Harassment, civic unraveling, and democratic resilience in the face of technology shocks

M. Amin Rahimian[1[0000−0001−9384−1041]] and Michael Colaresi[2[0000−0001−9574−5723]]

[1] Department of Industrial Engineering, University of Pittsburgh, Pittsburgh PA
[2] Department of Political Science, University of Pittsburgh, Pittsburgh PA

Abstract. In this paper, we focus on the potential fragility of democratic elections given modern information-communication technologies (ICT). Our work provides an explanation for the cascading attrition of public officials recently in the United States from a dynamic system's perspective. We propose that micro-level heterogeneities lead to vulnerabilities of election support systems at the macro-scale. Our analysis provides comparative statistics to measure the fragility of systems against targeted harassment and other adversarial manipulations that are now cheaper to scale and deploy.

Keywords: Online Harassment · Democratic Resilience · Civic Networks.

1 Introduction

Social networks are a critical substrate through which the recognition, adoption, and criticism of innovative technologies flow. Both in-person and digitally mediated communication can spread healthy behavior and democratic engagement but can also be used to target vulnerable communities with false and misleading content that sow discord and mistrust. Recent work has highlighted the role of social networks in spreading vaccine misinformation and violent conspiratorial content that disrupted our public health and national elections during the pandemic [1], [2]. Such concerns are amplified when it comes to the operation of democratic societies that rely on collective decision processes to implement effective policies [3].

We develop a dynamic model that allows for the attrition of election officials who leave their posts in the face of harassment and adverse work conditions. Each official is endowed with a threshold that is a random draw from a distribution of thresholds. The variability in the distribution of thresholds represents the heterogeneity in the officials' tolerance towards harassment, their dedication to the election system and their varying work conditions, e.g., salary and benefits. As some of the most at-risk officials leave the system, the remaining officials face higher pressure from the increased workload and more targeted harassment. We propose to characterize the resiliency of the election system as the proportion of officials that remain with the systems in the face of a fixed or time-varying harassment schedule and identify the factors that influence this resiliency including the structure of support networks among the officials and the state's policy in replacing officials who leave. Our results reveal the existence of a critical harassment level beyond which the election systems cannot survive and this critical value is directly related to the dispersion of the distribution of thresholds. The latter can be modified by offering incentives (e.g., increased salaries) to officials, which we propose to model as an optimal (least cost) transport problem from the existing distribution of threshold to a more resilient one. All of the factors and metrics that we identify with this analysis will be quantifiable with the data collected from administrative records, social media, and our survey instruments.

In the following sections, we address this issue by developing a model of election official harassment. We find that there exists a maximum of harassment that the system can tolerate, related to a proportion of officials that leave the community, based on the distribution of threshold that represents officials' tolerance. Following the basic model, we investigate how a targeting strategy from the attacker affects the model's findings.

1.1 Related Works

The concern about harassment of election officials is extremely important, however; to the best of our knowledge, there is not much research on harassment of election officials before. One of the closest papers is [4]. In this paper, the authors discuss the bribing and harassment of policymakers as opposed to election administrators. The authors in [5] illustrates bandwagons in large and especially close elections and [6] propose a two-stage election model and find the unique equilibrium related to different sample sizes. Other issues that are studied in this context include election fairness [7], candidates' honesty [8] and corruption [9, 10].

2 Modeling Attrition and Unraveling of Civic Networks

We will motivate our model in the context of election systems but the main conclusions apply broadly to civic networks when faced with harassment pressure. In our harassment model, the election officials are facing total harassment of H in their community, the local civic network with election officials, and voters. Political participation is costly [11, 12] and voters would like to participate in high-quality elections [13]. They rely on election officials for the administration of their elections, for information dissemination related to the election, and broadly to ensure the integrity and security of their votes. Each official has tolerance for a specific threshold of harassment that also depends on the support that they receive from other officials in the system. Once their tolerance for harassment is exceeded, the officials will leave, they will vacate their civic posts and we treat this retirement as permanent. We consider a continuum of agents (civic network participants) indexed by the unit interval $\mathcal{I} = (0, 1)$. We micro-found the agents' decisions to stay within the support network as follows. Each agent $i \in \mathcal{I}$ is endowed with a payoff-relevant type θ_i that is an i.i.d. draw from a distribution F. A high type represents increased utility that the agent receives from remaining within their civic network and the heterogeneity in θ_i values will imply the agents' differing tolerances for harassment. In particular, we assume that each agent i at time t has the option to either stay in the network $a_{i,t} = 1$ or leave $a_{i,t} = 0$. Given a harassment level H_t and assuming zero utility for the outside option $a = 0$, we can formulate the utility of agent i at time t as follows:

$$U_{i,t}(a) = a(\theta_i(1 - p_t) - H_t), a \in \{0, 1\}, \tag{1}$$

where $0 \leq p_t \leq 1$ is the fraction (Lebesgue measure on the unit interval) of the agents who have left the system at time t. Equation (1) can be interpreted as a threshold on the "experienced" harassment, $H_t/(1 - p_t)$, such that:

$$a_{i,t} = \arg\max_{a \in \{0,1\}} U_{i,t}(a) = \begin{cases} 0, & \text{if } \theta_i \leq H_t/(1 - p_t), \\ 1, & \text{otherwise.} \end{cases} \tag{2}$$

As more agents leave the civic network their positive externality on the remaining agents is lost, which in turn causes more departures and threatens a civic unraveling of the support network. The dynamics of the ensued unraveling are determined by distribution F, and we can use the distributional parameters to investigate the asymptotic fraction of agents that remain with the system. Another way to interpret the threshold behavior in (2) is to assume that the total amount of the harassment input to the system is equally distributed such that as agents leave the civic network, the harassment experienced by the remaining agents increases proportionally to weight of the remaining agent, i.e., $H_t/(1-p_t)$. The remaining agents tolerate their increasingly experienced harassment until their threshold is exceeded, $\theta_i \leq H_t/(1 - p_t)$, at which point they leave the network, $a_{i,t} = 0$. This situation is analogous to pre-social media harassment occurring at board meetings attended by election officials.

Initially ($t = 0$), all election official whose thresholds θ are less than $h_0 = H/1 = H$ will be removed from the community. Because of the heterogeneity of the thresholds, some proportion of officials p_0 will leave the community. Specifically, $p_0 = F(h_0)$, where F is the cumulative distribution function (CDF) of the thresholds. Once p_0 of the officials leave, the harassment faced by the remaining officials will increase based on the assumption. To be specific, at time $t = 1$, the harassment that the remaining officials experience is $h_1 = H_0/(1 - p_0)$. Since the tolerated harassment is increased, some additional proportion of officials will leave, which can be similarly expressed as a function of the distribution F.

Let $p_\infty(H) = \lim_{t\to\infty} p_t$ be the proportion that has left the system, when the process converges under H: the process will continue until no officials leave ($p_\infty(H) < 1$), or until all officials have left ($p_\infty(H) = 1$). The dynamic and the equilibrium are as follows:

$$p_{t+1} = F(\frac{H}{1 - p_t}), \text{ and } p_\infty(H) = F(\frac{H}{1 - p_\infty(H)}). \tag{3}$$

Given $H > 0$, the process starts with $p_0 = 0$. The support of θ is any positive number, so $F(x) = 0, \forall x \leq 0$. Note that $p_t \in [0, 1]$ and $p_\infty(H) = 1$ is always a fixed point of (3).

Based on the previous models, we attempt to find the relationship between the distribution of election officials' thresholds F and $p_\infty(H) = \lim_{t\to\infty} p_t$ in the model through asymptotic analysis. This allows to measure the the robustness of the system to harassment in terms of the distributional features of F.

Theorem 1 (Monotonicity and Convergence). *For any H and distribution F, the sequence $\{p_t\}$ is bounded, monotone increasing and converges to a limit $p_\infty(H)$.*

2.1 A New Resilience Metric

Based on (3), two questions arise: (1) What is the maximum total amount of harassment H^* that the community can tolerate without all election officials leaving? (2) What is the fraction of election officials that leaves, $p(H^*)$, if the total amount of harassment is H^*? This gives rise to the following definition of resilience:

Definition 1 (Resilience). *The Resilience of a Community with a threshold distribution, F, is defined as $\mathcal{R}(F) = \inf\{H > 0 \mid p_\infty(H) = 1\}$.*

It is clear that resilience is determined by the distribution function F. Note that for every $H > 0$, $p_\infty(H) = 1$ is always a fixed point of (3). Hence, if there are no roots in the interval $[0, 1)$, then $p_\infty(H) = 1$. On the other hand, if there exist roots that are strictly less than 1, the sequence $\{p_t\}$ under H must converge to the smallest root. Therefore, the goal is to find the smallest H such that (3) has no root in the interval $[0, 1)$. The following theorem gives us an expression of the Resilience $R(F)$ in terms of the quantile function Q of F:

Theorem 2. *Assume there exists a quantile function $Q = F^{-1}$, then the resilience of the harassment model is given by $\mathcal{R}(F) = \sup_{p\in(0,1)}(1 - p)Q(p)$.*

Stochastic Ordering of the Threshold Distributions Recall that given two distributions of thresholds whose cumulative distribution functions are F_1 and F_2, we say F_2 stochastically dominates F_1, if for all threshold values θ, $F_2(\theta) \leq F_1(\theta)$ with inequality strict for at least one value of θ. Second-order stochastic dominance is a relaxation of this definition that only requires $\int_{-\infty}^{\theta} [F_1(t) - F_2(t)]\, dt \geq 0$ for all θ with strict inequality at some θ. First-order stochastic dominance implies second-order dominance. our following result shows the implication of stochastic ordering of the distribution of threshold on resilience.

Corollary 1. *Given two distributions of thresholds F_1 and F_2, if F_1 stochastically dominates F_2 (second-order dominance suffices), then distribution F_1 can tolerate more harassment than distribution F_2 meaning that $\mathcal{R}(F_1) \geq \mathcal{R}(F_2)$.*

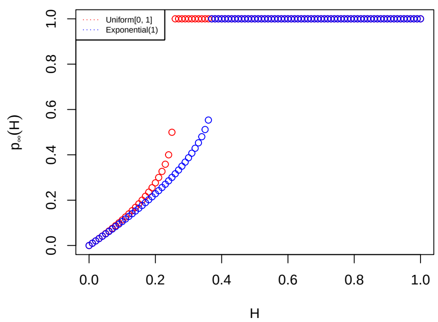

Fig. 1. We plot $p_\infty(H)$ as a function of H for $F_1 \sim$ Uniform$[0, 1]$ in red, $F_2 \sim$ Exponential(1) in blue, whose corresponding resilience are $1/4$ and $1/e$, respectively. Note the stochastic dominance relationship between these distributions: $F_2 \succ F_1$ which is reflected in their resilience metrics: $\mathcal{R}(F_2) = 1/e > \mathcal{R}(F_1) = 1/4$. We can compute $\mathcal{R}(F)$ explicitly as a function of distribution parameters, for example, $\mathcal{R}(F) = b^2/(4(b-a))$ for $\theta \sim$ Uniform$[a, b]$, whereas for exponential distribution $\mathcal{R}(F) = 1/(e\lambda)$ for $\theta \sim \text{Exp}(\lambda)$. One particular value of interest would be $p_\infty(\mathcal{R}(F))$: the fraction remaining in the system at the critical harassment value $\mathcal{R}(F)$.

3 Increased fragility from targeted harassment

A technology shock may enable adversaries to target at-risk officials at every time step and thus allocate the input harassment H more efficiently, e.g., by surveilling social media. If the adversary can observe the individual thresholds, then the adversary's optimum strategy at each time step is to target the maximum threshold of officials such that after distributing the net harassment H among the targeted group they all leave — the highest threshold of officials who can be included in the harassment target and still leave the population even after the net harassment H is distributed among the targets. In particular, at time zero, $\theta_0 = 0$, and at time one all officials with threshold less than $\theta_1 = \max\{\theta \geq \theta_0 = 0 : \theta \leq H/(F(\theta) - F(0)) = H/F(\theta)\}$ will be targeted and removed. At any following time step, $t+1$, the adversary will target all officials whose thresholds are less than θ_{t+1}, where θ_{t+1} is the largest number $\theta \geq \theta_t$ such that $\theta \leq H/(F(\theta) - F(\theta_t))$: $\theta_{t+1} = \max\{\theta \geq \theta_t : \theta \leq H/(F(\theta) - F(\theta_t))\}$.

This dynamic can be expressed as follows (in terms of the targeted thresholds θ_t or the proportions who have left the system p_t):

$$\theta_{t+1} = \frac{H}{F(\theta_{t+1}) - F(\theta_t)}, \text{ and } p_{t+1} = F(\frac{H}{p_{t+1} - p_t}). \tag{4}$$

It is clear that (4) is an extension of (3) and we can show following the above targeting strategy everyone in the community will leave after a finite time:

Theorem 3. *For all continuous distributions F under the targeted harassment model, we have $\mathcal{R}(F) = 0$.*

4 Conclusion

In this article, we study democratic resilience when civic networks are targeted with harassment, for example, in the case of election administrators. Even if the total harassment is constant and equally distributed, as some election officials leave the network those who remain in the system will experience increasing harassment which can result in a cascading failure that unravels the civic network (e.g., removing public support for a policy). This unravelling may be only partial if the input harassment is small enough. We characterize the resilience as the least harassment to cause total unravelling and relate it to network heterogeneity and distributional features. We also consider more sophisticated adversaries who can target their harassment and show the resultant fragility of the system (total unravelling) in the face of targeted attacks.

References

1. C. O'Connor, J. O. Weatherall, *The misinformation age: How false beliefs spread* (Yale University Press, 2019).
2. C. S. B. Ngai, R. G. Singh, L. Yao. Impact of COVID-19 vaccine misinformation on social media virality: Content analysis of message themes and writing strategies. *J Med Internet Res* **24**, e37806 (2022).
3. M. Colaresi. How our misunderstanding of the digital and computing revolutions puts democracy at risk (and what to do about it). *Critical Quarterly* **62**, 70–80 (2020).
4. E. Dal Bo, P. Dal Bo, R. Di Tella. "plata o plomo?": Bribe and punishment in a theory of political influence. *American Political Science Review* **100**, 41–53 (2006).
5. A. Leontiou, G. Manalis, D. Xefteris. Bandwagons in costly elections: The role of loss aversion. *Journal of Economic Behavior & Organization* **209**, 471–490 (2023).
6. C. Andonie, C. Kuzmics. Pre-election polls as strategic coordination devices. *Journal of Economic Behavior & Organization* **84**, 681–700 (2012).
7. E. Berman, M. Callen, C. C. Gibson, J. D. Long, A. Rezaee. Election fairness and government legitimacy in afghanistan. *Journal of Economic Behavior & Organization* **168**, 292–317 (2019).
8. J. Woon, K. Kanthak. Elections, ability, and candidate honesty. *Journal of Economic Behavior & Organization* **157**, 735–753 (2019).
9. A. Apffelstaedt, J. Freundt, C. Oslislo. Social norms and elections: How elected rules can make behavior (in)appropriate. *Journal of Economic Behavior & Organization* **196**, 148–177 (2022).
10. M. Le Moglie, G. Turati. Electoral cycle bias in the media coverage of corruption news. *Journal of Economic Behavior & Organization* **163**, 140–157 (2019).
11. A. Downs. An economic theory of political action in a democracy. *Journal of Political Economy* **65**, 135–150 (1957).
12. J. Ledyard. The pure theory of large two-candidate elections. *Public Choice* **44**, 7–41 (1984).
13. M. Mochtak, C. Lesschaeve, J. Glaurdić. Voting and winning: perceptions of electoral integrity in consolidating democracies. *Democratization* **28**, 1423–1441 (2021).