#change*

Deepal Basak
Indiana University

Joyee Deb
New York University

Aditya Kuvalekar
University of Essex

October 2, 2023

Abstract

Change may be socially beneficial, but causing change requires enough people to take a costly action—a classic collective action problem. More similar information about the benefits of regime change may help people coordinate and mobilize toward a shared goal, but it can also exacerbate free-riding. We propose a notion of information similarity to characterize when more similar information hinders or facilitates regime change. We show that more similar information helps overthrow stronger regimes but enables weaker regimes to survive. We apply our framework to collective action settings such as protests and voting in committees.

Keywords: collective action, information similarity, regime change

JEL Codes: D82, D83

*We are grateful to Nageeb Ali, Marco Battaglini, Steve Callander, Mehmet Ekmekci, Albin Erlanson, Daniel Garrett, Elliot Lipnowski, Meg Meyer, Doron Ravid, Phil Reny, Vasiliki Skreta, Yu Fu Wong, and conference audiences and seminar participants at University of Chicago, Collegio Alberto, Essex University, Kelley School of Business, Oxford University, Penn State University, Stony Brook Game Theory meeting, SITE Political Economy, University College London, University of Warwick, and Yale University for their helpful comments and suggestions. Basak: Indiana University, Kelley School of Business, email: dbasak@iu.edu; Deb: New York University, email: joyee.deb@nyu.edu, Kuvalekar: University of Essex, email: a.kuvalekar@essex.ac.uk.

The world has experienced more political uprisings in the past few years than ever before. In 2019 alone there were protests on every continent and across 114 countries—-from Hong Kong to Haiti, Bolivia to Britain... Smartphones and social media have revolutionised the way in which demonstrations are organised, advertised and sustained.... Mass protests have existed for centuries. But their increased frequency may be here to stay.

From "Political Protests Have Become More Widespread and More Frequent," *Economist*, March 2020.

1. Introduction

The popular press and academic discourse in political science and sociology suggest that mass protests have become more frequent. A recent report[1] writes, "We are living in an age of global mass protests that are historically unprecedented in frequency, scope, and size.... The size and frequency of recent protests eclipse historical examples of eras of mass protest, such as the late-1960s, late-1980s, and early-1990s." Social media is often claimed to be the primary driver behind this trend. The narrative is that increased connectedness—via social media, viral videos, or the internet—has made it easier for geographically dispersed groups to see similar information, arrive at a common understanding of the need for social change, and mobilize toward a shared goal. For example, in an interview with the *New York Times*, Omar Wasow says, "Social media radically simplified organizing and coordinating large groups.... Part of what social media does is allow us to see a reality that has been entirely visible to some people and invisible to others."[2]

The intuition that access to the same information helps people coordinate actions to achieve a common goal is compelling at first glance. But participating in a political protest is costly, and the benefits of regime change often accrue to everyone regardless of participation. In such a classic instrumental collective action problem (see Tullock (1971), Olson (1965)), access to similar information about others' participation could also cause people to free-ride. Indeed, in a recent experiment about mass protesters in Hong Kong, Cantoni et al. (2019) demonstrate the strategic substitutability of participants: citizens reduce their participation in a protest when they learn about others' intention to participate.

With both coordination and free-riding at play, similarity of information can have ambiguous effects on the size and success of protests. Potential participants face both fundamental uncertainty about the benefits of a regime change and strategic uncertainty about other participants' behavior. Similarity of information can mitigate strategic uncertainty, but this, in turn, can have two opposing

[1] Source: "The Age of Mass Protests," published by the Center of Strategic and International Studies, a Washington-based think tank.

[2] See https://www.nytimes.com/2020/06/18/technology/social-media-protests.html.

effects: On the one hand, it can help coordination because, now, an agent who is willing to participate because she believes a regime change is beneficial based on her information also believes that other agents think and behave similarly. On the other hand, it can exacerbate free-riding: a high likelihood of many agents' participating reduces the probability that one more person will make a difference, dampening the individual incentive for costly participation. This trade-off between aiding coordination and exacerbating free-riding due to increased similarity of information is at the heart of our paper.

We study this trade-off in a canonical regime-change game with incomplete information in which agents receive information about an unknown state of the world and decide whether to participate in a protest. Participation is costly, and successful regime change requires a minimum number of participants. We make three contributions. First, we propose a stochastic order to compare the similarity of joint distributions. Second, using this order, our main results characterize when increased similarity of information leads to larger or smaller protests. It turns out that even in a regime-change game featuring both coordination and free-riding, it is possible to partition the setting into two economically interpretable sub-environments: *encouragement environments*, in which greater turnout encourages participation by making it more likely that an individual can make a difference, and *discouragement environments*, in which greater turnout discourages participation by making it less likely that an agent can make a difference. We show that more similar information increases participation in encouragement environments and can decrease it in discouragement environments.

Our third contribution is a new and potentially testable insight that follows from the characterization: When agents consume more similar information, sufficiently resilient regimes—ones that require a large number of protesters to be overthrown—face larger protests. In contrast, increasing information similarity reduces the size of protests against weak regimes. In other words, as information becomes more similar, strong regimes become more vulnerable to the threat of regime change and weak regimes become less vulnerable. This result suggests that casual empirical observation about increased size and frequency of mass protests may be biased. While popular discourse has highlighted how a more connected world has enabled larger protests against regimes previously thought to be impregnable, it has not focused as much on the counterfactual events—how increased information similarity may have hindered protests with more modest goals.

Mass protests are not the only setting in which instrumental collective action problems arise under uncertainty. We apply our framework to voting in committees and show how increased sharing of information among committee members affects their voting given a fixed voting rule, and furthermore we show how optimal voting rules vary with increased information similarity.

Next, we describe our model and results more formally. There is a binary, unknown state of the world: good or bad. Regime change is socially efficient if the regime is bad, but not otherwise. Society comprises two geographically dispersed groups of agents who all have the same preferences but do not have access to the same information. Each group receives a signal about the state of the world. Agents observe only the signal received by their own group and must decide individually whether to participate in an action that can contribute toward changing the regime. Participating is costly, but the payoffs from a successful regime change accrue to all agents, not just the participants. A minimum threshold number of participants must be exceeded to cause a regime change. We allow agents to be uncertain about both the size of the groups and the participation threshold. In this setting, we ask whether increased similarity of the groups' signals increases participation.

We compare similarity of information using an order we call *Concentration Along Diagonal* (CAD).[3] It says that information is more similar when, conditional on observing signal x, an agent believes it is more likely that others also observed the signal x and less likely that others observed a signal different from x.

In a regime-change game, an agent's incentive to participate depends (non-monotonically) on how many others participate. For example, if she believes her participation can make a difference, then she has a strong incentive to join. However, if she expects the number of participants to exceed the threshold, then she will want to free-ride instead. We show that more similar information in the CAD order increases (decreases) participation in encouragement (discouragement) environments. Formally, for any strategy profile, we define the participation (non-participation) set as the set of signals for which agents participate (do not participate). Given any information structure, our game can admit multiple equilibria.[4] We show that more similar information enlarges the participation set in the equilibrium with maximal participation in the encouragement environment (Theorem 1) and shrinks it in the discouragement environment (Theorem 2).

To understand the intuition, consider a symmetric equilibrium in which agents participate if and only if they observe a signal in set P. If information similarity increases, an agent with a signal in P (not in P) assigns a higher probability to the possibility that everyone else sees the same signal and therefore everyone (no one) participates. In the encouragement environment, this reduction in strategic uncertainty increases (decreases) the incentive to participate for an agent who observes a signal in P (not in P). But in the discouragement environment, this reduction in strategic uncertainty is counteracted by the fact that if an agent is confident that many others will participate, then she is discouraged from participating because she is less likely to make a difference. Thus, more similar information helps in the encouragement environment but can hurt in the discouragement environment. We provide a sufficient condition on the information structure under which more similar information hurts in the discouragement environment.

[3] In two dimensions, this order is equivalent to one proposed by Meyer (1990).

[4] An equilibrium with no participation (regardless of signal) always exists. Non-monotone equilibria can also arise.

It is worth highlighting that existing literature often restricts attention to monotone, cut-off strategies for tractability. While it is known that in a game of pure complementary, the agents play a symmetric cutoff strategy in the maximal equilibrium, this is not true in a game that involves both complementarity and substitutability. Hence, we do not impose such restrictions on strategies—allowing non-monotonic and asymmetric strategies. We are still able to derive a characterization by focusing on how the equilibrium set changes with similarity of information. Moreover, to derive the characterization we propose the appropriate order of similarity of information, namely the CAD order, which partitions the environment into two economically interpretable sub-environments.

Equipped with our characterization, we consider two applications. In our first application, we analyze the effect of increased information similarity on different types of regimes: resilient regimes, which require a large number of protesters to cause a regime change, and weak regimes, which fall with even a small number of protesters. We show that under an intuitive single-crossing condition, in sufficiently resilient regimes, increased information similarity increases the maximal equilibrium turnout. However, for weak regimes, increased information similarity can have the opposite effect and reduce maximal turnout. The result suggests that the emergence of high-visibility protests against invincible regimes does not imply an *unconditional* increase in the size and frequency of mass protests in the age of more similar information. Increased information similarity made the observed protests bigger, but may also have prevented some protests from happening at all.

As a second application, we study the effect of increased information similarity on the incentives to vote in committees. Consider the non-executive board of a company deciding whether to replace the CEO. The existing CEO is either competent or incompetent. Board members have their own personal experiences with the CEO and must vote for or against the change. Effecting a change requires exceeding a minimum threshold of negative votes. The management literature documents that voting negatively is costly in such settings because voting is typically by open ballot and dissenting board members face retaliation from the management. The difference from our benchmark model is we no longer have two groups with population uncertainty but rather a fixed number ($N > 2$) of committee members. This requires us to extend our notion of information similarity to compare random variables of more than two dimensions. Specializing to binary signals, we establish results analogous to Theorems 1 and 2.[5] We also study optimal voting rules—rules that maximize the probability of replacing the CEO conditional on their being incompetent—and characterize how the optimal rule varies with increased information similarity. If the current threshold rule is sufficiently (not sufficiently) demanding, it is optimal to make it less (more) demanding when information similarity increases.

[5]Unlike Theorems 1 and 2, however, with more than two groups, the characterization is only partial, i.e., not every environment can be classified as encouragement or discouragement. In Section 6, we provide more details.

Finally, we discuss some extensions. For instance, we study a setting in which regime change does not exogenously occur with sufficient participation: Rather, citizens use protests as a tool to convey their private information to policymakers, and policymakers use the observed turnout to *infer* the state of the world and then endogenously decide whether to change the status quo.[6] We use our framework to study how increased similarity of information affects the informational content of turnout in such settings. We consider also questions of information design and extensions with state-dependent changes in information similarity.

1.1. Related Literature

Our paper relates to the literature on the value of information, which started with Hirshleifer (1971) and includes work by Morris and Shin (2002), Angeletos and Pavan (2007), and Bergemann and Morris (2013), who study a class of games with monotone best responses (pure complementarity or substitutability) in which agents receive private signals. They ask when new public information improves welfare. We explore two new directions. First, we study not the effect of new or more information, but rather the effect of *increased similarity* of information. Given the rise of social media, understanding how information similarity affects incentives in a strategic environment is a question of first-order importance. We consider an arbitrary signal structure and propose the CAD as a natural order of information similarity. We do not use correlation as a measure of similarity since it is not invariant to relabeling of signals. The existing literature contains other measures of the interdependence of joint distributions (e.g., Müller and Stoyan (2002), Meyer and Strulovici (2012)), but none are appropriate for comparing the *conditional* belief distributions that arise naturally in strategic settings with incomplete information.[7] For the bivariate case (the focus of this paper), the CAD order is the same as that in Meyer (1990). Awaya and Krishna (2022) study the effect of the interdependence of signals on common learning as do Cripps et al. (2008). They show that essentially any interdependence obstructs common learning. Second, we study a new, important game: the classic collective action game (Olson (1965), Tullock (1971)), where agents have both coordination and free-riding motives, making the best response non-monotonic. There can be multiple equilibria. The standard tools of iterated elimination of strictly dominated strategies cannot be extended to non-monotone best responses. Nevertheless, we provide a characterization using the maximal equilibrium.

We also contribute to the sizable literatures on protests and voting. For instance, Manacorda and Tesei (2020) provide empirical evidence that mobile phones facilitated protests in Africa. Edmond (2013) considers a theoretical model of

[6]Some recent papers (for example, Battaglini (2017) and Ekmekci and Lauermann (2019)) focus on this informational role of turnout.

[7]In a companion paper, we study similarity orders in general and explain why existing orders do not suffice for studying conditional beliefs in games with strategic uncertainty.

protest in which the regime can manipulate information. These papers consider protest as a game of pure complementarity. However, Cantoni et al. (2019) demonstrate, in a recent experiment about the mass protesters in Hong Kong, that the free-riding incentive is important in practice. Some recent papers—e.g., Shadmehr (2021), Dziuda et al. (2021), and Park and Smyrniotis (2022)—incorporate free-riding and construct cutoff equilibria under some restrictive assumptions. Dziuda et al. (2021) show that when agents have the incentive to free-ride, a lower required participation threshold might not increase the likelihood of change. The literature on voting focuses on questions of information aggregation (e.g., Battaglini (2017) and Ekmekci and Lauermann (2019)) or the optimal voting rule (e.g., Kattwinkel and Winter (2023)). Chemmanur and Fedaseyeu (2018) present a model of voting on a corporate board that closely resembles our application. These papers do not consider changing information environments, which is the central focus here.

2. A Regime-Change Game

There are two states of the world: $\boldsymbol{\theta} \in \Theta = \{0, 1\}$. In $\boldsymbol{\theta} = 1$, the regime is bad and it is socially beneficial to change the regime. In $\boldsymbol{\theta} = 0$, the regime is good and it is not beneficial to change the regime. Society consists of G groups, $g \in \mathcal{G} := \{1, 2, \ldots, G\}$. We introduce population uncertainty à la Myerson (1998): the number of agents in any group g, denoted by $\mathbf{N}_g$, is a $\mathbb{Z}_+$-valued random variable with probability mass function $\eta(\cdot)$ and mean N. Agents do not observe the size of their own or other group. We let $\eta^A(\cdot)$ denote the conditional probability mass function of $\mathbf{N}_g - 1$ according to an agent in group g.[8] Throughout this paper, we consider $G = 2$; the analysis for more than two groups is relegated to the online appendix.

Each agent decides whether to take a costly action: participating (choosing $a = 1$) costs $c > 0$, while not participating ($a = 0$) is costless. Regime change occurs only if enough agents participate. We assume that the threshold participation required to change the regime is an $\mathbb{N}$-valued random variable, $\bar{\boldsymbol{n}} + 1$, where $\bar{\boldsymbol{n}}$ follows probability mass function $\phi(\cdot)$, and agents do not observe the realization of $\bar{\boldsymbol{n}} + 1$. We call $\bar{\boldsymbol{n}}$ the *resilience of the regime*. We allow population uncertainty and a random (unobserved) resilience because this is more realistic in many regime-change settings. However, in Section 5 we present an application with observable and deterministic group sizes and participation threshold.

Note that we assume that agents take participation decisions simultaneously, thus abstracting from dynamic considerations that may be important in many protest settings. Examples of our setting would be protests that are scheduled on a particular day, e.g., the women's march in the US on Jan 21st, 2017 or the

[8] In general, $\eta^A(\cdot)$ can be different from $\eta(\cdot)$ because an agent may be more or less optimistic about the size of her group, conditional on belonging to the group herself, compared to what someone outside the group believes about the group size. As an example, if $\eta(\cdot)$ were drawn from a Poisson distribution, then $\eta^A(\cdot)$ and $\eta(\cdot)$ would coincide.

farmer protests in India on Republic Day. Like our model, these examples also feature agents that are geographically dispersed but have common interests.

We summarize the payoffs in the matrix below, in which $\bar{n}+1$ denotes the realized resilience (threshold required for regime change) and A denotes the number of agents who participate.

	$A \geqslant \bar{n}+1$	$A \leqslant \bar{n}$
$a=1$	$\theta - c$	$-c$
$a=0$	θ	0

2.1. Information Structure

Before deciding whether to participate, agents receive information about the state of the world. Each group receives a signal $\mathbf{X}_g$ drawn from a finite ordered set $\mathcal{X}$, and every agent in group g observes only the signal received by the group. Let $\mathbf{X} := (\mathbf{X}_g)_{g \in \mathcal{G}}$. We denote the joint distribution of $(\boldsymbol{\theta}, \mathbf{X})$ by $\mathscr{P}(\cdot) \in \Delta(\Theta \times \mathcal{X}^G)$, and the distribution of $\mathbf{X}$ conditional on $\boldsymbol{\theta} = \theta$ by $\mathscr{P}^\theta \in \Delta(\mathcal{X}^G)$. Let $\mu_1 := \mathscr{P}(\{\boldsymbol{\theta} = 1\})$. We denote by $\mathbb{P}_g^\theta \in \Delta(\mathcal{X})$ the distribution of $\mathbf{X}_g$ given that $\boldsymbol{\theta} = \theta$. We assume that the distribution of $\mathscr{P}^\theta(\cdot)$ is exchangeable so that we have $\mathbb{P}_g^\theta = \mathbb{P}_{g'}^\theta =: \mathbb{P}^\theta$ for all $g, g' \in \mathcal{G}$ and all $\theta \in \Theta$.

Let $\mu(x) := \mathscr{P}(\{\boldsymbol{\theta} = 1\} | \{\mathbf{X}_g = x\})$ denote the posterior probability that any agent in group g assigns to the state's being 1 given a realized signal x. We assume that $\mu : \mathcal{X} \to [0,1]$ is injective. Given group g, we let $\mathbf{X}_{-g}$ be the random variable denoting the signal of the other group. Let $\mathscr{P}_x^\theta \in \Delta(\mathcal{X})$ denote the conditional distribution of $\mathbf{X}_{-g}$ given state $\boldsymbol{\theta} = \theta$ and $\mathbf{X}_g = x$. Since $\mathbb{P}^\theta$ is exchangeable, $\mathscr{P}_x^\theta(\cdot)$ is independent of the group g.

2.2. Strategies and Turnout

Strategies: A (pure) strategy of agent i from group g is a mapping,

$$\sigma_g^i : \mathcal{X} \to \{0,1\}.$$

Given a strategy profile $\sigma = (\sigma_1, \sigma_2)$, we define the *participation set of σ_g for group g*, denoted by $P(\sigma_g)$, to be the set of signals such that $\sigma_g(x) = 1$. Analogously, we define the *nonparticipation set of σ*, denoted by$NP(\sigma_g) := \mathcal{X} \backslash P(\sigma_g)$. When the dependence on σ_g is obvious, we denote $P(\sigma_g)$ and $NP(\sigma_g)$ by P_g and NP_g, respectively.

Turnout: Since group size is random, we define $\mathbf{A}$ to be the random variable corresponding to the total number of agents who participate, given strategy profile σ:

$$\mathbf{A} := \sum_{g=1}^{G} \mathbf{N}_g \sigma_g(\mathbf{X}_g)$$

We call $\mathbf{A}$ the turnout. Notice that $\mathbf{A}$ is $(\mathbf{N}_1, \mathbf{N}_2, \mathbf{X})$-measurable and depends on σ. When this is obvious, we suppress σ. All agents in group g receive the same signal and have the same belief. Let $\mathbf{A}_{-g}(x_g, x_{-g}; \sigma)$ be the turnout according to an agent in group g, excluding herself. Then we have

$$\begin{aligned}\mathbf{A}_{-g}(x_g, x_{-g}; \sigma) &:= \mathbf{N}_{-g}\sigma_{-g}(x_{-g}) + (\mathbf{N}_g - 1)\sigma_g(x_g) \\ &= \mathbf{N}_{-g}\mathbb{1}_{x_{-g}\in P_{-g}} + (\mathbf{N}_g - 1)\mathbb{1}_{x_g \in P_g}.\end{aligned}$$

Expected turnout in $\boldsymbol{\theta} = 1$: For any σ, define

$$\mathscr{V}(\sigma) := \mathbb{E}[\mathbf{A}(\mathbf{X}; \sigma) | \boldsymbol{\theta} = 1] \tag{1}$$

to be the expected turnout in state 1, which is the state in which it is beneficial to change the regime. For a fixed σ, $\mathscr{V}(\sigma)$ depends only on the marginal distribution of the signals and not on the joint distribution. So information similarity affects $\mathscr{V}$ only by affecting the equilibrium σ. We state this in the following lemma. The proof is in the appendix. With some abuse of notation, let $\mathbb{P}^1(S)$ denote $\mathscr{P}(\mathbf{X}_g \in S | \boldsymbol{\theta} = 1)$.

Lemma 1: *For any σ, with associated participation sets $(P_g)_{g\in\mathcal{G}}$,*

$$\mathscr{V}(\sigma) = N \sum_{g=1}^{G} \mathbb{P}^1(P_g).$$

Therefore, when $G = 2$, we can say $\mathscr{V}(\sigma) = 2[\mathbb{P}^1(P_1 \cup P_2) + \mathbb{P}^1(P_1 \cap P_2)]$.

Solution concept: We consider Bayes Nash equilibria in pure strategies. Multiple equilibria may exist, including one in which no one participates regardless of the signal. We do not impose any additional structure on the equilibrium, such as monotone or symmetric strategies. Let $\mathcal{E}(\mathscr{P})$ be the set of strategy profiles that constitute an equilibrium under information structure $\mathscr{P}$.

Given multiple equilibria, we focus on how increased similarity affects the maximal possible participation in any equilibrium. Accordingly, we define the following.

Definition 1 (Maximal Participation Equilibrium and Maximal Equilibrium Turnout): *We say that an equilibrium σ^* is a maximal participation equilibrium if $\mathscr{V}(\sigma^*) \geqslant \mathscr{V}(\sigma)$ for all $\sigma \in \mathcal{E}$. Let $\mathscr{V}^*(\mathscr{P})$ denote the expected turnout (in state 1) in the maximal participation equilibrium given information structure $\mathscr{P}$, and call it the maximal equilibrium turnout.*

Unlike in several regime-change games featuring only strategic complementarities (Morris and Shin (2002) e.g.), best responses in our game are not monotonic in aggregate participation. Therefore, it is not clear whether equilibria can be ordered in any natural way. This means we cannot use existing tools—such as those used in supermodular games—directly.

Both the expected turnout given a strategy $\mathscr{V}(\sigma)$ and the maximal equilibrium turnout for an information structure $\mathscr{V}^*(\mathscr{P})$ are defined *conditional on* $\boldsymbol{\theta} = 1$; that is, when change is beneficial. For brevity, henceforth, we do not mention this explicitly. $\mathscr{V}(\cdot), \mathscr{V}^*(\cdot)$ also depend on other parameters, such as $\bar{\boldsymbol{n}}$. We typically suppress this dependence and only make the dependence on the information structure explicit.[9,10]

2.3. A Measure of Information Similarity

Given our central question, we need a notion of informational similarity. Agents make participation decisions based on their beliefs about the state of the world and the expected turnout. So, they must reason about the *conditional* probability of others' information given their own. We define a similarity order for two-dimensional random variables using such conditional beliefs.

Definition 2 (Concentration Along Diagonal (CAD)): *Let $\mathcal{Y} \subset \mathbb{R}$ be a finite set. Let Y and $\widehat{Y}$ be two $\mathcal{Y}^2$-valued random variables whose distributions are given by $\mathscr{D}$ and $\widehat{\mathscr{D}}$, respectively. We say Y is more similar than $\widehat{Y}$ in the CAD order, denoted by $Y \succcurlyeq_{CAD} \widehat{Y}$ or $\mathscr{D} \succcurlyeq_{CAD} \widehat{\mathscr{D}}$, if the following two conditions hold.*

1. *Y_i and $\widehat{Y}_i$ are identically distributed for all $i \in \{1, 2\}$.*
2. *For $y \in \mathcal{Y}$ and $T \subseteq \mathcal{Y}$,*
   (a) $\mathscr{D}(Y_2 \in T | Y_1 = y) \geqslant \widehat{\mathscr{D}}(\widehat{Y}_2 \in T | \widehat{Y}_1 = y)$ if $y \in T$.
   (b) $\mathscr{D}(Y_2 \in T | Y_1 = y) \leqslant \widehat{\mathscr{D}}(\widehat{Y}_2 \in T | \widehat{Y}_1 = y)$ if $y \notin T$.

Notice that, by the exchangeability of the distributions, we can interchange Y_1 and Y_2 in the definition.

The CAD order captures the intuitive notion that when information similarity increases, any agent believes that it is more likely that others received the same signal as her. A natural interpretation is that in a more connected world, with the possibility of information going viral, when an agent sees a post on social media, she assigns a higher probability to the event that others, even in distant places, have received exactly the same information. This definition of the CAD order for two-dimensional random variables is equivalent to an order due to Meyer (1990). The lemma below derives a useful property of the CAD order.

Lemma 2: *If $\mathscr{D} \succcurlyeq_{CAD} \widehat{\mathscr{D}}$, then, for every $T \subseteq \mathcal{Y}$, $\exists\ \alpha_T \geqslant 0$ such that $\mathscr{D}(T, T) = \widehat{\mathscr{D}}(T, T) + \alpha_T$, and $\mathscr{D}(T, \mathcal{Y} \backslash T) = \widehat{\mathscr{D}}(T, \mathcal{Y} \backslash T) - \alpha_T$.*

The proof of the lemma is in the appendix. Formally, we use the CAD order to compare $\mathscr{P}_x^\theta(\cdot)$, the beliefs of players conditional on a state and a realized signal,

[9] We focus on turnout in the state in which regime change is beneficial (state $\boldsymbol{\theta} = 1$). Changing a regime when it is not beneficial to do so ($\boldsymbol{\theta} = 0$) entails no additional costs over and above the costs of participation. We relax this assumption in Section 6.

[10] For some applications, the object of interest could be maximal welfare of protesters instead of maximal turnout. We discuss this in Section 6.

keeping the marginal beliefs conditional on $\boldsymbol{\theta} = \theta$ by $\mathbb{P}^\theta$ unchanged,[11] i.e., agents receive more similar information, but the total informativeness of the signal is unaltered. Arguably, social media has changed not just the similarity but also the amount of information available. There is a large literature that studies the value of more information in games. Our formulation ensures that each agent has the same information about fundamentals ($\boldsymbol{\theta}$) and that changes in information similarity affect only strategic uncertainty—belief about others' information.

We construct posteriors explicitly using the marginal distributions, rather than model signals as posteriors, as is now standard, following Kamenica and Gentzkow (2011). Our assumption that $\mu(\cdot)$ is injective implies that signals and posteriors are interchangeable. We could have, alternatively, chosen signals as posteriors and started with a *feasible* joint distribution over posterior beliefs instead.[12] By working with signals directly, we have feasible joint distributions by construction.

3. Information Similarity and Equilibrium Turnout

In this section, we present our main results that characterize how increased similarity of information affects turnout in equilibrium.

3.1. Preliminaries

A regime-change game has flavors of both complementarity and substitutability. Accordingly, an agent's incentive to participate is nonmonotonic in others' participation. In equilibrium, an agent participates and bears the cost if and only if she believes that it is sufficiently likely that a change is beneficial and that her participation will make a difference.

Consider an agent in group $g \in \{1, 2\}$. She believes that the size of the other group is $\mathbf{N}_{-g} \sim \eta(.)$. However, that she belongs to group g may change her belief about the size of her own group. She believes that $\mathbf{N}_g - 1 \sim \eta^A(\cdot)$. Let $\eta_2^A(\cdot)$ denote her belief about $\mathbf{N}_g - 1 + \mathbf{N}_{-g}$.[13] The agent's participation incentive depends on whether she expects her own participation to make a difference (to the protest's outcome). Below, we define expressions Λ_2, Λ_1, and Λ, which denote, respectively, the probabilities of an agent making a difference when both groups participate, when only her own group participates, and when only the other group participates.

$$\Lambda_2 := \sum_{k=0}^{\infty} \phi(k)\eta_2^A(k), \;\; \Lambda_1 := \sum_{k=0}^{\infty} \phi(k)\eta^A(k), \;\; \Lambda := \sum_{k=0}^{\infty} \phi(k)\eta(k) \tag{2}$$

[11]We analyze the effect of increasing similarity of $\mathscr{P}_x^1(\cdot)$. This is because, given the definitions of states and payoffs in our game, changing $\mathscr{P}_x^0(\cdot)$ is not payoff relevant if marginals $\mathbb{P}^\theta$ are unaltered. See Section 6 for a more detailed discussion.

[12]With more than one agent, characterizing the feasible joint distributions over posteriors is not trivial. Recently, Arieli et al. (2021) characterize the set of feasible two-dimensional joint distributions.

[13]Since $\mathbf{N}_1$ and $\mathbf{N}_2$ are i.i.d., this distribution does not need to be indexed by group g.

In collective action games, agents take a costly action when they believe that they can make a difference (are pivotal). While models with pivotatility are ubiquitous in the voting literature, one common critique is that voters in large electorates are unlikely to ever be pivotal. But in our setting with uncertainty about the population size and the required threshold, Λ_2, Λ_1, and Λ are smooth functions of pivotal probabilities and reflect agents' beliefs about whether their participation can make a difference. Indeed, Cantoni et al. (2019) validate such strategic considerations on the part of the potential participants empirically. So, Λ_2, Λ_1, and Λ serve as a way of modeling this incentive to avoid costly participation in any collective action game with strategic uncertainty.

We assume that, from any agent's perspective, the probability of making a difference when only the other group participates is weakly smaller than the probability of making a difference when only her own group participates or when both groups participate.[14]

Assumption 1: $\max\{\Lambda_1, \Lambda_2\} \geqslant \Lambda$

We first write down the conditions for a strategy profile to be an equilibrium.

Proposition 1: *A strategy profile σ is an equilibrium if and only if for all $g \in \{1,2\}$ and for all $x \in \mathcal{X}$,*

$$\mathscr{P}^1_x(P_{-g})\Lambda_2 + (1 - \mathscr{P}^1_x(P_{-g}))\Lambda_1 \geqslant \frac{c}{\mu(x)} \text{ if } x \in P_g \tag{IC:P}$$

$$\mathscr{P}^1_x(P_{-g})\Lambda \leqslant \frac{c}{\mu(x)} \text{ if } x \in NP_g. \tag{IC:NP}$$

The intuition is straightforward. Consider an agent in group g with signal $x \in P_g$. If the other group also receives a signal in its participation set, which occurs with probability $\mathscr{P}^1_x(P_{-g})$, then this agent can make a difference with probability Λ_2. If the other group does not receive a signal in P_{-g}, then the agent can make a difference with probability Λ_1. (IC:P) simply says that the agent has an incentive to incur the cost of participating if she believes she can make a difference with a sufficiently high probability. The logic behind (IC:NP) is similar.

To capture how the incentives change with similarity of information, it is convenient to partition the model primitives into two environments.

Definition 3 (Encouragement/Discouragement)**:**

- *We say we are in the **encouragement environment** if*

$$\Lambda_2 > \Lambda_1. \tag{E}$$

[14] This assumption is satisfied if, for example, $\mathbf{N}_1, \mathbf{N}_2$ are Poisson random variables, as in, for example, (Myerson, 1998). In contrast, if $\mathbf{N}_1, \mathbf{N}_2$ are deterministic, say N, then Assumption 1 is violated if $Prob(\bar{\boldsymbol{n}} = N + 1) > \max\{Prob(\bar{\boldsymbol{n}} = N), Prob(\bar{\boldsymbol{n}} = 2N\}$. Assumption 1 rules out cases in which an agent from a group that does not participate has the strongest incentive to participate, knowing that no one in his group will participate.

In this case, an agent is more likely to make a difference when both groups participate than if only her own group participates; that is, higher turnout encourages participation.

- *We say we are in the* ***discouragement environment*** *if*

$$\Lambda_1 > \Lambda_2. \tag{D}$$

In this case, an agent is more likely to make a difference when only her group participates than if both groups participate; that is, higher turnout discourages participation.

At first glance, the encouragement and discouragement environments may seem to be environments of strategic complementarity and substitutability, respectively. This is not quite true. The discouragement environment does not feature strategic substitutability, because a nonparticipating agent in group g (with $x_g \in NP_g$) has a *stronger* incentive to participate if the other group is more likely to participate.

Theorems 1 and 2 establish that if we compare information similarity in the sense of CAD, then the above simple condition about primitives (whether $\Lambda_2 > \Lambda_1$ or $\Lambda_1 > \Lambda_2$) yields a complete characterization of when increased information similarity facilitates or hinders participation.

3.2. Encouragement Environment

Theorem 1: *In the encouragement environment, the maximal equilibrium turnout increases when information becomes more similar. That is,*

$$\mathscr{P}^\theta \geqslant_{CAD} \widehat{\mathscr{P}}^\theta \text{ for all } \theta \implies \mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}}) \text{ in (E)}.$$

This is true regardless of how $\mathscr{P}^0$ changes, as long as the marginals $\mathbb{P}^\theta$ are unaltered.

The proof is in the appendix and proceeds in two steps. First, we show that in the encouragement environment, the maximal participation equilibrium must be in symmetric strategies. Then we show that any maximal participation equilibrium remains an equilibrium when information similarity increases. Suppose that in the maximal equilibrium, in each group, an agent participates if and only if $x \in P$. If information similarity increases, an agent with $x \in P$ now assigns a higher probability that the other group also sees a signal $x \in P$ that induces them to participate. That is, $\mathscr{P}^1_x(P)$ increases for $x \in P$. Since $\Lambda_2 > \Lambda_1$, we can see from (IC:P) that such an agent has an even stronger incentive to participate. Analogously, if information similarity increases, a nonparticipating agent with $x \in NP$ now assigns a lower probability that the other group sees $x \in P$. That is, $\mathscr{P}^1_x(P)$ decreases for $x \in NP$. We can see from (IC:NP) that a nonparticipating agent has an even weaker incentive to participate.

3.3. Discouragement Environment

Next, we analyze the discouragement environment, in which $\Lambda_1 > \Lambda_2$. The maximal equilibrium might no longer be symmetric. Further, a symmetric equilibrium under $\widehat{\mathscr{P}}$ might no longer be an equilibrium under $\mathscr{P}$ when $\mathscr{P} \succcurlyeq_{CAD} \widehat{\mathscr{P}}$. This alone does not imply a smaller maximal equilibrium turnout under $\mathscr{P}$, because new equilibria may arise under $\mathscr{P}$ that were not sustainable under $\widehat{\mathscr{P}}$. Given any information structure $\mathscr{P}$ and a maximal equilibrium σ^*, we define a condition that describes why σ^* is maximal under $\mathscr{P}$.

Definition 4 (Condition M)**:** *Let σ^* be a maximal equilibrium for $\mathscr{P}$ with participation sets (P_1^*, P_2^*). We say that $\mathscr{P}$ satisfies condition M if, for any strategy profile $\hat{\sigma}$ with participation sets $(\hat{P}_1, \hat{P}_2)$ such that $\mathscr{V}(\hat{\sigma}) > \mathscr{V}(\sigma^*)$, at least one of the following holds.*

(M1) $\exists x \in \hat{P}_1 \cap \hat{P}_2$ *such that*

$$\min_{i \in \{1,2\}} \left\{ \mathscr{P}_x^1(\hat{P}_i)\Lambda_2 + (1 - \mathscr{P}_x^1(\hat{P}_i))\Lambda_1 \right\} < \frac{c}{\mu(x)}.$$

Or

(M2) $\exists x \in (\hat{P}_1 \cup \hat{P}_2) \backslash (\hat{P}_1 \cap \hat{P}_2)$ *such that*

$$\mathbb{1}_{x \in \hat{P}_1} \mathscr{P}_x^1\left(\hat{P}_1\right) + \mathbb{1}_{x \in \hat{P}_2} \mathscr{P}_x^1\left(\hat{P}_2\right) > \frac{c}{\Lambda \mu(x)}.$$

Condition M says there are two reasons why any strategy profile $\hat{\sigma}$ with a larger expected turnout than the maximal equilibrium σ^* fails to be an equilibrium. Either (IC:P) is violated for some signal that prescribes both groups to participate under $\hat{\sigma}$ or (IC:NP) is violated at a signal at which exactly one group is prescribed to participate under $\hat{\sigma}$. Below we establish that in the discouragement environment, if the information structure satisfies condition M, increasing information similarity can lead to lower maximal equilibrium turnout.

Theorem 2: *In a discouragement environment that satisfies condition M, the maximal equilibrium turnout decreases when information becomes more similar. That is,*

$$\mathscr{P}^\theta \succcurlyeq_{CAD} \widehat{\mathscr{P}}^\theta, \text{ for all } \theta \implies \mathscr{V}^*(\mathscr{P}) \leqslant \mathscr{V}^*(\widehat{\mathscr{P}}) \text{ under (D) if } \widehat{\mathscr{P}} \text{ satisfies condition } M.$$

Moreover, the inequality can be strict. The result is true regardless of how $\mathscr{P}^0$ changes, as long as the marginals $\mathbb{P}^\theta$ are unaltered.

The proof is in the appendix. The argument involves two steps. First, we argue that the maximal equilibrium might no longer be an equilibrium when information becomes more similar. Let σ^* with participation sets (P_1^*, P_2^*) be the maximal equilibrium under $\widehat{\mathscr{P}}$. Consider a participating agent with a signal $x \in P_1^* \cap P_2^*$. If information becomes more similar, this agent assigns a higher probability to the event that the other group also receives a signal in their respective participation set. However, unlike in the encouragement environment, this reduces her incentive to participate since $\Lambda_1 > \Lambda_2$. As a result, this agent's (IC:P) may be violated. Indeed, a nonparticipant's (IC:NP) may also fail. Consider a signal $x \in P_1^* \backslash P_2^*$. An agent in group 2 who receives such a signal is prescribed to not participate. However, with increased similarity of information, this agent assigns a higher probability that group 1 will participate. This, in turn, makes her more likely to participate, which may violate (IC:NP). So the maximal equilibrium σ^* under $\widehat{\mathscr{P}}$ may no longer be an equilibrium under $\mathscr{P}$.

In the second step, we establish that no new equilibrium with larger expected turnout arises under $\mathscr{P}$. Suppose, for a contradiction, there is an equilibrium σ' with $\mathscr{V}(\sigma') > \mathscr{V}(\sigma^*)$. By the maximality of σ^*, we know that σ' is not an equilibrium under $\widehat{\mathscr{P}}$. By condition M, two cases arise. In case (i), σ' is not an equilibrium under $\widehat{\mathscr{P}}$ because an agent's incentive to participate (IC:P) is violated at some signal $x \in P_1' \cap P_2'$, where both groups are prescribed to participate. With more similar information, (IC:P) would continue to be violated. To see why, note that (IC:P) is a convex combination of Λ_2 and Λ_1, and with more similar information, she assigns a higher weight to Λ_2. Now if (IC:P) was violated under $\widehat{\mathscr{P}}$, then it will also be violated under $\mathscr{P}$ because in the discouragement environment, $\Lambda_2 \leqslant \Lambda_1$. In case (ii), σ' is not an equilibrium under $\widehat{\mathscr{P}}$ because (IC:NP) is violated for some agent, say, from group 2 with a signal in $(P_1' \backslash P_2')$. Such an agent wishes to participate under $\widehat{\mathscr{P}}$ because she assigns a high probability to the event that group 1 participates. Under $\mathscr{P}$, when information is more similar, she has an even stronger incentive to participate (since she believes that the other group is more likely to participate and her own group is not). So, in both cases, if σ' is not an equilibrium under $\widehat{\mathscr{P}}$, then it cannot be an equilibrium under $\mathscr{P}$ either.

Two observations are worth highlighting. First, existing literature often restricts attention to monotone, cut-off strategies for the sake of tractability. In a game of pure complementarity (see Morris and Shin (2002)), the best and worst equilibrium are in cutoff strategies. However, with both, complementarities and substitutabilities, this is no longer true (see Shadmehr (2021)). We do not impose such a restriction on strategies—allowing non-monotonic and asymmetric strategies. We also do not impose restrictions on the signal structure, except finiteness. Consequently, there is little hope of characterizing all equilibria. However, we are able to characterize the effects of changing similarity of information through an indirect approach, in which we explore how the equilibrium set and maximal equilibria change. Second, this comparative static is achieved by choosing an order of similarity of information, namely the CAD order, that seems particularly well suited for the task. This order partitions the environment into two economically interpretable sub-environments—encouragement and discouragement—delivering a complete characterization of when information similarity helps or harms in terms of the primitives of the game.

4. Application 1: Which Regimes Suffer from Increased Information Similarity?

In this section, we apply our characterization to ask a new qualitative question: does increased similarity of information have a differential effect on different types of regimes? Proposition 2 shows that as information becomes more similar, resilient regimes face larger protests while weak regimes become less vulnerable and face smaller protests.

We say a regime with participation threshold $\bar{\boldsymbol{n}}'$ is more resilient than another with threshold $\bar{\boldsymbol{n}}$, denoted by $\bar{\boldsymbol{n}}' \geqslant_{st} \bar{\boldsymbol{n}}$, if $\bar{\boldsymbol{n}}'$ first-order stochastically dominates $\bar{\boldsymbol{n}}$, and we make the following assumption.

Assumption 2 (Single-crossing)**:** *If $(\Lambda_2 - \Lambda_1)(\bar{\boldsymbol{n}}) \geqslant 0$, then $(\Lambda_2 - \Lambda_1)(\bar{\boldsymbol{n}}') \geqslant 0$ for all $\bar{\boldsymbol{n}}' \geqslant_{st} \bar{\boldsymbol{n}}$.*

This assumption captures the idea that if an agent is more likely to make a difference when both groups join than when only her group joins, this continues to be true when it is even more difficult to bring about a change. That is, if $\Lambda_2 > \Lambda_1$ for a given regime, then $\Lambda_2 > \Lambda_1$ also for more resilient regimes. Assumption 2 is satisfied, for example, if the group sizes $\mathbf{N}_1, \mathbf{N}_2$ are drawn from a Poisson distribution and the resilience of the regime $\bar{\boldsymbol{n}}$ is deterministic. In this case, $\Lambda_2 > \Lambda_1$ if and only if $\bar{n} \geqslant n^* = \frac{N}{\ln 2}$.

Proposition 2: *Let Assumption 2 hold. Define $\bar{\boldsymbol{n}}$ such that $(\Lambda_2 - \Lambda_1)(\bar{\boldsymbol{n}}) = 0$.*

1. *For any $\bar{\boldsymbol{n}}'$ with $\bar{\boldsymbol{n}}' \geqslant_{st} \bar{\boldsymbol{n}}$,*

$$\mathscr{P}^1 \geqslant_{CAD} \widehat{\mathscr{P}}^1 \implies \mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}}).$$

2. *For any $\bar{\boldsymbol{n}}'$ with $\bar{\boldsymbol{n}} \geqslant_{st} \bar{\boldsymbol{n}}'$,*

$$\mathscr{P}^1 \geqslant_{CAD} \widehat{\mathscr{P}}^1 \implies \mathscr{V}^*(\mathscr{P}) \leqslant \mathscr{V}^*(\widehat{\mathscr{P}}) \quad \textit{if } \widehat{\mathscr{P}} \textit{ satisfies condition M.}$$

In words, for sufficiently resilient regimes, increased information similarity increases the maximal equilibrium turnout. However, for weak regimes, increased information similarity can have the opposite effect and reduce maximal turnout. The proof directly follows from Theorems 1 and 2.

This result suggests that the casual empirical observation that mass protests have increased in size and frequency may be biased. While popular discourse has highlighted how a more connected world has enabled larger protests against regimes previously thought to be invincible, it has not focused as much on how increased information similarity may have hindered the size and frequency of movements with ex ante easier goals.

Relatedly, increased similarity of information enables coordination of actions across groups, increasing the probability of both groups' participating or not participating. We show (in Proposition 5 in the appendix) that if a strategy profile constitutes a maximal equilibrium under two information structures ranked according to CAD, then *conditional on there being a protest*, the maximal equilibrium turnout (and the probability of a regime change when $\boldsymbol{\theta} = 1$) is strictly higher under the more similar information structure. This exacerbates the observation bias: focusing our attention on the coverage of high-visibility protests against resilient regimes may lead us to the inaccurate conclusion that the *unconditional* size of mass protests has increased in the age of more similar information.

5. Application 2: Open-Ballot Voting in Corporate Boards

Next, we apply our framework to voting in committees (à la Palfrey and Rosenthal (1985)). Think about a corporate board that must vote on whether to fire the current CEO. The management literature documents that board members often have no incentive to cast a vote against the CEO because voting on corporate boards is mostly by open ballot and the management can impose high costs on dissenting board members.[15]

We consider a board comprising $G > 2$ members who must vote for or against the current leadership. There are two states: the current leadership is either competent ($\boldsymbol{\theta} = 0$) or not ($\boldsymbol{\theta} = 1$). Let $\mu(1)$ and $\mu(0)$ denote $\mathscr{P}(\boldsymbol{\theta} = 1)$ and $\mathscr{P}(\boldsymbol{\theta} = 0)$, respectively. Each board member i receives a private, noisy binary signal $\mathbf{X}_i \in \mathcal{X} = \{0, 1\}$ about the state $\boldsymbol{\theta}$. Conditional on the state $\boldsymbol{\theta}$, the signals are drawn from a joint distribution $\mathscr{P}^{\theta}$. These signals can be interpreted as each board member's personal experience with the leadership. Increased similarity of information can arise from the members' sharing their experiences publicly before the vote. Suppose the voting rule is fixed: if more than $\bar{n} + 1$ members cast negative votes, then the current leadership is replaced. If the leadership changes, the members get a payoff of θ. Voting against the leadership entails a cost of c because of possible retaliation. For simplicity, we assume $\frac{c}{\mu(0)} > 1$. This guarantees that in any equilibrium, a member who experiences $\mathbf{X}_i = 0$ never casts a negative vote. We also restrict attention to symmetric strategies.

In this setting, we first characterize when increased information similarity

[15]The costs may be related to retaliation from senior management or to reputational damage from being perceived as contrarian, which may result in not being nominated again for a seat on the same or other boards. See Chemmanur and Fedaseyeu (2018) and references therein for both empirical evidence and theoretical models of board voting with costs of dissent. Chemmanur and Fedaseyeu (2018) refer to a coordination problem among directors: "To decide whether or not to vote against the CEO, a director who received a bad private signal about the CEO's quality needs to assess the expected number of other directors who also received bad private signals and therefore the probability of his being in the majority. This is because, if he votes against the CEO but fails to oust her, he will incur dissent costs."

strengthens or weakens the incentive to cast a negative vote conditional on a bad experience. Moreover, we investigate how the optimal voting rule changes with increased information similarity.

Note that, unlike in the baseline model, there is no population uncertainty and the participation threshold is not random. But most importantly, since we have information being received across $G > 2$ agents, we now need a notion of information similarity to compare random variables with $G > 2$ dimensions. Consider a strategy profile, denoted by σ^1, in which members vote against the CEO whenever they experience $\mathbf{X}_i = 1$. Let $I = \sum_{j \in G} \mathbb{1}_{\mathbf{X}_j = 1}$ denote the number of board members who cast a negative vote and $I_{-i} = \sum_{j \in G \setminus \{i\}} \mathbb{1}_{\mathbf{X}_j = 1}$ denote the number of board members other than i who cast a negative vote. For the strategy profile to be an equilibrium, we need

$$\gamma_1^1(\bar{n}) \geqslant \frac{c}{\mu(1)}, \tag{IC-voting}$$

where γ_1^1 is the probability distribution function of I_{-i} conditional on $\boldsymbol{\theta} = 1$ and $\mathbf{X}_i = 1$. We extend our notion of information similarity as follows.

Definition 5: *$\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$ if there exists $k^* \in \{0, 1, \ldots, G-2\}$ such that*

$$\gamma_1^1(k) \leqslant \hat{\gamma}_1^1(k) \text{ for all } k \leqslant k^*$$

$$\text{and } \gamma_1^1(k) \geqslant \hat{\gamma}_1^1(k) \text{ for all } k > k^*.$$

We call k^* the index of sign change between $\mathscr{P}^1$ and $\widehat{\mathscr{P}}^1$.

In words, if information similarity among members increases, then conditional on having a bad experience ($\mathbf{X}_i = 1$), a board member assigns a higher probability to more than k^* others having also observed a bad signal and a lower probability to fewer than k^* others having observed a good experience, conditional on the CEO being incompetent.[16]

Proposition 3 characterizes how increased similarity of information affects the maximal equilibrium (the equilibrium with the largest number of no votes, conditional on the CEO being incompetent). The only candidate equilibrium with any negative votes is σ^1. Therefore, studying maximal equilibria reduces to starting out with a σ^1 that is an equilibrium and checking whether it remains an equilibrium when information similarity increases.

Proposition 3: *Fix a voting rule $\bar{n}$. Suppose $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$, and let k^* be the associated index of sign change between $\mathscr{P}^1$ and $\widehat{\mathscr{P}}^1$. Suppose $\sigma^1 \in \mathcal{E}(\widehat{\mathscr{P}})$.*

[16]Notice that we do not restrict the conditional beliefs after $\mathbf{X}_i = 0$. This is because, given $c > \mu(0)$, an agent never casts a negative vote after observing $\mathbf{X}_i = 0$. For more general environments, we would need restrictions on the conditional beliefs after any signal realization. In a companion paper, we present extensions of the CAD order for more than two groups and study their implications for equilibrium behavior in a class of binary-action games.

1. *If $k^* < \bar{n}$, then $\sigma^1 \in \mathcal{E}(\mathscr{P})$.*
2. *If $k^* \geqslant \bar{n}$, then it is possible that $\sigma^1 \notin \mathcal{E}(\mathscr{P})$.*

Consider two extreme cases. First, suppose a board member had a bad experience and unanimous negative votes are required to replace the CEO (that is, $\bar{n}+1 = G$). Then the member's vote is relevant only when all the others also vote against the CEO, which requires that all others also have had bad experiences. With increased information similarity, $\gamma_1^1(G-1)$ increases, which makes it more likely that her vote is relevant. Thus, increased information similarity increases her incentive to cast a negative vote conditional on a bad experience, regardless of k^*. But if replacing the CEO requires just one negative vote ($\bar{n}+1 = 1$), then her vote is relevant only if all the others have observed good experiences. Increased information similarity means $\gamma_1^1(G-1)$ decreases, and therefore the incentive to vote against the CEO is diminished, regardless of k^*. In general, whether similar information helps or hurts depends on k^*. If the number of votes required to replace the CEO is sufficiently (insufficiently) demanding, then information similarity increases (reduces) the maximal equilibrium number of negative votes.[17]

The design of optimal voting rules in committees is a long-studied normative question, and our framework allows us to also analyze how the optimal rule changes with increased information similarity.[18] Recall that there are multiple equilibria in our setting. We assume board members will play the maximal equilibrium, and the optimal threshold is one that maximizes the probability that the CEO is replaced conditional on being incompetent; that is, the optimal threshold $\bar{n}^*(\mathscr{P})$ is given by

$$\bar{n}^*(\mathscr{P}) := \operatorname{argmax}_{\bar{n}} \mathscr{P}^1(I \geqslant \bar{n}+1 | \boldsymbol{\theta} = 1),$$

subject to the (IC-voting) constraint.[19]

In the absence of the incentive constraint, the designer would naturally set $\bar{n}$ as low as possible. But a low $\bar{n}$ reduces an individual member's incentive to cast a vote against the CEO. So the optimal threshold is the lowest one that satisfies the incentive constraint. Suppose the optimal voting rule is sufficiently demanding. It follows from Proposition 3 that increased information similarity will increase the incentive to cast a vote against the CEO conditional on having had a bad experience. This implies it can be optimal for the designer to make the threshold less demanding without violating the (IC-voting) constraint. An analogous argument holds when the voting rule is insufficiently demanding.

[17] See Section 6 and the appendix for more on settings with $G > 2$ groups.

[18] In a recent paper, Kattwinkel and Winter (2023) study the optimal decision mechanism for juries, allowing for general mechanisms but keeping the information structure fixed with independent signals across jurors. Our framework considers a special class of decision mechanisms, namely those in which a minimum threshold number of negative votes is required for a decision, and asks how the optimal rule changes with a changing information structure.

[19] The qualitative argument is unchanged if we assume a small negative payoff from replacing a competent CEO.

Proposition 4: *Suppose that $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$ and k^* is the index of sign change between $\mathscr{P}^1$ and $\widehat{\mathscr{P}}^1$.*

1. *If $k^* < \bar{n}^*(\widehat{\mathscr{P}})$, then $\bar{n}^*(\mathscr{P}) \leqslant \bar{n}^*(\widehat{\mathscr{P}})$.*
2. *If $k^* \geqslant \bar{n}^*(\widehat{\mathscr{P}})$, then $\bar{n}^*(\mathscr{P}) \geqslant \bar{n}^*(\widehat{\mathscr{P}})$.*

6. Discussion

Finally, we discuss some extensions. Formal results are in the online appendix.

State-dependent changes in similarity: There is much public discussion of how undesirable (autocratic) regimes restrict the flow of information via social media, making coordination among citizens harder. We could ask how protest size changes if the similarity of information is state dependent: information similarity increases when the regime change is not beneficial ($\boldsymbol{\theta} = 0$) but not when the regime change is beneficial ($\boldsymbol{\theta} = 1$).

In our model, changes in information similarity in state 0 are irrelevant, as they do not affect agents' incentives, i.e., we implicitly assumed that overthrowing a bad regime is welfare-improving but overthrowing a good regime is not welfare-reducing. We could have relabeled state 0 as a negative number—say, $\Theta = \{-1, 1\}$—so that overthrowing a good regime led to a welfare loss. We can use our framework to demonstrate qualitatitively different effects in this case. Proposition 6 in the online appendix shows that if information similarity increases in the good regime ($\boldsymbol{\theta} = -1$) and is unchanged in the bad regime ($\boldsymbol{\theta} = 1$), then the expected turnout falls in the encouragement environment and increases in the discouragement environment.

More than two groups: With two groups, comparisons in the CAD order allowed for a complete characterization of when more similar information increased or decreased maximal equilibrium participation. Our application of voting on a board illustrates that our results extend qualitatively to some settings with more than two agents. However, in general, with more than two groups, no order of similarity can yield a complete characterization. This is because the probability of making a difference in a regime-change game is often quasi-concave—first increasing in the number of groups with similar information, and then decreasing—and random variables are ranked according to the quasi-concave order if and only if they have the same distribution.[20] This means that, in general, the participation incentives will not be ranked as we vary information similarity. In the online appendix, we propose a notion of similarity of n-dimensional random variables and

[20] To be ranked according to the quasi-concave order means that the expectations of any quasi-concave function are ranked. To see why this ranking implies the same distribution, notice that $f(x) := \mathbb{1}_{x \geqslant z}$ and $f(x) := \mathbb{1}_{x \leqslant z}$ are both quasi-concave.

show (in Proposition 7) that results qualitatively similar to Theorems 1 and 2 are valid, even though the characterization is partial. We use an n-dimensional order due to Meyer (1990), which is stronger than the one in Section 5.

Mixed strategies: We restricted attention to pure-strategy equilibria. In the encouragement environment this is without loss, as maximal equilibria are pure and symmetric. This is not true in the discouragement environment. In the online appendix, we demonstrate in a simple example with binary signals that the qualitative insight remains valid—increasing similarity may cause harm in the discouragement environment—when we allow for mixed strategies. The general analysis is challenging with a richer signal space because the maximal equilibrium could involve mixing on some signals with different probabilities. Therefore, it is no longer possible to partition $\mathcal{X}$ cleanly into participation and nonparticipation sets P and NP. Recall that much of the analysis examined how $\mathscr{P}^1_x(P)$ changes for any x. But when agents mix with different probabilities after different signals, we need to understand how $\mathscr{P}^1_x(x')$ changes for each $x, x' \in \mathcal{X}$. Developing tools to completely characterize how the set of mixed-strategy equilibria varies with information similarity is left for future work.

Welfare In this paper we focus on the size of participation. A natural question is also how increased information similarity affects welfare. A priori increased information similarity does not necessarily increase welfare. To see this note that more similar information, in a symmetric equilibrium, enables coordination of actions, increasing the probability of both groups participating and neither group participating. Therefore, even if maximal equilibrium turnout increases, it is unclear whether the protest is more likely to be successful. Moreover, if more agents protest than required, they bear additional cost.

Specializing to the environment with Poisson population uncertainty[21] as described in Section 4, we study the relationship between information similarity and welfare. Exactly as in Theorem 1, welfare unambiguously (weakly) increases in the encouragement environment. On the other hand, it is straightforward to construct an example where welfare reduces with increased similarity of information in the discouragement environment. Notice that such a possibility result—that welfare may go down in the discouragement environment—is weaker than Theorem 2. The reason is that even if increased similarity of information were to decrease the expected turnout, protests may exhibit more coordination between two groups, thereby still increasing welfare.

Cost incurred only if the oppressive regime survives: In many regime-change settings, protesters are deterred not by the costs of participating but by the costs of retaliation from the regime if the protest fails. To capture such situations, consider the following game.

[21] Both $\mathbf{N}_1, \mathbf{N}_2$ follow a Poisson distribution with mean N and $\bar{\boldsymbol{n}}$ is a constant.

	$A \geqslant \bar{n}+1$	$A \leqslant \bar{n}$
$a=1$	θ	$-\theta c$
$a=0$	θ	0

Essentially identical arguments as in Theorem 1 lead to the conclusion that increased similarity of information leads to larger protests in the encouragement environment. The effect of increased similarity of information in the discouragement environment is more nuanced. On the one hand, increasing information similarity makes an agent less likely to make a difference (assuming symmetric strategies). This dampens her incentive to participate. On the other hand, the expected cost of participating is lower because a protest is more likely to be successful with more people and because the costs are incurred only when a protest fails. Therefore, it is possible that even in the discouragement environment, increasing information similarity can lead to larger protests.

Informativeness of turnout: In many settings, individuals use protests and petitions to convey dispersed private information to policymakers in order to influence policy, and in turn, policymakers use the observed turnout of protests or the number of signatures on a petition to *infer* the state of the world and then decide whether to change a particular policy. Our baseline model abstracts from this issue as regime change occurs whenever turnout exceeds an exogenous threshold. We consider a version of our model in which a strategic policymaker observes the realized turnout and updates her belief about the state of the world, and then decides whether to change the regime. Recent work by Battaglini (2017) and Ekmekci and Lauermann (2019) use a similar setup and focus on the informational role of turnout. We can apply our framework to ask how similarity of information affects the informational content of turnout: Does increasing information similarity improve information aggregation?

We first show that informativeness of equilibria can decrease with increasing similarity of information.[22] The intuition is that when information becomes more similar, holding the strategies fixed, the policymaker wants to lower the threshold, and this has two opposing effects. On the one hand, a lower threshold can encourage more participation by making it more likely that individual agents can make a difference. On the other hand, a lower threshold can exacerbate free-riding, eliminating some equilibria. We also ask when increasing information similarity can help information aggregation, and show that when the threshold belief at which the policymaker changes the regime is intermediate (within a certain range), increasing information similarity can enable information aggregation that was otherwise impossible when the signals are conditionally independent.

[22]This is consistent with current public discourse. For instance, in a piece about technology and protests in the *Atlantic*, Zeynep Tufekci writes, *"Protests are signals: 'We are unhappy, and we won't put up with things the way they are.' But for that to work, the 'We won't put up with it' part has to be credible. Nowadays, large protests sometimes lack such credibility, especially because digital technologies have made them so much easier to organize."*(See https://www.theatlantic.com/technology/archive/2020/06/why-protests-work/613420/)

Information design: A natural question is how a designer might choose the optimal level of similarity of information, given a certain objective. In the online appendix, we derive the information structure that maximizes expected protest size when change is beneficial (in $\boldsymbol{\theta} = 1$). We show that in the encouragement environment, the optimal information structure is full correlation: both groups receive identical signals. In the discouragement environment, interior levels of similarity—neither conditionally independent signals nor full correlation—can be optimal if the conditionally independent signals do not satisfy condition M. Our analysis restricts attention to information structures that are (weakly) more similar than conditionally independent signals. More generally, in the discouragement environment, some negative interdependence may be desirable.

References

G.-M. Angeletos and A. Pavan. Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142, 2007. 6

I. Arieli, Y. Babichenko, F. Sandomirskiy, and O. Tamuz. Feasible joint posterior beliefs. *Journal of Political Economy*, 129(9):2546–2594, 2021. 11

Y. Awaya and V. Krishna. Commonality of information and commonality of beliefs. *Working Paper*, 2022. 6

M. Battaglini. Public protests and policy making. *The Quarterly Journal of Economics*, 132(1):485–549, 2017. 6, 7, 22

D. Bergemann and S. Morris. Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308, 2013. 6

D. Cantoni, D. Y. Yang, N. Yuchtman, and Y. J. Zhang. Protests as strategic games: experimental evidence from hong kong's antiauthoritarian movement. *The Quarterly Journal of Economics*, 134(2):1021–1077, 2019. 2, 7, 12

T. J. Chemmanur and V. Fedaseyeu. A theory of corporate boards and forced ceo turnover. *Management Science*, 64(10):4798–4817, 2018. 7, 17

M. W. Cripps, J. C. Ely, G. J. Mailath, and L. Samuelson. Common learning. *Econometrica*, 76(4):909–933, 2008. 6

W. Dziuda, A. A. Gitmez, and M. Shadmehr. The difficulty of easy projects. *American Economic Review: Insights*, 3(3):285–302, 2021. 7

C. Edmond. Information manipulation, coordination, and regime change. *Review of Economic studies*, 80(4):1422–1458, 2013. 6

M. Ekmekci and S. Lauermann. Informal elections with dispersed information. 2019. 6, 7, 22, 38

J. Hirshleifer. The private and social value of information and the reward to inventive activity. *The American Economic Review*, 61(4):561–574, 1971. 6

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(October):2590–2615, 2011. 11

D. Kattwinkel and A. Winter. Optimal decision mechanisms for juries: Acquitting the guilty. *Working Paper*, 2023. 7, 19

M. Manacorda and A. Tesei. Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567, 2020. 6

M. Meyer and B. Strulovici. Increasing interdependence of multivariate distributions. *Journal of Economic Theory*, 147(4):1460–1489, 2012. 6

M. A. Meyer. *Interdependence in multivariate distributions: stochastic dominance theorems and an application to the measurement of ex post inequality under uncertainty*. Nuffield College, 1990. 4, 6, 10, 21, 26, 35

S. Morris and H. S. Shin. Social value of public information. *american economic review*, 92(5):1521–1534, 2002. 6, 9, 15

A. Müller and D. Stoyan. *Comparison methods for stochastic models and risks*, volume 389. Wiley, 2002. 6

R. B. Myerson. Population uncertainty and poisson games. *International Journal of Game Theory*, 27(3):375–392, 1998. 7, 12

M. Olson. The logic of collective action [1965]. *Contemporary Sociological Theory*, 124, 1965. 2, 6

T. R. Palfrey and H. Rosenthal. Voter participation and strategic uncertainty. *American political science review*, 79(1):62–78, 1985. 17

I.-U. Park and E. Smyrniotis. Global games without dominance solvable games. *Working Paper*, 2022. 7

M. Shadmehr. Protest puzzles: Tullock's paradox, hong kong experiment, and the strength of weak states. *Quarterly Journal of Political Science*, 16(3):245–264., 2021. 7, 15

G. Tullock. The paradox of revolution. *Public Choice*, pages 89–99, 1971. 2, 6

A. Appendix: Proofs

A.1. Proof of Lemma 1

Proof. Let $\vec{x} = (x_1, x_2, \ldots x_G)$ be a profile of signal realizations. By definition,

$$\begin{aligned}
\mathscr{V}(\sigma) &= \sum_{\vec{x}\in\mathcal{X}^G} \mathscr{P}^1(\vec{x}) \left[\sum_{g=1}^{G} N\mathbb{1}_{x_g\in P_g}\right] \\
&= \sum_{\vec{x}\in\mathcal{X}^G} \sum_{g=1}^{G} \mathscr{P}^1(\vec{x}) N\mathbb{1}_{x_g\in P_g} \\
&= \sum_{g=1}^{G} \sum_{x_g\in\mathcal{X}} \mathbb{P}^1_g(x_g) N\mathbb{1}_{x_g\in P_g} \\
&= N\sum_{g=1}^{G} \mathbb{P}^1(P_g).
\end{aligned}$$

where the last equality is due to the exchangeability of the distribution. $\square$

A.2. Proof of Lemma 2

Proof. Suppose $\mathcal{Y} \subset \mathbb{R}$ is finite, and Y and $\widehat{Y}$ are two $\mathcal{Y}^2$-valued random variables with joint distributions $\mathscr{D}$ and $\widehat{\mathscr{D}}$ respectively, and identical marginals. Consider any two distinct points in the support of Y, say y_j, y_k. Define an "elementary transformation along identical intervals" (ETI) as an operation in which, for some $\alpha > 0$, we increase the probability mass on points (y_j, y_j) and (y_k, y_k) each by α, and reduce the probability mass on (y_j, y_k) and (y_k, y_j) each by α. An alternative characterization of our CAD order in two dimensions is that $\mathscr{D} \succcurlyeq_{CAD} \widehat{\mathscr{D}}$ if and only if $\mathscr{D}$ can be derived from $\widehat{\mathscr{D}}$ by a finite sequence of ETIs. This follows from Meyer (1990) (Proposition 1). We use this characterization to establish Lemma 2.

Let $(y_{i,k}, y_{j,k})_k$, $k = 1, \ldots n$, be a finite set of points in $\mathcal{Y}^2$ describing a sequence of ETIs, each with a mass α_k, to obtain $\mathscr{D}$ from $\widehat{\mathscr{D}}$. Let the resulting distribution after the k-th ETI be $\widehat{\mathscr{D}}_k$. So, $\widehat{\mathscr{D}}_1 = \widehat{\mathscr{D}}$ and $\widehat{\mathscr{D}}_n = \mathscr{D}$. If $(y_{i,k}, y_{j,k}) \in T \times T$ or $(\mathcal{Y}\backslash T) \times (\mathcal{Y}\backslash T)$, then $\widehat{\mathscr{D}}_k(T,T) = \widehat{\mathscr{D}}_{k-1}(T,T)$. On the other hand, if exactly one of $\{y_{i,k}, y_{j,k}\}$ is in T for some k, then $\widehat{\mathscr{D}}_k(T,T) = \widehat{\mathscr{D}}_{k-1}(T,T) + \alpha_k$. Therefore, $\mathscr{D}(T,T) = \widehat{\mathscr{D}}(T,T) + \sum_{k=1}^{n} \alpha_k$. Since an ETI leaves the marginal distribution unchanged, therefore $\mathscr{D}(T, \mathcal{Y}\backslash T) = \widehat{\mathscr{D}}(T, \mathcal{Y}\backslash T) - \sum_{k=1}^{n} \alpha_k$. The lemma follows.

$\square$

A.3. Proof of Proposition 1

Proof. Consider the payoff difference for any agent between participating and not. Let $u_g(a, x; \sigma)$ be the expected payoff of an agent from group g by playing action

a given that $\mathbf{X}_g = x$, and the players are following σ. We define the net expected payoff from participation as

$$\Delta_g(x;\sigma) := u_g(1,x;\sigma) - u_g(0,x;\sigma)$$

Consider an agent in group g. Suppose that $\mathbf{X}_g = x$. If she participates, she incurs a cost c and gets a positive payoff only if the regime is bad ($\boldsymbol{\theta} = 1$) and at least $\bar{\boldsymbol{n}}$ other agents participate. If she does not participate then she gets a positive payoff only if $\boldsymbol{\theta} = 1$ and the turnout is at least $\bar{\boldsymbol{n}} + 1$ without her participation. So, assuming that players play according to σ, we have

$$\Delta_g(x;\sigma)$$

$$= \mathscr{P}\left(\{\boldsymbol{\theta}=1\}\bigcap\{\mathbf{A}_{-g} \geqslant \bar{\boldsymbol{n}}\}\middle|\mathbf{X}_g = x\right) - c - \mathscr{P}\left(\{\boldsymbol{\theta}=1\}\bigcap\{\mathbf{A}_{-g} \geqslant \bar{\boldsymbol{n}}+1\}\middle|\mathbf{X}_g = x\right)$$

$$= \mathscr{P}(\boldsymbol{\theta}=1|\mathbf{X}_g = x)\Big[\mathscr{P}(\mathbf{A}_{-g} \geqslant \bar{\boldsymbol{n}}|\mathbf{X}_g = x, \boldsymbol{\theta}=1) - \mathscr{P}(\mathbf{A}_{-g} \geqslant \bar{\boldsymbol{n}}+1|\mathbf{X}_g = x, \boldsymbol{\theta}=1)\Big] - c$$

$$= \mu(x)\mathscr{P}(\mathbf{A}_{-g} = \bar{\boldsymbol{n}}|\boldsymbol{\theta}=1, \mathbf{X}_g = x) - c$$

To simplify the above expression further, consider two cases. (i) Suppose $x \in P_g$. If the realized signal for the other group $x_{-g} \in P_{-g}$, then $\mathbf{A}_{-g} = \mathbf{N}_g - 1 + \mathbf{N}_{-g}$, and if $x_{-g} \notin P_{-g}$, then $\mathbf{A}_{-g} = \mathbf{N}_g - 1$. (ii) Suppose $x \notin P_g$. Then, $x_{-g} \in P_{-g} \implies \mathbf{A}_{-g} = \mathbf{N}_{-g}$, and $x_{-g} \notin P_{-g} \implies \mathbf{A}_{-g} = 0$. Therefore,

$$\Delta_g(x;\sigma) = \begin{cases} \mu(x)\Big[\mathscr{P}^1_x(P_{-g})Prob\,(\mathbf{N}_g - 1 + \mathbf{N}_{-g} = \bar{\boldsymbol{n}} \mid \boldsymbol{\theta}=1, \mathbf{X}_g = x) & \\ \quad +\mathscr{P}^1_x(NP_{-g})Prob(\mathbf{N}_g - 1 = \bar{\boldsymbol{n}} \mid \boldsymbol{\theta}=1, \mathbf{X}_g = x)\Big] & \text{if } x \in P_g \\ \mu(x)\mathscr{P}^1_x(P_{-g})Prob(\mathbf{N}_{-g} = \bar{\boldsymbol{n}} \mid \boldsymbol{\theta}=1, \mathbf{X}_g = x) & \text{if } x \notin P_g \end{cases}$$

Finally, σ is an equilibrium if, for all $x \in \mathcal{X}$ and for all $g \in \{1,2\}$,

1. $\sigma_g(x) = 1 \implies \Delta_g(x;\sigma) \geqslant 0$.
2. $\sigma_g(x) = 0 \implies \Delta_g(x;\sigma) \leqslant 0$.

The expression in the proposition follows from noting that according to an agent in group g, $\mathbf{N}_g - 1 + \mathbf{N}_{-g} \sim \eta_2^A(\cdot)$, $\mathbf{N}_g - 1 \sim \eta^A(\cdot)$, $\mathbf{N}_{-g} \sim \eta(\cdot)$, and $\bar{\boldsymbol{n}} \sim \phi(\cdot)$.

□

A.4. Proof of Theorem 1

Proof. We prove this result using two steps. In Step 1, we establish that the maximal equilibrium is symmetric. In step 2, we show that the symmetric equilibrium under $\widehat{\mathscr{P}}^1$ remains an equilibrium under $\mathscr{P}^1$.

Step 1: With some abuse of notation, we say that $(P_1, P_2) \in \mathcal{E}(\mathscr{P})$ to mean that $\sigma := (\mathbb{1}_{\mathbf{X}_1 \in P_1}, \mathbb{1}_{\mathbf{X}_2 \in P_2}) \in \mathcal{E}(\mathscr{P})$.

LEMMA 3: *There is a unique maximal equilibrium in the encouragement environment, and it is symmetric.*

Proof. Suppose that σ is some asymmetric equilibrium with participation sets P_1 and P_2 for groups 1 and 2 respectively. We show that $\exists$ a symmetric equilibrium $\widehat{\sigma}$ with a participation set $P \supseteq P_1 \cup P_2$.

For any set S, define

$$\mathcal{T}(S) := S \bigcup \left\{ x \in \mathcal{X} : \mathscr{P}_x^1(S)\Lambda \geqslant \frac{c}{\mu(x)} \right\}. \tag{3}$$

In words, $\mathcal{T}(S)$ adds those signals to S (if there are any) at which an agent wants to participate if he believes that his group will not participate but the other group will participate if they receive a signal in S.

CLAIM 1: *Let $\sigma = (\mathbb{1}_{P_1}, \mathbb{1}_{P_2})$ be a strategy profile such that (IC:P) is satisfied for all $x \in P_1 \cup P_2$ given that players follow σ. Then, for all $x \in \mathcal{T}(P_1 \cup P_2)$,*

$$\mathscr{P}_x^1(\mathcal{T}(P_1 \cup P_2))\Lambda_2 + (1 - \mathscr{P}_x^1(\mathcal{T}(P_1 \cup P_2)))\Lambda_1 \geqslant \frac{c}{\mu(x)}$$

Proof. Since $(P_1, P_2) \in \mathcal{E}(\mathscr{P})$, (IC:P) implies,

$$\begin{aligned} \mathscr{P}_x^1(P_2)\Lambda_2 + (1 - \mathscr{P}_x^1(P_2))\Lambda_1 &\geqslant \frac{c}{\mu(x)} && \text{if } x \in P_1 \\ \mathscr{P}_x^1(P_1)\Lambda_2 + (1 - \mathscr{P}_x^1(P_2))\Lambda_1 &\geqslant \frac{c}{\mu(x)} && \text{if } x \in P_2 \end{aligned}$$

Since $\Lambda_2 > \Lambda_1$ and $\mathscr{P}_x^1(\cdot)$ is monotonic (in the set inclusion order),

$$\mathscr{P}_x^1(P_1 \cup P_2)\Lambda_2 + (1 - \mathscr{P}_x^1(P_1 \cup P_2))\Lambda_1 \geqslant \frac{c}{\mu(x)} \quad \forall x \in P_1 \cup P_2$$

If $\mathcal{T}(P_1 \cup P_2) \neq P_1 \cup P_2$, then, for all $x \in \mathcal{T}(P_1 \cup P_2) \backslash (P_1 \cup P_2)$, we have,

$$\begin{aligned} \mathscr{P}_x^1(P_1 \cup P_2)\Lambda &\geqslant \frac{c}{\mu(x)} \\ \implies \mathscr{P}_x^1(P_1 \cup P_2)\Lambda_2 + (1 - \mathscr{P}_x^1(P_1 \cup P_2))\Lambda_1 &\geqslant \frac{c}{\mu(x)} \end{aligned}$$

where the inequality is due to $\Lambda_2 > \Lambda_1 > 0$ and $\max\{\Lambda_2, \Lambda_1\} \geqslant \Lambda$ (see assumption 1). □

Define, $\mathcal{T}^*(P_1 \cup P_2) := \mathcal{T}^{|\mathcal{X}|}(P_1, P_2)$. First, notice that $\mathcal{T}(\cdot)$ is an increasing (in the set-inclusion order) map. Therefore, either $\mathcal{T}^*(S) = \mathcal{X}$ (due to the finiteness of $\mathcal{X}$), or $S \subseteq \mathcal{T}^*(S) \subset \mathcal{X}$ for any S. If $\mathcal{T}^*(P_1 \cup P_2) \neq \mathcal{X}$, then, by definition, (IC:NP) is satisfied for all $x \notin \mathcal{T}^*(P_1 \cup P_2)$ when players play $(\mathbb{1}_{\mathbf{X}_1 \in \mathcal{T}^*(P_1 \cup P_2)}, \mathbb{1}_{\mathbf{X}_2 \in \mathcal{T}^*(P_1 \cup P_2)})$. Moreover, (IC:P) is satisfied when both the groups play $a = 1$ on $\mathcal{T}^*(P_1 \cup P_2)$ by Claim 1. Therefore, given any equilibrium (P_1, P_2), $\mathcal{T}^*(P_1 \cup P_2)$ is a larger symmetric equilibrium.

Finally, let (P, P) and (P', P') be two maximal symmetric equilibria with $P \neq P'$. First, Claim 1 establishes that (IC:P) is satisfied for all signals in $\mathcal{T}(P \cup P')$, and hence for all the signals in $\mathcal{T}^*(P \cup P')$. By construction, (IC:NP) is satisfied for all the signals outside of $\mathcal{T}^*(P \cup P')$. Therefore, $\mathcal{T}^*(P \cup P')$ is an equilibrium, and $P, P' \subseteq \mathcal{T}^*(P \cup P')$, contradicting the maximality of P, P'. Therefore, $P = P'$, i.e., there is a unique maximal equilibrium. □

Step 2: By definition, $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$ implies the following:

1. $\mathscr{P}^1_x(P(\sigma)) \geqslant \widehat{\mathscr{P}}^1_x(P(\sigma))$ for all $x \in P(\sigma)$.
2. $\mathscr{P}^1_x(P(\sigma)) \leqslant \widehat{\mathscr{P}}^1_x(P(\sigma))$ for all $x \in NP(\sigma)$.

Consider (IC:P), the incentive constraint for an agent who participates. The left hand side is a convex combination of Λ_2 and Λ_1. In the encouragement environment, a higher weight on Λ_2 increases the LHS of (IC:P), making the constraint easier to satisfy. Therefore,

$$\begin{aligned}
\mathscr{P}^1_x(P(\sigma))\Lambda_2 + (1 - \mathscr{P}^1_x(P(\sigma)))\Lambda_1 &\geqslant \widehat{\mathscr{P}}^1_x(P(\sigma))\Lambda_2 + (1 - \widehat{\mathscr{P}}^1_x(P(\sigma)))\Lambda_1 \\
&\geqslant \frac{c}{\mu(x)} \quad \text{since } \sigma \in \mathcal{E}(\widehat{\mathscr{P}}).
\end{aligned}$$

Therefore, (IC:P) is satisfied for all $x \in P(\sigma)$ under $\mathscr{P}$. Similarly, for all $x \in NP(\sigma)$, $\mathscr{P}^1_x(P(\sigma)) \leqslant \widehat{\mathscr{P}}^1_x(P(\sigma))$. Therefore, (IC:NP) is satisfied for all $x \in NP(\sigma)$ under signal $\mathscr{P}$. Therefore, $\sigma \in \mathcal{E}(\mathscr{P})$.

Finally, since $\mathscr{V}(\cdot)$ depends only the marginal distributions (by Lemma 1)—which are the same in $\mathscr{P}$ and $\widehat{\mathscr{P}}$—it follows that $\mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}})$. □

A.5. Proof of Theorem 2

Proof of Theorem 2. We prove this result using two steps. In step 1, we show that the maximal equilibrium under $\widehat{\mathscr{P}}^1$ may no longer be an equilibrium under $\mathscr{P}^1$. In step 2, we show that no larger equilibrium can emerge when information becomes more similar.

Step 1: Let σ^* be a maximal equilibrium with the associated participation sets (P_1^*, P_2^*).

Case 1. $P_1^* = P_2^*$

Let $P := P_1^* = P_2^*$. By definition, $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$ implies the following:

(a) $\mathscr{P}^1_x(P) \geqslant \widehat{\mathscr{P}}^1_x(P)$ for all $x \in P$.

(b) $\mathscr{P}^1_x(P) \leqslant \widehat{\mathscr{P}}^1_x(P)$ for all $x \in NP$.

Consider (IC:P), the incentive constraint for an agent who participates. The left hand side is a convex combination of Λ_2 and Λ_1. In the discouragement environment, a higher weight on Λ_2 decreases the LHS of (IC:P). Therefore, for all $x \in P$,

$$\mathscr{P}^1_x(P)\Lambda_2 + (1 - \mathscr{P}^1_x(P))\Lambda_1 \leqslant \widehat{\mathscr{P}}^1_x(P)\Lambda_2 + (1 - \widehat{\mathscr{P}}^1_x(P))\Lambda_1$$

Therefore, (IC:P) may fail for some $x \in P$ under $\mathscr{P}$, in which case σ may no longer be in $\mathcal{E}(\mathscr{P})$.

Case 2. $P^*_1 \neq P^*_2$

Then, $\mathscr{P}^1_x(P^*_i) \geqslant \widehat{\mathscr{P}}^1_x(P^*_i)$ for all $i \in \{1, 2\}$ and $x \in P^*_i$. Notice that at least one of $P^*_1 \backslash P^*_2$ and $P^*_2 \backslash P^*_1$ is not $\varnothing$. Let $P^*_1 \backslash P^*_2 \neq \varnothing$ wlog. Consider some $x \in P^*_1 \backslash P^*_2$. Agents in group 2 must find it incentive compatible to not participate when they receive a signal in P^*_1. Therefore, we must have,

$$\mathscr{P}^1_x(P^*_1)\Lambda \leqslant \frac{c}{\mu(x)}$$

for all $x \in P^*_1$. However, since $\mathscr{P}^1_x(P^*_1) \geqslant \widehat{\mathscr{P}}^1_x(P^*_1)$ for all $x \in P^*_1$, (IC:NP) is harder to satisfy for any $x \in P^*_1 \backslash P^*_2$, and for any $x \in P^*_2 \backslash P^*_1$. Hence, σ^* may no longer be in $\mathcal{E}(\mathscr{P})$.

Step 2: Next, consider any $(\widehat{P}_1, \widehat{P}_2) \geqslant (P^*_1, P^*_2)$. For $(\widehat{P}_1, \widehat{P}_2)$ to be an equilibrium, one necessary condition is (IC:NP) for signals in P_1 and P_2 for groups 2 and 1 respectively. That is, we need that, at least one of the following holds:

$$\mathscr{P}^1_x(\widehat{P}_1)\Lambda \leqslant \frac{c}{\mu(x)} \text{ if } x \in \widehat{P}_1$$
$$\mathscr{P}^1_x(\widehat{P}_2)\Lambda \leqslant \frac{c}{\mu(x)} \text{ if } x \in \widehat{P}_2$$

Suppose that $\widehat{P}_1 \cap \widehat{P}_2 = \varnothing$. Since $\widehat{\mathscr{P}}$ satisfies condition M (Definition 4), part (ii) implies that, for some $x \in \widehat{P}_1 \cup \widehat{P}_2$,

$$\widehat{\mathscr{P}}^1_x(\widehat{P}_1)\Lambda > \frac{c}{\mu(x)} \text{ if } x \in \widehat{P}_1$$
$$\widehat{\mathscr{P}}^1_x(\widehat{P}_2)\Lambda > \frac{c}{\mu(x)} \text{ if } x \in \widehat{P}_2$$

By CAD, $\mathscr{P}^1_x(\widehat{P}_1) \geqslant \widehat{\mathscr{P}}^1_x(\widehat{P}_1)$ if $x \in \widehat{P}_1$ (and analogously for $\widehat{P}_2$). Therefore, $(\widehat{P}_1, \widehat{P}_2)$ cannot be an equilibrium in $\mathscr{P}$ if $\widehat{P}_1 \cap \widehat{P}_2 = \varnothing$.

Suppose that $\widehat{P}_1 \cap \widehat{P}_2 \neq \varnothing$. Since $\widehat{\mathscr{P}}$ satisfies Condition M, if (i) holds for some

$x \in \widehat{P}_1 \cap \widehat{P}_2$, then,

$$\min_{i\in\{1,2\}} \left\{ \widehat{\mathscr{P}}^1_x(\widehat{P}_i)\Lambda_2 + (1 - \widehat{\mathscr{P}}^1_x(\widehat{P}_i))\Lambda_1 \right\} < \frac{c}{\mu(x)}$$

By CAD, $\mathscr{P}^1_x(\widehat{P}_i) \geqslant \widehat{\mathscr{P}}^1_x(\widehat{P}_i)$ for $i \in \{1,2\}$. Since $\Lambda_1 > \Lambda_2$, this implies that,

$$\min_{i\in\{1,2\}} \left\{ \mathscr{P}^1_x(\widehat{P}_i)\Lambda_2 + (1 - \mathscr{P}^1_x(\widehat{P}_i))\Lambda_1 \right\} < \frac{c}{\mu(x)}$$

Therefore, (IC:P) fails for such an x.

Finally, if (IC:P) is satisfied for all $x \in \widehat{P}_1 \cap \widehat{P}_2$, then, by Condition M (ii), the exact same argument as in the case when $\widehat{P}_1 \cap \widehat{P}_2 = \varnothing$ implies that $(\widehat{P}_1, \widehat{P}_2)$ cannot be an equilibrium under $\mathscr{P}$. Therefore, no larger equilibrium can exist under $\widehat{\mathscr{P}}$, i.e., $\mathscr{V}^*(\widehat{\mathscr{P}}) \leqslant \mathscr{V}(\mathscr{P})$. □

A.6. Turnout conditional on protests

In the main text, we focus on how information similarity affects the equilibrium turnout. However, even when we fix an equilibrium, information similarity also has an additional impact on turnout. It makes the participation more coordinated. Accordingly, conditional on there being a protest, we may see that the protests are more likely to bring about social changes. The following proposition formalizes this intuition.

Proposition 5: *Suppose $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}^1}$, and $\sigma \in \mathcal{E}(\mathscr{P}) \cap \mathcal{E}(\widehat{\mathscr{P}})$ and σ is symmetric. Then,*

$$\mathscr{P}^1 \left[\mathbf{A}(\sigma) > \bar{n} \middle| \mathbf{A}(\sigma) > 0 \right] \geqslant \widehat{\mathscr{P}^1} \left[\mathbf{A}(\sigma) > \bar{n} \middle| \mathbf{A}(\sigma) > 0 \right].$$

Proof. Let $H_2(\cdot), H_1(\cdot)$ be the CDFs of $\mathbf{N}_1 + \mathbf{N}_2$ and $\mathbf{N}_1$ (and $\mathbf{N}_2$ by symmetry) respectively.

Suppose that $\sigma \in \mathcal{E}(\mathscr{P}) \cap \mathcal{E}(\widehat{\mathscr{P}})$ and $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}^1}$. Under $\widehat{\mathscr{P}}$, the probability of there being a protest at all when $\boldsymbol{\theta} = 1$ is $\mathscr{P}^1(\{\mathbf{A} > 0\}) = 1 - \widehat{\mathscr{P}^1}(NP, NP)$. Here $\mathscr{P}^1(NP, NP)$ means $\mathscr{P}^1(\{\mathbf{X}_1 \in NP(\sigma), \mathbf{X}_2 \in NP(\sigma)\})$. Therefore,

$$\mathscr{P}^1 \left(\{\mathbf{A}(\sigma) > \bar{n})\} \middle| \{\mathbf{A}(\sigma) > 0\} \right) = \frac{[(1 - H_2(\bar{n}))\mathscr{P}^1(P,P) + 2(1 - H_1(\bar{n}))\mathscr{P}^1(P,NP)]}{1 - \mathscr{P}^1(NP,NP)}.$$

Then, by Lemma 2, there exists $\alpha > 0$ such that

$$\begin{aligned}
&\mathscr{P}^1\left(\{\mathbf{A}(\sigma) > \bar{n})\}\Big|\{\mathbf{A}(\sigma) > 0\}\right) \\
&= \frac{\left[(1 - H_2(\bar{n}))(\widehat{\mathscr{P}}^1(P,P) + \alpha) + 2(1 - H_1(\bar{n}))(\widehat{\mathscr{P}}^1(P,NP) - \alpha)\right]}{1 - \widehat{\mathscr{P}}^1(NP,NP) - \alpha} \\
&= \underbrace{\frac{\widehat{\mathscr{P}}^1(P,P) + \alpha}{1 - \widehat{\mathscr{P}}^1(NP,NP) - \alpha}}_{\text{Increasing in } \alpha}(1 - H_2(\bar{n})) + \underbrace{\left(1 - \frac{\widehat{\mathscr{P}}^1(P,P) + \alpha}{1 - \widehat{\mathscr{P}}^1(NP,NP) - \alpha}\right)}_{\text{Decreasing in } \alpha}(1 - H_1(\bar{n})))
\end{aligned}$$

The equality follows from noting that $(\widehat{\mathscr{P}}^1(P,P) + \alpha) + 2(\widehat{\mathscr{P}}^1(P,NP) - \alpha) = 1 - \widehat{\mathscr{P}}^1(NP,NP) - \alpha$. Notice that the above expression is a convex combination of $1 - H_2(\bar{n})$ and $1 - H_1(\bar{n})$. Since $\mathbf{N}_1, \mathbf{N}_2 \geqslant 0$ a.s., $H_2(\cdot) \leqslant H_1(\cdot)$. Therefore, the LHS puts a larger weight on the larger term for any $\alpha \geqslant 0$, which implies

$$\begin{aligned}
&\mathscr{P}^1\left(\{\mathbf{A}(\sigma) > \bar{n})\}\Big|\{\mathbf{A}(\sigma) > 0\}\right) \\
&> \frac{\widehat{\mathscr{P}}^1(P,P)}{1 - \widehat{\mathscr{P}}^1(NP,NP)}(1 - H_2(\bar{n})) + \left(1 - \frac{\widehat{\mathscr{P}}^1(P,P)}{1 - \widehat{\mathscr{P}}^1(NP,NP)}\right)(1 - H_1(\bar{n})) \\
&= \widehat{\mathscr{P}}^1\left(\{\mathbf{A}(\sigma) > \bar{n})\}\Big|\{\mathbf{A}(\sigma) > 0\}\right).
\end{aligned}$$

□

A.7. Proof of Proposition 3

Proof. Suppose that $k^* < \bar{n}$. Then, $\sigma^1 \in \mathcal{E}(\widehat{\mathscr{P}})$ implies that $\hat{\gamma}_1^1(\bar{n}) \geqslant \frac{c}{\mu(1)}$. By CAD, $\gamma_1^1(\bar{n}) \geqslant \hat{\gamma}_1^1(\bar{n})$. Therefore, (IC-voting) continues to be satisfied under $\mathscr{P}$, and hence $\sigma^1 \in \mathcal{E}(\mathscr{P})$.

On the other hand, if $k^* \geqslant \bar{n}$, then $\gamma_1^1(\bar{n}) \leqslant \hat{\gamma}_1^1(\bar{n})$. Therefore, it is possible that $\sigma^1 \in \mathcal{E}(\widehat{\mathscr{P}})$ but $\sigma^1 \notin \mathcal{E}(\mathscr{P})$. Hence, part (2) of the proposition follows. □

A.8. Proof of Proposition 4

Proof. Consider $k^* < \bar{n}^*(\widehat{\mathscr{P}})$. Then, by definition,

$$\gamma_1^1(\bar{n}^*(\widehat{\mathscr{P}})) \geqslant \hat{\gamma}_1^1(\bar{n}^*(\widehat{\mathscr{P}})) \geqslant \frac{c}{\mu(1)}.$$

This means under the same policy threshold $\bar{n}^*(\widehat{\mathscr{P}})$, the incentive constraint is satisfied even under more similar experiences $(\mathscr{P})$. Since the designer's objective $\mathscr{P}^1(I \geqslant \bar{n} + 1)$ is decreasing in $\bar{n}$, we have $\bar{n}^*(\mathscr{P}) \leqslant \bar{n}^*(\widehat{\mathscr{P}})$.

Next, consider $k^* \geqslant \bar{n}^*(\widehat{\mathscr{P}})$. Recall that $\bar{n}^*(\widehat{\mathscr{P}})$ is the lowest $\bar{n}$ that satisfies the incentive constraint under $\widehat{\mathscr{P}}$. Therefore, for any $\bar{n} < \bar{n}^*(\widehat{\mathscr{P}})$,

$$\widehat{\gamma}_1^1(\bar{n}) < \frac{c}{\mu(1)}.$$

Since $k^* \geqslant \bar{n}^*(\widehat{\mathscr{P}}) > \bar{n}$, by definition,

$$\gamma_1^1(\bar{n}) \leqslant \widehat{\gamma}_1^1(\bar{n}) < \frac{c}{\mu(1)}.$$

This means for any policy $\bar{n} < \bar{n}^*(\widehat{\mathscr{P}})$, under more similar experiences ($\mathscr{P}$), the incentive constraint does not hold. Moreover, since $\gamma_1^1(\bar{n}^*(\widehat{\mathscr{P}})) \leqslant \widehat{\gamma}_1^1(\bar{n}^*(\widehat{\mathscr{P}}))$, even under policy $\bar{n}^*(\widehat{\mathscr{P}})$, the incentive constraint may no longer be satisfied under more similar experiences ($\mathscr{P}$). Therefore, $\bar{n}^*(\mathscr{P}) \geqslant \bar{n}^*(\widehat{\mathscr{P}})$. $\square$

B. Online Appendix

B.1. On state-dependent changes in similarity

Suppose that $\Theta = \{-1, 1\}$. We keep the model unchanged in all respects otherwise. Now a regime change when $\boldsymbol{\theta} = -1$ is welfare-reducing. We demonstrate how increasing similarity in state $\boldsymbol{\theta} = -1$ has the opposite effects from what we showed in the main paper. To this end, we specialize to symmetric strategies. Then, given any strategy σ, we have the associated participation and nonparticipation sets given by $P(\sigma)$ and $NP(\sigma)$ respectively.

For σ to be an equilibrium, the IC constraints for protesting and not-protesting are:

$$\begin{aligned}&\mu(x)\left[\mathscr{P}_x^1(P)\Lambda_2 + (1-\mathscr{P}_x^1(P))\Lambda_1\right]\\ &-(1-\mu(x))\left[\mathscr{P}_x^{-1}(P)\Lambda_2 + (1-\mathscr{P}_x^{-1}(P))\Lambda_1\right] \geqslant c \qquad \text{if } x \in P \end{aligned} \tag{IC:P-S}$$

$$\mu(x)\mathscr{P}_x^1(P)\Lambda - (1-\mu(x))\mathscr{P}_x^{-1}(P)\Lambda \leqslant c \qquad \text{if } x \notin P \tag{IC:NP-S}$$

The only difference from our benchmark setup is the second term in the incentive constraints. This captures the probability of making a difference in state $\boldsymbol{\theta} = -1$. It is straightforward to see that an increase in similarity in state 1 (i.e., CAD increases of $\mathscr{P}^1$) has the same impact as in the main paper (for the natural modification of Condition M for this environment). But now consider the effects of increases in similarity in $\mathscr{P}^{-1}$. We can interpret

$$\mathscr{P}_x^{-1}(P)\Lambda_2 + (1-\mathscr{P}_x^{-1}(P))\Lambda_1$$

as the cost of making a difference in state $\theta = -1$ for a participant. In the encouragement (discouragement) environment, an increase in similarity increases (decreases) this cost, thus reducing (increasing) the incentive for participation among participants. In other words, CAD increase of $\mathscr{P}^{-1}$ has the opposite impact, compared to CAD increases of $\mathscr{P}^1$, on the incentive of the participants ((IC:P-S)).

For nonparticipants ($x \notin P$), higher similarity in state $\boldsymbol{\theta} = 1$ reduces the LHS in (IC:NP-S) while higher similarity in state $\boldsymbol{\theta} = -1$ increases it. Under the following assumption, the incentive constraint of the nonparticipants is always satisfied regardless of $\mathscr{P}^{-1}$.

Assumption 3: *For any σ with $\mathscr{V}(\sigma) > \mathscr{V}(\sigma^*)$, and any $x \notin P(\sigma)$,*

$$\mu(x)\mathscr{P}_x^1(P(\sigma))\Lambda < c.$$

We can again use CAD to characterize the effect of information similarity.

Proposition 6: *Suppose $\widehat{\mathscr{P}}^1$ satisfies assumption 3. Let $\mathscr{P} := (\widehat{\mathscr{P}}^1, \mathscr{P}^{-1})$ be an information structure such that $\mathscr{P}^{-1} \succcurlyeq_{CAD} \widehat{\mathscr{P}}^{-1}$, then*

1. *$\mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}})$ if $\Lambda_1 > \Lambda_2$. And,*
2. *$\mathscr{V}^*(\mathscr{P}) \leqslant \mathscr{V}^*(\widehat{\mathscr{P}})$ if $\Lambda_2 > \Lambda_1$.*

Proof. Let σ^* be a maximal equilibrium under $\mathscr{P}$ with associated participation and nonparticipation sets P^* and NP^* respectively.

Suppose that $\Lambda_1 > \Lambda_2$. Then, for all $x \in P^*$

$$\mathscr{P}_x^{-1}(P^*)\Lambda_2 + (1 - \mathscr{P}_x^{-1}(P^*))\Lambda_1 \leqslant \widehat{\mathscr{P}}_x^{-1}(P^*)\Lambda_2 + (1 - \widehat{\mathscr{P}}_x^{-1}(P^*))\Lambda_1$$

And, for all $x \notin P^*$, $\mu(x)\mathscr{P}_x^1(P^*) < c$ by Assumption 3. Therefore, $\sigma^* \in \mathcal{E}(\mathscr{P})$, proving part (1).

For (2), for all $x \in P^*$,

$$\mathscr{P}_x^{-1}(P^*)\Lambda_2 + (1 - \mathscr{P}_x^{-1}(P^*))\Lambda_1 \geqslant \widehat{\mathscr{P}}_x^{-1}(P^*)\Lambda_2 + (1 - \widehat{\mathscr{P}}_x^{-1}(P^*))\Lambda_1.$$

Therefore, σ^* may not be in $\mathcal{E}(\mathscr{P})$. Finally, consider any strategy profile σ such that $\mathscr{V}(\sigma) \geqslant \mathscr{V}(\sigma^*)$. By Assumption 3, (IC:NP-S) holds for all $x \in P$ in both, $\widehat{\mathscr{P}}$ and $\mathscr{P}$ (Since $\mathscr{P}^1 = \widehat{\mathscr{P}}^1$). Since σ^* is a maximal equilibrium under $\widehat{\mathscr{P}}$, $\sigma \notin \mathcal{E}(\widehat{\mathscr{P}})$. Therefore, $\exists x \in P$ for whom (IC:P-S) fails. Finally, for all $x \in P$,

$$\mathscr{P}_x^{-1}(P)\Lambda_2 + (1 - \mathscr{P}_x^{-1}(P))\Lambda_1 \geqslant \widehat{\mathscr{P}}_x^{-1}(P)\Lambda_2 + (1 - \widehat{\mathscr{P}}_x^{-1}(P))\Lambda_1.$$

Therefore, $\sigma \notin \mathcal{E}(\mathscr{P})$. The proposition follows.

□

B.2. On more than 2 groups

To establish that the qualitative results extend to settings with more than two groups, we consider a specialized environment in which the group sizes $\mathbf{N}_g$ are drawn from Poisson distributions with means N and the resilience $\bar{\boldsymbol{n}}$ is a deterministic $\bar{n}$, and restrict ourselves to symmetric equilibrium. Let $\psi(k, N) = \frac{\exp(-N)N^k}{k!}$ be the probability that nature chooses a group to have k agents. We use the following strong notion of similarity for n-dimensional random variables with $n > 2$, given by Meyer (1990). Let $\mathcal{Y}$ be a finite subset of $\mathbb{R}^n$, and let $\mathscr{Y}$ be a set of $\mathcal{Y}$-valued random variables with exchangeable distributions, i.e., the marginal distribution of Y_i and Y_j are equal for any i, j.

Definition 6: *Let $Y, \widehat{Y} \in \mathscr{Y}$ be two random variables with distributions $\mathscr{D}, \widehat{\mathscr{D}}$ respectively. Then, $Y \succcurlyeq_{sCAD} \widehat{Y}$ if,*

$$\mathscr{D}(\{Y_1 = y_{i_1}, Y_2 = y_{i_2}, \ldots, Y_n = y_{i_n}\}) \leqslant \widehat{\mathscr{D}}(\{\widehat{Y}_1 = y_{i_1}, \widehat{Y}_2 = y_{i_2}, \ldots, \widehat{Y}_n = y_{i_n}\})$$

for all $(i_1, \ldots, i_n)$ *for which it is not the case that* $i_1 = i_2 = \ldots = i_n$.

With this order, we can obtain results analogous to Theorem 1 and 2.

Proposition 7: *Suppose that* $\mathscr{P}^1 \succcurlyeq_{sCAD} \widehat{\mathscr{P}}^1$.

1. *If* $\psi(\bar{n}, GN) \geqslant \psi(\bar{n}, kN)$ *for all* $k < G-1$, *then* $\mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}})$.
2. *If* $\psi(\bar{n}, GN) < \psi(\bar{n}, kN)$ *for all* $k < G-1$, *and* $\widehat{\mathscr{P}}$ *satisfies condition M, then* $\mathscr{V}^*(\mathscr{P}) \leqslant \mathscr{V}^*(\widehat{\mathscr{P}})$.

Proof. For any set $T \subset \mathcal{X}$, we define $I(T)$ as the random variable denoting the number of other groups that receive a signal $x \in T$. Formally, for any group $g = 1$ (say), $I(T) := \sum_{g \neq 1} \mathbb{1}_{\mathbf{X}_g \in T}$. Abusing notation, we define $\mathscr{P}^1_x(T,k) = \mathscr{P}(I(T) = k | \mathbf{X}_1 = x, \boldsymbol{\theta} = 1)$ as the probability that $k = 0, 1, \ldots G-1$ other groups see a signal in T when group $g = 1$ (say) sees the signal x and the state is $\boldsymbol{\theta} = 1$. Notice that $\mathscr{P}^1 \succcurlyeq_{sCAD} \widehat{\mathscr{P}}^1$ implies that

1. For $x \in T$, $\mathscr{P}^1_x(T,k) \leqslant \widehat{\mathscr{P}}^1_x(T,k)$ for all $k = 0, 1, 2 \ldots G-2$ and $\mathscr{P}^1_x(T, G-1) \geqslant \widehat{\mathscr{P}}^1_x(T, G-1)$.
2. For $x \notin T$, $\mathscr{P}^1_x(T,k) \leqslant \widehat{\mathscr{P}}^1_x(T,k)$ for all $k = 0, 1, 2 \ldots G-1$

Let σ^* be the maximal participation equilibrium in $\widehat{\mathscr{P}}$, and let P^* and NP^* be the associated participation and not-participation sets respectively. For more than 2 groups, the IC for protesting and not-protesting can be modified as follows

$$\sum_{k=0}^{G-1} \widehat{\mathscr{P}}^1_x(T,k)\psi(\bar{n}, (k+1)N) \geqslant \frac{c}{\mu(x)} \quad \text{if } x \in P^* \tag{IC:P-G}$$

$$\sum_{k=0}^{G-1} \widehat{\mathscr{P}}^1_x(T,k)\psi(\bar{n}, kN) \leqslant \frac{c}{\mu(x)} \quad \text{if } x \in NP^* \tag{IC:NP-G}$$

It is easy to see that when the similarity increases ($\mathscr{P}^1 \succcurlyeq_{sCAD} \widehat{\mathscr{P}}^1$), if $\psi(\bar{n}, GN) \geqslant \psi(\bar{n}, kN)$ for all $k < G-1$, then the LHS increases in (IC:P-G) and decreases in (IC:NP-G). Therefore, $\sigma^* \in \mathcal{E}(\mathscr{P})$ and accordingly $\mathscr{V}^*(\mathscr{P}) \geqslant \mathscr{V}^*(\widehat{\mathscr{P}})$. On the other hand, if $\psi(\bar{n}, GN) \leqslant \psi(\bar{n}, kN)$ for all $k < G-1$, then under $\mathscr{P}^1$ the LHS decreases in (IC:P-G), which can violate the incentive of the participant, and accordingly, σ^* may no longer be an equilibrium. Moreover, given that $\widehat{\mathscr{P}}$ satisfies Condition M, given any σ such that $P(\sigma) \geqslant P^*$, $\exists x \in P(\sigma)$ such that,

$$\begin{aligned}
&\sum_{k=0}^{G-1} \widehat{\mathscr{P}}^1_x(T,k)\psi(\bar{n}, (k+1)N) < \frac{c}{\mu(x)} \\
\implies &\sum_{k=0}^{G-1} \mathscr{P}^1_x(T,k)\psi(\bar{n}, (k+1)N) < \frac{c}{\mu(x)}.
\end{aligned}$$

Therefore, such σ cannot constitute an equilibrium under $\mathscr{P}^1$, and accordingly, $\mathscr{V}^*(\mathscr{P}) \leqslant \mathscr{V}^*(\widehat{\mathscr{P}})$ with the inequality being strict whenever $\sigma^* \notin \mathcal{E}(\mathscr{P})$. $\square$

B.3. On mixed strategies

Below, in a simple two-player, two-signal, and deterministic threshold setting, we demonstrate that increasing similarity can reduce turnout in the maximal participation equilibrium when agents can play mixed strategies. Formally, consider $\mathbf{N}_1 = \mathbf{N}_2 = 1$ a.s., $\bar{\boldsymbol{n}} = 1$ a.s., and $\mathcal{X} = \{0, 1\}$, i.e., a two-player game where participation by one person is enough to change the regime. Evidently, $\Lambda_2 = 0 < 1 = \Lambda_1$, and therefore, we are in the discouragement environment.

We further assume $\mu(0) < c$, which means that the agents will not participate when they receive a signal 0, and $\mu(1)(1 - \mathscr{P}_1^1(\{1\})) < c$, which means always participating on receiving signal of 1 is not an equilibrium. A symmetric mixed strategy equilibrium exists, in which agents participate with probability β when they receive signal 1, where

$$\beta = \frac{\mu(1) - c}{\mu(1)\mathscr{P}_1^1(\{1\})}.$$

With increased similarity, $\mathscr{P}_1^1(\{1\})$ increases, and therefore, β decreases. That is, the expected turnout or the probability of a regime change reduces when regime change is beneficial ($\boldsymbol{\theta} = 1$). A similar argument (with additional restrictions) shows that expected participation may go down in the discouragement environment even if the group sizes are distributed according to Poisson distribution.

B.4. Informativeness of turnout

Consider our baseline model with two groups, two states, and a finite set of signals. For simplicity, we assume that $\mathbf{N}_1, \mathbf{N}_2$ are Poisson distributed with mean N; that is, let $\psi(k, N) = \frac{\exp(-N)N^k}{k!}$ be the probability that nature chooses a group to have k agents. However, now assume that a strategic policymaker observes the realized turnout and then decides whether to change the regime. We restrict attention to symmetric strategies, and as before, we denote the associated participation and non-participation sets of any strategy by P and NP, respectively.

Given an information structure $\mathscr{P}$, agents' strategy σ, and aggregate turnout $\mathbf{A}$, the policymaker's belief about the state of the world is given by the likelihood function

$$\beta(\cdot; \mathscr{P}, P) := \frac{Prob(\boldsymbol{\theta} = 1 | \mathbf{A} = \cdot, P)}{1 - Prob(\boldsymbol{\theta} = 1 | \mathbf{A} = \cdot, P)}.$$

The policymaker changes the status quo only if she is sufficiently confident that the state is 1; that is, there is a cutoff belief, $\underline{\beta}$, such that the policymaker changes the status quo if $\beta(k) \geqslant \underline{\beta} > 0$. Therefore, the policymaker's preferences are (ordinally) aligned with the citizens'.

We define informativeness of turnout, given a strategy σ, and its associated participation set P, as in Ekmekci and Lauermann (2019):

$$I(P) := \mathbb{P}^1(P) - \mathbb{P}^0(P)$$

Define $\bar{P} := \{x \in \mathcal{X} : \mathbb{P}^1(\{x\}) > \mathbb{P}^0(\{x\})$. It is easy to see that the informativeness of any strategy is bounded $I(\bar{P})$. Therefore, for a given information structure $\mathscr{P}$ with fixed marginals, we say that *information aggregates* if $\mathbb{1}_{\bar{P}}$ is an equilibrium under $\mathscr{P}$.

We fix $\mathscr{P}^0$ to investigate when, if at all, increasing similarity of information facilitates information aggregation. To this end, define

$$\begin{aligned} \underline{l} &:= \frac{\mu_1}{1-\mu_1}\frac{\mathbb{P}^1(\bar{P})^2}{\mathscr{P}^0(\bar{P},\bar{P})} \\ \bar{l} &:= \frac{\mu_1}{1-\mu_1}\frac{\mathbb{P}^1(\bar{P})}{\mathscr{P}^0(\bar{P},\bar{P})}. \end{aligned}$$

Proposition 8: *Suppose that $\mathbb{P}^1(\bar{P})\mathbb{P}^0(\bar{P}) > \mathscr{P}^0(\bar{P},\bar{P})$.*[23]

1. *If $\underline{l} \leqslant \underline{\beta} < \bar{l}$, then information does not aggregate if $\mathscr{P}^1$ has conditionally independent signals (denoted by $\mathscr{P}^{1,CI}$) as long as $c > 0$; and $\exists c > 0$ and a signal $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$, such that information aggregates under $\mathscr{P}^1$.*
2. *If $\underline{\beta} > \bar{l}$, then information does not aggregate for any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$ and any $c > 0$.*

Proof of Proposition 8. Substituting the expression for Poisson pdf, we get

$$\begin{aligned} \beta(k,\mathscr{P},\bar{P}) &= \frac{\mu_1}{1-\mu_1}\frac{\mathscr{P}^1(\bar{P},\bar{P})\psi(k,2N) + 2\mathscr{P}^1(\bar{P},N\bar{P})\psi(k,N)}{\mathscr{P}^0(\bar{P},\bar{P})\psi(k,2N) + 2\mathscr{P}^0(\bar{P},N\bar{P})\psi(k,N)} \\ &= \frac{\mu_1}{1-\mu_1}\frac{\mathbb{P}^1(\bar{P}) + \mathscr{P}^1(\bar{P},\bar{P})(e^{-N}2^{k-1}-1)}{\mathbb{P}^0(\bar{P}) + \mathscr{P}^0(\bar{P},\bar{P})(e^{-N}2^{k-1}-1)} \end{aligned}$$

Since 2^{k-1} is increasing in k,

$$\text{sign}\left(\frac{\partial\beta(k,\mathscr{P},\bar{P})}{\partial k}\right) = \text{sign}(\mathscr{P}^1(\bar{P},\bar{P})\mathbb{P}^0(\bar{P}) - \mathbb{P}^1(\bar{P})\mathscr{P}^0(\bar{P},\bar{P})).$$

When signals are conditionally independent in state 1, $\mathscr{P}^1(\bar{P},\bar{P}) = \mathbb{P}^1(\bar{P})^2$. Moreover, for any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$, $\mathscr{P}^1(\bar{P},\bar{P}) \geqslant \mathbb{P}^1(\bar{P})^2$. This implies

$$\text{sign}\left(\frac{\partial\beta(k,\mathscr{P},\bar{P})}{\partial k}\right) = \text{sign}(\mathbb{P}^1(\bar{P})\mathbb{P}^0(\bar{P}) - \mathscr{P}^0(\bar{P},\bar{P})) > 0.$$

The last inequality is true by hypothesis. Therefore, for any $\mathscr{P}$ such that $\mathscr{P}^1 \succcurlyeq_{CAD}$

[23] It is easy to generate examples in which this inequality is satisfied. For example, this inequality holds if the signals are independent conditional on the state.

$\mathscr{P}^{1,CI}$, $\beta(k, \mathscr{P}, \bar{P})$ is increasing in k. Moreover,

$$\lim_{k\to\infty} \beta(k, \mathscr{P}, \bar{P}) = \frac{\mu_1}{1-\mu_1} \frac{\mathscr{P}^1(\bar{P}, \bar{P})}{\mathscr{P}^0(\bar{P}, \bar{P})}$$

When $\mathscr{P}^1 = \mathscr{P}^{1,CI}$,

$$\lim_{k\to\infty} \beta(k, \mathscr{P}, \bar{P}) = \underline{l} < \underline{\beta}$$

Therefore, $\beta(k, \mathscr{P}, \bar{P}) < \underline{\beta}$ for all k whenever $\mathscr{P}^1 = \mathscr{P}^{1,CI}$. Therefore, $\mathbb{1}_{\bar{P}} \notin \mathcal{E}(\mathscr{P})$ if $\mathscr{P}^1 = \mathscr{P}^{1,CI}$ as long as $\underline{c} > 0$. That is, information aggregation fails when signals are conditionally independent.

In contrast, when $\mathscr{P}^1 = \mathscr{P}^{1,corr}$, where $\mathscr{P}^{1,corr}$ means signals being fully correlated in state 1,

$$\lim_{k\to\infty} \beta(k, \mathscr{P}, \bar{P}) = \bar{l}.$$

If $\bar{l} > \underline{\beta}$, $\exists k^*$ such that $\beta(k, \mathscr{P}, \bar{P}) > \underline{\beta}$ for all $k \geqslant k^*$. For any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$, by Lemma 2, $\mathscr{P}^1(\bar{P}, \bar{P}) = \mathscr{P}^{1,CI}(\bar{P}, \bar{P}) + \alpha$, for some $\alpha \geqslant 0$. We know that, when $\alpha = 0$, $\beta(k, \mathscr{P}, \bar{P}) < \underline{\beta}$ for all k, and, when $\alpha = \mathbb{P}^1(\bar{P}) - \mathbb{P}^1(\bar{P})^2$, $\exists k^* \in \mathbb{N}$ such that $\beta(k, \mathscr{P}, \bar{P}) > \underline{\beta}$ whenever $k \geqslant k^*$. Therefore, we can choose an $\alpha > 0$ small enough to construct $\mathscr{P}^1$ so that $\mathscr{P}^1(\bar{P}, \bar{P}) = \mathscr{P}^{1,CI}(\bar{P}, \bar{P}) + \alpha$ and $\beta(k, \mathscr{P}, \bar{P}) \geqslant \underline{\beta}$ if and only if $k > k^* > 2N$. Therefore, the policymaker would use a threshold of $\bar{n} = k^* > 2N$ when $\mathscr{P}^1$ constructed using α described above. It is easy to check that $\psi(\bar{n}, 2N) > \psi(\bar{n}, N)$ in this case. Claim 2 then establishes that,

$$\min_{x\in\bar{P}} \mu(x) \left[\mathscr{P}^1_x(\bar{P})\psi(\bar{n}, 2N) + \mathscr{P}^1_x(\mathcal{X}\backslash\bar{P})\psi(\bar{n}, N)\right] > \max_{x\in\mathcal{X}\backslash\bar{P}} \mu(x)\mathscr{P}^1_x(\bar{P})\psi(\bar{n}, N)$$

Therefore, by letting c to be strictly between the LHS and the RHS of the above, we get that $\mathbb{1}_{\bar{P}}$ is an equilibrium, for it satisfies (IC:P) for all $x \in \bar{P}$ and (IC:NP) for all $x \in \mathcal{X}\backslash\bar{P}$. Therefore, information aggregates under $\mathscr{P}$ wherein $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$.

Finally, if $\underline{\beta} > \bar{l}$, then the policymaker would not change the status quo regardless of the turnout for any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$ establishing the last part of the Proposition.

Claim 2: *If $\psi(\bar{n}, 2N) > \psi(\bar{n}, N)$, then*

$$\min_{x\in\bar{P}} \mu(x) \left[\mathscr{P}^1_x(\bar{P})\psi(\bar{n}, 2N) + \mathscr{P}^1_x(\mathcal{X}\backslash\bar{P})\psi(\bar{n}, N)\right] > \max_{x\in\mathcal{X}\backslash\bar{P}} \mu(x)\mathscr{P}^1_x(\bar{P})\psi(\bar{n}, N)$$

for any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$.

Proof. First, by definition of $\bar{P}$, $\min_{x\in\bar{P}} \mu(x) > \max_{x\in\mathcal{X}\backslash\bar{P}} \mu(x)$. Also, $\mathscr{P}^{1,CI}_x(\bar{P}) = \mathbb{P}^1(\bar{P})$ is independent of x, and, $\mathscr{P}^1_x(\bar{P}) \geqslant \mathbb{P}^1(\bar{P})$ for all $x \in \bar{P}$ and $\mathscr{P}^1_x(\bar{P}) \leqslant \mathbb{P}^1(\bar{P})$ for all $x \notin \bar{P}$. Therefore, $\min_{x\in\bar{P}} \mathscr{P}^1_x(\bar{P}) \geqslant \max_{x\in\mathcal{X}\backslash\bar{P}} \mathscr{P}^1_x(\bar{P})$. The claim, then, follows due to $\psi(\bar{n}, 2N) > \psi(\bar{n}, N)$.

$\square$

$\square$

Given an information structure $\mathscr{P}$, we say an equilibrium σ^* has "*maximally informative turnout*" if $I(\sigma^*) \geqslant I(\sigma)$ for all $\sigma \in \mathcal{E}(\mathscr{P})$. We denote $I(\sigma^*)$ by $\mathcal{I}(\mathscr{P})$. Proposition 9 below, shows that increasing similarity of information can reduce the informativeness of turnout.

Proposition 9: *Suppose that $\mathscr{P}^1 \succcurlyeq_{CAD} \widehat{\mathscr{P}}^1$ and $\mathscr{P}^0 = \widehat{\mathscr{P}}^0$. Let $\bar{n}^*$ be the optimal participation threshold for the equilibrium with maximally informative turnout under $\widehat{\mathscr{P}}$.*

1. *If $\psi(\bar{n}^*, 2N) > 2\psi(\bar{n}^*, N)$, then $\mathcal{I}(\mathscr{P}) \geqslant \mathcal{I}(\widehat{\mathscr{P}})$ if $\max_{T \subset \mathcal{X}} \mathscr{P}^1(T,T) - \widehat{\mathscr{P}}^1(T,T)$ is sufficiently small.*
2. *If $\psi(\bar{n}^*, 2N) < \psi(\bar{n}^*, N)$, then it is possible that $\mathcal{I}(\mathscr{P}) < \mathcal{I}(\widehat{\mathscr{P}})$.*

Proof of Proposition 9. Let σ^* be the maximally informative equilibrium under $\widehat{\mathscr{P}}$. If the policymaker continues to use $\bar{n}^*$ as the cutoff, then σ^* continues to remain an equilibrium under $\mathscr{P}$ due to Theorem 1. While this takes care of the incentives of the participants, unlike the earlier arguments, we also need to ensure that a cutoff of $\bar{n}^*$ is indeed a best response for the policymaker. Since $\bar{n}^*$ is the cutoff for the maximally informative equilibrium $\beta(\bar{n}^*; \widehat{\mathscr{P}}) \geqslant \underline{\beta}$ and $\beta(k; \widehat{\mathscr{P}}) < \underline{\beta}$ for all $k < \bar{n}^*$. By Bayes' rule, we have,

$$\frac{\beta(k;\widehat{\mathscr{P}})}{1-\beta(k;\widehat{\mathscr{P}})} = \frac{\mu}{1-\mu}\frac{\widehat{\mathscr{P}}^1(P,P)\psi(k,2N) + 2\widehat{\mathscr{P}}^1(P,NP)\psi(k,N)}{\widehat{\mathscr{P}}^0(P,P)\psi(k,2N) + 2\widehat{\mathscr{P}}^0(P,NP)\psi(k,N)}$$

Since $\psi(\bar{n}^*, 2N) > 2\psi(\bar{n}^*, N)$, $\mathscr{P}^1(P,P) = \widehat{\mathscr{P}}^1(P,P) + \alpha$ and $\mathscr{P}^1(P,NP) = \widehat{\mathscr{P}}^1(P,NP) - \alpha$ for some $\alpha > 0$ by Lemma 2, and $\mathscr{P}^0 = \widehat{\mathscr{P}}^0$, $\beta(\bar{n}^*; \mathscr{P}) > \beta(\bar{n}^*; \widehat{\mathscr{P}}) \geqslant \underline{\beta}$. However, it is now also possible that $\beta(k; \mathscr{P}) \geqslant \underline{\beta}$ for some $k < \bar{n}$[8]. Simply lowering the threshold in this case is not an option either as it affects the incentives of the agents, possibly destroying σ^* as an equilibrium. However, when $\max_{T \subseteq \mathcal{X}} \mathscr{P}^(T,T) - \widehat{\mathscr{P}}^1(T,T)$ is sufficiently small, $\beta(k; \mathscr{P}) < \underline{\beta}$. Finally, since $\mathcal{I}(\cdot)$ only depends on the marginal distributions, we obtain the desired inequality.

For the second part, suppose that $N = 20$, $\mathcal{X} = \{0, 1\}$, $c = 0.0368$, and $\underline{\beta} = 0.7281$. Signals are conditionally independent in state 0 with the marginal distribution $\mathbb{P}^0(1) = 0.3$. In state 1, $\widehat{\mathscr{P}}^1(1,1) = 0.66$, $\widehat{\mathscr{P}}^1(1,0) = 0.15$. $\mathscr{P}^1$ is constructed from $\widehat{\mathscr{P}}^1$ by using $\alpha = 0.05$. It is easy to see that $\bar{n}^*(\widehat{\mathscr{P}}) = 28$, while the same no longer constitutes an equilibrium under $\mathscr{P}$. In this case, if an informative equilibrium exists, it must involve mixing. It is easy to check that mixing can only happen on signal 1, and agents continue to not participate when they receive a signal of 0. Therefore, informativeness under $\mathscr{P}$ is strictly lower than under $\widehat{\mathscr{P}}$.

$\square$

B.5. On optimal information similarity

Consider two extreme cases: conditionally independent signals and perfectly correlated signals. Suppose $Y = (Y_1, Y_2)$ where Y_i is distributed according to $\mathbb{P}^1$, and Y_1, Y_2 are independent. Denote this joint distribution by $\mathscr{P}^{1,CI}$. Analogously, let $Y = (Y_1, Y_2)$ be a random variable such that $Y_1 = Y_2$ a.s., and Y_i is distributed according to $\mathbb{P}^1$. We denote this joint distribution by $\mathscr{P}^{1,corr}$. Given a conditionally independent signal distribution $\mathscr{P}^{1,CI}$, define

$$CI^{\uparrow} := \{\mathscr{D} \in \Delta(\mathcal{X} \times \mathcal{X}) : \mathscr{D} \succcurlyeq_{CAD} \mathscr{P}^{1,CI}\}$$

as all the signal distributions that are more similar (in the CAD sense) than $\mathscr{P}^{1,CI}$. Recall that by definition of CAD, all such distributions have the same marginal, and in this case, $\mathscr{P}^0$ does not affect $\mathscr{V}(\cdot)$. Therefore, the designer solves the following problem:

$$\sup_{\mathscr{P}^1 \in CI^{\uparrow}} \mathscr{V}^*(\mathscr{P}^1, \mathscr{P}^0)$$

Proposition 10 (Optimal information similarity)**:** *An optimal information structure exists. In the encouragement environment, fully correlated signals are optimal. In the discouragement environment, if conditionally independent signals satisfy Condition M, then they are optimal. In other cases, intermediate levels of similarity can be optimal.*

Proof. We prove this using three steps. Steps 1 and 2 establish the existence of an optimal information structure, while Step 3 describes it.

Step 1: We show that $CI^{\uparrow}$ is weak-* compact. Consider a sequence $\{\mathscr{D}_m\}$ from $CI^{\uparrow}$ that converges to $\mathscr{D}$ in the sense that for all $f \in C(\mathcal{X} \times \mathcal{X})$, $\int f \mathrm{d}\mathscr{D}_m \to \int f \mathrm{d}\mathscr{D}$. Consider a symmetric $\alpha \in \mathbb{R}_+^{\mathcal{X} \times \mathcal{X}}$, i.e., $\alpha(i,j) = \alpha(j,i)$ for all i, j, and $\widehat{\mathscr{D}} \in \Delta(\mathcal{X} \times \mathcal{X})$. Define,

$$\mathscr{D}(i,j) = \widehat{\mathscr{D}}(i,j) - \alpha(i,j)\mathbb{1}_{i \neq j} + \sum_{k \neq i} \alpha(i,k)\mathbb{1}_{i=j}$$

If $\mathscr{D} \in \Delta(\mathcal{X} \times \mathcal{X})$, then we say that "$\mathscr{D}$ is obtained from $\widehat{\mathscr{D}}$ by an ETI given by α", denoted by $\mathscr{D} = \widehat{\mathscr{D}} \biguplus \alpha$. Recall from the proof of Lemma 2 that an alternative characterization of the CAD order (from Proposition 1 in **?**) is

$$\mathscr{D} \succcurlyeq_{CAD} \widehat{\mathscr{D}} \iff \exists \alpha \in \mathbb{R}_+^{\mathcal{X} \times \mathcal{X}} \text{ such that } \mathscr{D} = \widehat{\mathscr{D}} \biguplus \alpha.$$

Since $\mathscr{D}_m \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$, we have a sequence $(\alpha_m) \in \mathbb{R}_+^{\mathcal{X} \times \mathcal{X}}$ such that $\mathscr{D}_m = \mathscr{P}^{1,CI} \biguplus \alpha_m$. Due to finiteness of $|\mathcal{X} \times \mathcal{X}|$, and boundedness of α_m, $\{\alpha_m\}$ has a convergent subsequence, $\{\alpha_{m_k}\}$. Let α be a limit of one such convergent subsequence. Let $\tilde{\mathscr{D}} := \mathscr{P}^{1,CI} \biguplus \alpha$. Suppose, for contradiction, that $\tilde{\mathscr{D}} \neq \mathscr{D}$. Then, there exists some $(x, y) \in \mathcal{X} \times \mathcal{X}$ such that $\tilde{\mathscr{D}}(x, y) \neq \mathscr{D}(x, y)$. Consider a continuous function $f \in C(\mathcal{X} \times \mathcal{X})$ such that $f(x', y') = 1$ if $x' = x$ and $y' = y$, and $f(x', y') = 0$, otherwise.[24] By construction, $\int_{\mathcal{X}\times\mathcal{X}} f \mathrm{d}\mathscr{D}_{m_k} \to \int_{\mathcal{X}\times\mathcal{X}} f \mathrm{d}\tilde{\mathscr{D}} \neq \int_{\mathcal{X}\times\mathcal{X}} f \mathrm{d}\mathscr{D}$. Hence, a contradiction. Therefore, $\tilde{\mathscr{D}} = \mathscr{D} = \mathscr{P}^{1,CI} \biguplus \alpha$, i.e., $\mathscr{D} \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$, and therefore, $\mathscr{D} \in CI^{\uparrow}$. Therefore, $CI^{\uparrow}$ is closed. Finally, since $\mathcal{X} \times \mathcal{X}$ is compact, $\Delta(\mathcal{X} \times \mathcal{X})$ is weak$-^*$ compact. This makes $CI^{\uparrow}$ weak$-^*$ compact.

Step 2: Given compactness, consider a sequence $\{\mathscr{D}_m\} \in CI^{\uparrow}$ such that $\mathscr{V}(\mathscr{D}_m) \to \bar{\mathscr{V}} = \sup_{\mathscr{D}\in CI^{\uparrow}} \mathscr{V}^*(\mathscr{D})$. By compactness, we can (wlog) assume $\mathscr{D}_m$ converges to some $\mathscr{D} \in CI^{\uparrow}$. For each $\mathscr{D}_m$, let σ^*_m be a maximal turnout equilibrium, with participation and not-participation sets (P^*_m, NP^*_m). Due to the finiteness of $\mathcal{X}$, $|2^{\mathcal{X}}|$ is finite, and therefore, wlog, $P^*_m = P^*$ for a sufficiently large m.[25] Since the marginals are unchanged, we have $\mathscr{V}(\mathscr{D}_m) = 2N\mathbb{P}^1(P^*)$ for a sufficiently large m. Therefore, $\mathscr{V}(\mathscr{D}_m) = \bar{\mathscr{V}}$ for a sufficiently large m.

Step 3: Since more similarity increases maximal equilibrium turnout in the encouragement environment (when $\Lambda_2 > \Lambda_1$), we have $\mathscr{V}^*(\mathscr{P}^{1,corr}, \mathscr{P}^0) \geqslant \mathscr{V}^*(\mathscr{P}^1, \mathscr{P}^0)$ for any $\mathscr{P}^1 \in CI^{\uparrow}$. In the discouragement environment (when $\Lambda_1 > \Lambda_2$), since $\mathscr{P}^{1,CI}$ satisfies Condition M, any $\mathscr{P}^1 \succcurlyeq_{CAD} \mathscr{P}^{1,CI}$ has $\mathscr{V}^*(\mathscr{P}^1, \mathscr{P}^0) \leqslant \mathscr{V}^*(\mathscr{P}^{1,CI}, \mathscr{P}^0)$. To see that intermediate levels of similarity can be optimal otherwise, consider the example below.

Example 1: $\mathbf{N}_1, \mathbf{N}_2$ *follow a Poisson distribution with mean* $N = 15$ *and deterministic threshold* $\bar{n} = 20$. *The cost of participation is* $c = 0.009$ *and the signal structure is as follows:* $\mathcal{X} = \{1, 2, 3\}$, $\mu_0 = \frac{1}{2}$, $\mathbb{P}^1 = [0.25, 0.3, 0.45], \mathbb{P}^0 = [0.6, 0.35, 0.05]$.

With $\mathscr{P}^{1,CI}$, the unique equilibrium is $\sigma = \mathbb{1}_{s=3}$. However, if we perform ETI on the square $\{(1, 1), (1, 2), (2, 1), (2, 2)\}$ with $\alpha = 0.005$, we can support an equilibrium $\sigma = \mathbb{1}_{s\in\{2,3\}}$. Finally, it is easy to see that $\mathscr{P}^{1,corr}$ also cannot support $\mathbb{1}_{s\in\{2,3\}}$ as an equilibrium. We can verify that $\mu = [0.2941, 0.4615, 0.9000]$. When $P = \{2, 3\}$, $\mathbb{P}^1(P) = 0.75 = \mathscr{P}^{1,CI}_x(P)$ for all x since signals are conditionally independent. $\psi(\bar{n}, N) = 0.0418, \psi(\bar{n}, 2N) = 0.0134$. □

[24] Such a function obviously exists because of the finiteness of $\mathcal{X}$.

[25] To be precise, it may be necessary to pass onto a subsequence for this to be true.