence from Zimbabwe. *American Political Science Review*, 118(2), 584–601
- Lee, T. E. (2018). The thin blue line between protesters and their counter-protesters. *Journal of Artificial Societies and Social Simulation*, 21(2), 10
- Linz, J. J. (1964). An authoritarian regime: Spain. In E. Allardt & Y. Littunen (Eds.), *Cleavages, Ideologies, and Party Systems: Contributions to Comparative Political Sociology*. Helsinki: The Academic Bookstore
- Lührmann, A., Tannenberg, M. & Lindberg, S. I. (2018). Regimes of the world (RoW): Opening new avenues for the comparative study of political regimes. *Politics and Governance*, 6(1), 60
- Makowsky, M. D. & Rubin, J. (2013). An agent-based model of centralized institutions, social network technology, and revolution. *PLoS One*, 8(11), e80380
- Morozov, E. (2009). From slacktivism to activism. Foreign Policy. Available at: <https://foreignpolicy.com/2009/09/05/from-slacktivism-to-activism/>
- Noelle-Neumann, E. (1974). The spiral of silence: A theory of public opinion. *Journal of Communication*, 24(2), 43–51
- O'Reilly, C., Maher, P. J., Lüders, A. & Quayle, M. (2022). Sharing is caring: How sharing opinions online can connect people into groups and foster identification. *Acta Psychologica*, 230, 103751
- Petrov, A., Akhremenko, A. & Zheglov, S. (2023). Dual identity in repressive contexts: An agent-based model of protest dynamics. *Social Science Computer Review*, 41(6), 2249–2273

- Rød, E. G. & Weidmann, N. B. (2015). Empowering activists or autocrats? The Internet in authoritarian regimes. *Journal of Peace Research*, 52(3), 338–351
- Ross Arguedas, A., Robertson, C. T., Fletcher, R. & Nielsen, R. K. (2022). Echo chambers, filter bubbles, and polarisation: A literature review. Available at: <https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review>
- Ruess, C., Hoffmann, C. P., Boulianne, S. & Heger, K. (2021). Online political participation: The evolution of a concept. *Information, Communication & Society*, 26(8), 1495–1512
- Schäfer, M. S. (2015). Digital public sphere. In G. Mazzoleni, K. G. Barnhurst, K. Ikeda, R. C. M. Maia & H. Wessler (Eds.), *The International Encyclopedia of Political Communication*, (pp. 322–328). London: Wiley Blackwell
- Sohn, D. & Choi, Y.-S. (2023). Silence in social media: A multilevel analysis of the network structure effects on participation disparity in Facebook. *Social Science Computer Review*, 41(5), 1767–1790
- Sohn, D. & Leckenby, J. D. (2007). A structural solution to communication dilemmas in a virtual community. *Journal of Communication*, 57(3), 435–449
- Tarrow, S. (1996). Social movements in contentious politics: A review article. *American Political Science Review*, 90(4), 874–883
- Teorell, J., Torcal, M. & Montero, J. R. (2007). Political participation: Mapping the terrain. In J. W. van Deth, J. R. Montero & A. Westholm (Eds.), *Citizenship and Involvement in European Democracies: A Comparative Analysis*, (pp. 334–357). London & New York: Routledge
- Terren, L. & Borge, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, 9
- Tufekci, Z. (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven & London: Yale University Press
- Verba, S. & Nie, N. H. (1972). *Participation in America: Political Democracy and Social Equality*. New York, NY: Harper & Row
- Waldherr, A. (2014). Emergence of news waves: A social simulation approach. *Journal of Communication*, 64(5), 852–873
- Waldherr, A., Stoltenberg, D., Maier, D., Keinert, A. & Pfetsch, B. (2024). Translocal networked public spheres: Spatial arrangements of metropolitan Twitter. *New Media & Society*, 26(11), 6636–6657
- Waldherr, A. & Wijermans, N. (2017). Modelling the role of social media at street protests. In W. Jager, R. Verbrugge, A. Flache, G. de Roo, L. Hoogduin & C. Hemelrijk (Eds.), *Advances in Social Simulation 2015*, vol. 528. Cham: Springer
- Wijermars, M. & Lokot, T. (2022). Is Telegram a “harbinger of freedom”? The performance, practices, and perception of platforms as political actors in authoritarian states. *Post-Soviet Affairs*, 38(1–2), 125–145
- Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University
- Williams, D. L., Crittenden, V. L., Keo, T. & McCarty, P. (2012). The use of social media: An exploratory study of usage among digital natives. *Journal of Public Affairs*, 12(2), 127–136
- Zeng, J. (2020). #MeToo as connective action: A study of the anti-sexual violence and anti-sexual harassment campaign on Chinese social media in 2018. *Journalism Practice*, 14(2), 171–190