Decentralized Social Movement and Tullock's Paradox: Theory and Evidence from Protests in Hong Kong

Pui-Hang Wong
Maastricht University

Tsz-Nga Wong
FRB Richmond

This version: March 26, 2026.
Check the latest version

According to Tullock's paradox of revolution, mass movements cannot be explained by rational choice models with strategic substitutability, which is inconsistent with the recent field evidence. Motivated by the fact that many protests and movements are decentralized and leaderless nowadays, we devise a search-and-matching model of protest and voting which features endogenous strategic interaction and multiple threshold equilibria. In the "quiet" equilibrium, the comparative statics behave like the standard global games featuring strategic complementarity. In this case, protest turnouts are small, crackdown reduces turnouts with amplifiying marginal effects, and regime changes are rare. The opposite occurs in the "unrest" equilibrium that features strategic substitutability, which is plausible when the search frictions are low (e.g. via social media) and the goal of decentralized protests is modest. To test the model's implications, we construct a new database detailing the distribution of Facebook users and the deployment of tear gas during the 2019 protests in Hong Kong. We use the location of police stations to instrument crackdown. Results from the instrumental variable estimation support our model's implications and provide strong evidence of unrest equilibrium featuring strategic substitutability. Our paper provides a new resolution to Tullock's paradox.

JEL Classification: D72, D74, D82, D83

Keywords: search-and-matching, privotal protests, voting, global games, multiplicity.

Acknowledgments: The views expressed herein are those of the authors and not necessarily those of the Federal Reserve System or the Federal Reserve Bank of Richmond. We are grateful for the comments from Noam Yuchtman, Ping Wang and Wing Suen. We are also grateful for the comments from participants at ASSA, APSA and other seminars and conferences.

I. INTRODUCTION

Protests involve confrontation and are often perceived as deviant. Though dissidents almost always exist, they hide among the silent majority and face a constant challenge of finding linkminded people to spark a movement. Thus, a major barrier to collective actions is finding ways to overcome this 'frictional' environment. Historically, successful collective actions often benefited from skilled and charismatic organizers who mobilized people to join a movement (Olson 1965). Benefited from communication technology, protesters nowadays are easier to find one another through the internet. This can explain why protests are increasingly decentralized, in which telecommunication technologies have played an increasingly significant role in more recent movements (Castells 2015; Enikolopov et al. 2020; Manacorda and Tesei 2020). Social media has significantly contributed to the sucess of many contemporary movements. Acemoglu et al. (2018), for example, showed that the number of Twitter posts containing "Tahrir hashtags" predicted the protests in Tahrir Square, Egypt, during the Arab Spring.

In political economy literature, there is broad consensus that large-scale collective actions often emerge under conditions of strategic complementarities, and social media is thought to play a crucial role in coordinating participants' actions (Manacorda and Tesei 2020; Zhuravskaya et al. 2020). However, evidence of strategic substitutability has also been reported in some mass movements (Cantoni et al 2019; Hager et al 2022). This raises Tullock's paradox of revolution once again: if communication facilitates coordination, yet people tend to free ride under conditions of strategic substitutability, why are massive turnouts still observed in equilibrium, especially when the risks of repression are heightened, making an environment increasingly frictional and large-scale events more difficult to emerge?

To analyze the dynamics between search frictions and decentralized social movements, we introduce the first search-theoretic model of decentralized collective actions. The advantage of decentralized movements is that protesters can find and mobilize one another without the need for leaders. However, the downside lies in search frictions and strategic uncertainty among protesters, challenges that were mitigated by organizers before the age of social media. Similar to traditional models that characterize protests as a global game, individuals in our model observe a private signal about the fundamental and adopt a threshold strategy in equilibrium (Morris and Shin 1998; Shadmehr 2018). Unlike traditional models, however, search friction now endogenizes strategic interactions. Before the rise of social media, the flow of information was more limited, and search frictions were high. Under these circumstances, mobilization heavily depended on social movement organizations, and mass protests occurred only under strategic complementarity. The advent of social media reduces search frictions and introduces an additional equilibrium. When individuals perceive their participation as pivotal (Myatt 2017), they would join protests strategically, balanc-

ing the odds of achieveing a successful movement against the risk of triggering costly repression. We prove that in a low search-friction environment, mass mobilization can arise under strategic substitutability.

Our new model also yields distinct predictions regarding protest turnout in the presence of government crackdowns. The comparative statics enable us to derive testable implications, facilitating equilibrium identification and empirical testing of the model using real-world data. Specifically, our model predicts that when governments repress, turnout will decrease under strategic complementarity but increase under strategic substitutability. It also predicts that the marginal impact of crackdowns on turnout is positive but diminishing.

Using hand-collected data from social media and news reports, we test our model by examining the impact of tear gas deployment on support for pro-protest candidates during a local election held amid the 2019-2020 mass protests in Hong Kong. Since teat gas could be more likely to be deployed in pro-protester areas and thus potentially endogenous, we instrument tear gas deployment using the geographical distance between a constituency and its nearest police station. The instrument is relevant because tear gas deployment is more likely to occur in constituencies closer to police stations. It is also exogenous, as the location of police stations is historically predetermined and unrelated to the political alignment of an area. Estimates based on instrumental regressions reveal that tear gas deployment widens support for pro-protest candidates by 57 log points and raises their probability of winning in constituency where tear-gas was deployed by 33%. Results from a placebo test using data from the 2015 election data corroborate predictions of our model. Our findings align with a conclusion from Cantoni et al. (2019), who used data from a lab-in-the-field experiment during the 2016 Umbrella Movement and found that protest participation acted as a strategic substitute among university students in Hong Kong.

The model offers a solution to the paradox of revolution (Tullock 1971). A key assumption underlying this paradox is about people's belief about their marginal contribution to the success of an event is negligible. Similar to protest participation, the significance of a single vote to a candidate's victory is also marginal. If people hold this belief, they are unlikely to vote or participate. However, this assumption has been challenged by some empirical studies. For example, a laboratory experiment by Duffy and Tavits (2008) estimated that individuals' subjective probability of their votes being decisive can be as high as 0.33. This belief in being pivotal to the success of an endeavor provides a plausible explaination for why citizens choose to join a protest despite the possibility of free-riding. The belief also explains why repression may backfire. In the presence of strategic substitutability, the expectation that repression will reduce turnout can motivate individuals to show up, effecreduces countering the intended repression effect.

Our paper also contributes to two strands of literature on protests. First, the debate over whether massive turnout results from strategic substitutability or strategic complementarity re-

mains a focus in recent research. Since Olson (1965) and Tullock (1971), political activists have often been thought of as free-riders. While no empirical evidence have been found in the context of protests, this idea has recently been validated using data from a natural field experiment conducted among political activists during a political campaign in Western Europe (Hager et al., 2023). Evidence of strategic consideration has also been identified in the Umbrella Movement in Hong Kong (Cantoni et al., 2019; Bursztyn et al., 2021). Nevertheless, the assumption of strategic complementarity has often been made in some theoretical contributions (Bueno de Mesquita 2010; Edmond 2013; Passarelli and Tabellini 2017). Drawing on our model and data from a more recent protest in Hong Kong, we provide both theoretical support and empirical evidence for the role of strategic substitutability in this debate. Although our original data are observational, findings from our instrumental variable estimation encompass a broader population spectrum, addressing a limitation often in experimental studies.

Second, our paper is among the first studies in the economics literature to provide direct empirical evidence on the counter-productiveness of repression. Discussions on the relationship between repression and protest dynamics remain scant in the economics literature (Cantoni et al., 2024). To date, repression has primarily been studied for its role in maintaining regime stability through censorship and political terror (Shadmehr and Bernhardt, 2015), without a focus on its relationship with protests (Besley and Persson, 2011). Existing contributions are often theoretical and largely situated within political science (e.g. Shadmehr and Bernhardt, 2011). The backfire effect can be explained by the logic of strategic substitution and the perception of being pivotal. Citizens are more likely to protest against an incumbent regime when they believe that repression will deter participation and cause turnout to fall short of the threshold needed for success. Repression works when participation is strategic complement, as citizens believe they can hide among the crowd to reduce their chances of being caught. Conversely, when individuals think the outcome hinges on their involvement, especially if the required threshold is relatively low, they are more likely to show up. Our model captures this dynamics, which is further supported by additional anecdotal evidence. A central question is that if governments know that repression can backfire, why would they choose to do so in the first place? The existence of multiple equilibria in our model provides a plausible explanation. Governments do not know *ex ante* which equilibrium they are in until the first bullet is fired. The plausibility of strategic complementarity and substitution puts governments in a dilemma. Our study challenges the deterrence assumption and offers a novel mechanism to explain why repression may ultimately undermine the goals of an oppressive government

II. BASELINE MODEL

Decentralized movement. There is a large number, n, of citizens with heterogeneous payoffs from the outcomes of a decentralized movement opposing the incumbent government. The payoff

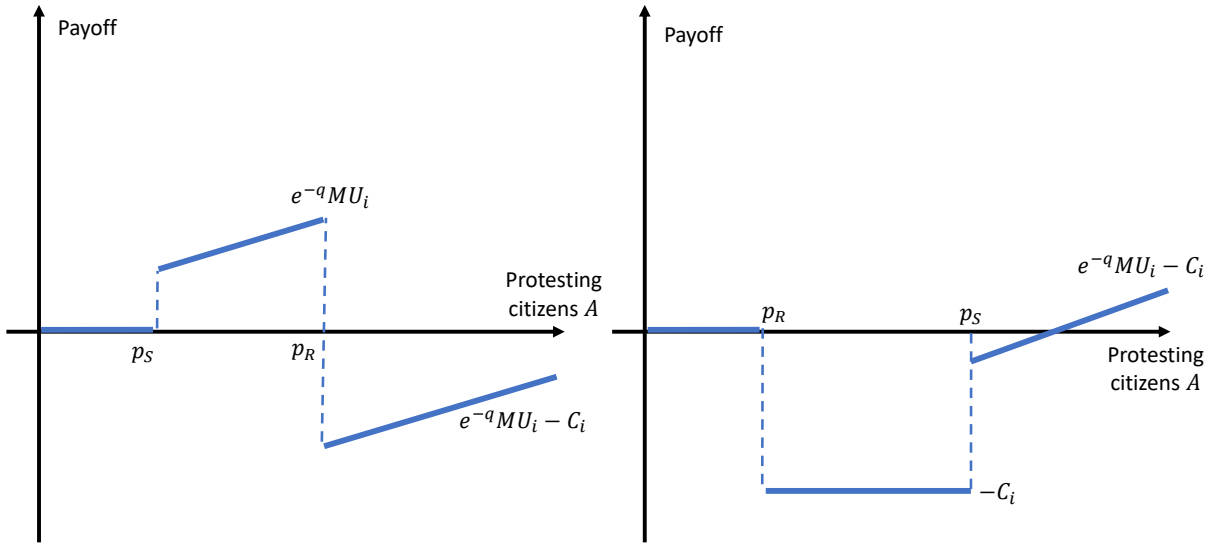


Figure 1: Individual payoff.

for citizen i is given by

$$\mathcal{U}_i(a_i) = e^{-q} \Pr_i \left\{ \frac{A}{n} > p_S \right\} M_i U_i - \Pr_i \left\{ \frac{A}{n} > p_R \right\} C_i, \text{ where } A = a_i + \sum_{j \neq i} a_j,$$

where $a_i \in \{0, 1\}$ indicates that citizen i's decision to protest. A decentralized movement is composed of a social network of protests, which can form through various channels such as schools, workplaces, neigbhorhoods, and online communities. Let M_i represent citizen i's expected effective fraction of protesting citizens within his social network (excluding himself). If the decentralized movement succeeds, citizen i receives a payoff of $M_i U_i$, which is increasing M_i, reflecting the idea that a higher expected effective fraction of protesters in his social network amplifies his marginal benefit of protest U_i.

A decentralized movement succeeds if the *total* number of protesters across social groups, $A \equiv \sum_i a_i$, exceeds the threshold $p_S \in (0, 1)$ (i.e., $A/n > p_S$). However, the government represses when the size of the protest becomes sufficiently large, such that $A/n > p_R$, where $p_R \in (0, 1)$. In this case, citizen i's incurs a loss of C_i, which varies across citizens. The left panel of Figure 1 illustrates the payoff when $p_S < p_R$, where the government tolerates some successful protests without repression. The contrasting case $p_S > p_R$ is illustrated in the right panel of Figure 1.

We model a decentralized movement using a random matching model. Each citizen engages with his social groups randomly with probability e^{-q}, where $q > 0$ represents the search friction, capturing the difficulty of meeting and networking with other citizens. The social network of citizen

$i$is represented by the adjacency matrix $G_{ij}^n \in [0,1]$, where G_{ij}^n measures the importance of citizen j in citizen i's social network, where $G_{ii}^n = 0$ and $\sum_j G_{ij}^n = 1$. The expected effective fraction of protesting citizens within citizen i's social network is

$$M_i = \mathbb{E}_i \sum_j G_{ij}^n 1_{a_j=1}.$$

Dispersed information. Although citizens are heterogenous, the "sufficient statistics" for a citizen's action u_i, as we will see later, is given by

$$u_i \equiv \log\left(\frac{U_i}{C_i}\right),$$

which is referred to as citizen i's private interest in protesting. A high level of private interest indicates that the citizen derives greater utility from participating in protests compared to abstaining, after accouting for the potential consequences of repression. Citizens' private interest are heterogenous, private, and follow a normal distribution:

$$u_i = \theta + \epsilon_i,\ \epsilon_i \sim N\left(0, \sigma^2\right),$$

where ϵ_i is i.i.d. normal among citizens and σ captures the dispersion of private interests among citizens. Additionally, the social interest in the protest is captured by θ, which reflects the aggregate state of this society. However, there is an aggregate uncertainty and a lack of common knowledge about the social interest, leading to strategic uncertainty among citizens. This encapsulates, for instance, the situation where citizen i is enthusiatic about protesting, but he is uncertain whether other individuals share the same enthusiatism. This uncertainty may cause hestitation to participate. The lack of common knowledge about social saliencies is particularly significant for decentralized movements. In particular, θ is given by

$$\theta = \bar{\theta} + \varepsilon,\ \varepsilon \sim N\left(0, \frac{\sigma^2}{\psi}\right),$$

where $\bar{\theta}$ is the prior interest. Unknown to citizens, the error term ε catpures the aggregate uncertainty regarding the social interest, which is common for all citizens. Specifically, ψ measures the confidence in the social interest relative to the private interest.

After observing his private interest u_i, citizen i's posterior belief about the social interest is

$$\theta|_{u_i} \sim N\left(\theta_i, \frac{\sigma^2}{1+\psi}\right),\ \text{where} \tag{1}$$

$$\theta_i \equiv \frac{\psi\bar{\theta} + u_i}{\psi + 1}. \tag{2}$$

Essentially, θ_i is citizen i's posterior on θ, which puts a weight of $\psi/(1+\psi)$ on the prior interest $\bar{\theta}$ and a weight of $1/(1+\psi)$ on his private interest u_i. The weight on private interest is smaller when a citizen has greater confidence ψ on the prior interest $\bar{\theta}$.

Pivotal protester. Citizens are pivotal, which means that they internalize the impact of their choice of a_i on A as $A = a_i + \sum_{j \neq i} a_j$. Rearraning the expectation operator, citizen i's payoff becomes[1]

$$\mathcal{U}_i(a_i) = e^{-q} \Pr_i \{A_{-i} + a_i > \lceil p_S n \rceil\} M_i U_i - \Pr_i \{A_{-i} + a_i > \lceil p_R n \rceil\} C_i$$

where A_{-i} is the random variable of the number of protesting citizens excluding citizen i. Denote citizen i's gain from protest as $\Delta_i \equiv \mathcal{U}_i(1) - \mathcal{U}_i(0)$, which is given by

$$\Delta_i \equiv \mathcal{U}_i(1) - \mathcal{U}_i(0) = e^{-q} P_i^S M_i U_i - P_i^R C_i, \tag{3}$$

where

$$P_i^S \equiv \Pr_i \{A_{-i} + 1 > \lceil p_S n \rceil\} - \Pr_i \{A_{-i} > \lceil p_S n \rceil\} = \Pr_i \{A_{-i} = \lceil p_S n \rceil\}, \tag{4}$$

$$P_i^R \equiv \Pr_i \{A_{-i} + 1 > \lceil p_R n \rceil\} - \Pr_i \{A_{-i} > \lceil p_R n \rceil\} = \Pr_i \{A_{-i} = \lceil p_R n \rceil\}. \tag{5}$$

We refer P_i^S and P_i^R as to the citizen i's *pivotal* probability of success and of repression respectively. In particular, P_i^S is the probability that citizen i is the marginal participant whose involvement determines the success of the protest. Similarly P_i^R is the probability that citizen i is the marginal participant whose involvement triggers government repression. Arranging terms, it is strictly optimal for a citizen i to protest upon observing u_i iff

$$\Delta_i > 0 \Leftrightarrow u_i > \log\left(\frac{P_i^R}{P_i^S}\right) - \log M_i + q. \tag{6}$$

Given the protest interest θ, the probability that a citizen will protest is given by[2]

$$P(\theta) \equiv \Pr\{a_j = 1 | \theta\}. \tag{7}$$

After observing his private interest u_i, citizen i has a different posterior of θ, which is denoted as $f_i(\theta)$ following (1). We are going to solve $P(\theta)$ in an equilibrium shortly. Given $P(\theta)$ and $f_i(\theta)$, the pivotal probabilities and expected size of protesting citizens are given by the following combinatorics:

$$P_i^S = \frac{(n-1)!}{\lceil p_S n \rceil! (n - 1 - \lceil p_S n \rceil)!} \int_0^1 P(\theta)^{\lceil p_S n \rceil} [1 - P(\theta)]^{n-1-\lceil p_S n \rceil} f_i(\theta) \, d\theta, \tag{8}$$

$$P_i^R = \frac{(n-1)!}{\lceil p_R n \rceil! (n - 1 - \lceil p_R n \rceil)!} \int P(\theta)^{\lceil p_R n \rceil} [1 - P(\theta)]^{n-1-\lceil p_R n \rceil} f_i(\theta) \, d\theta, \tag{9}$$

$$M_i \equiv \sum_j G_{ij}^n \Pr_i \{a_j = 1\} = \int P(\theta) f_i(\theta) \, d\theta. \tag{10}$$

[1] Recall that $\lceil x \rceil$ is the smallest integer that is weakly greater than x.

[2] We assume a tie-breaking rule where citizens choose not to protest if they are indifferent.

Counting the combinatorics for the event of $\{A_{-i} = \lceil p_S n \rceil\}$ for P_i^S, there are $C_{\lceil p_S n \rceil}^{n-1}$ combinations among $n-1$ citizens (excluding citizen i himself) where $\lceil p_S n \rceil$ of them protest but $n-1-\lceil p_S n \rceil$ of them do not. Each combination has probability $P(\theta)^{\lceil p_S n \rceil} [1-P(\theta)]^{n-1-\lceil p_S n \rceil}$ for a given θ. The same logic applies to P_i^R. It is important to note that, up to this point, we have not imposed any assumptions about normal distributions. Therefore, the above formulation of the model holds for any distribution of the preference parameters.

III. EQUILIBRIUM

In a perfect Bayesian equilbrium, $P(\theta)$ is the fixed point such that for any θ,

$$P(\theta) = \Pr\{\Delta_j > 0 | \theta\},$$

where Δ_j follows (6) given $P(\theta)$. Consider a threshold equilibrium (to be verified later) in which there exists a marginal citizen such that any citizen j will protest iff their private interest u_j exceeds marginal citizen's u^*:

$$\Delta_j > 0 \Leftrightarrow u_j > u^*. \tag{11}$$

Then the probability that a citizen protests, given the social interest, becomes

$$P(\theta) = \Pr\{u_j > u^* | \theta\} = \Phi\left(\frac{\theta - u^*}{\sigma}\right), \tag{12}$$

where Φ is the cumulative distribution function of the standard normal distribution. By symmetry, $1 - \Phi(x) = \Phi(-x)$.

Lemma 1 *Under a threshold equilibrium, if exists, the log odds ratio of pivotal probabilities of citizen i at the limit is*

$$\lim_{n\to\infty} \log\left(\frac{P_i^R}{P_i^S}\right) = \frac{1+\psi}{\sigma}(z_R - z_S)(\theta_i - u^*) - \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right), \tag{13}$$

where the posterior θ_i is given by (2), $z_S \equiv \Phi^{-1}(p_S)$ and $z_R \equiv \Phi^{-1}(p_R)$. The expected fraction of protesters within citizen i's social group is

$$M_i = \Phi\left(\frac{\theta_i - u^*}{\sigma\sqrt{1 + (1+\psi)^{-1}}}\right). \tag{14}$$

According to Lemma 1, after observing his private interest u_i, citizen i expects that the fraction of protesting citizens in a threshold equilibrium is $M_i = \Phi(\tau_i)$, where

$$\tau_i \equiv \frac{\theta_i - u^*}{\sigma\sqrt{1 + (1+\psi)^{-1}}} = \frac{u_i + \psi\bar{\theta} - (1+\psi)u^*}{\sigma\sqrt{(1+\psi)(2+\psi)}}. \tag{15}$$

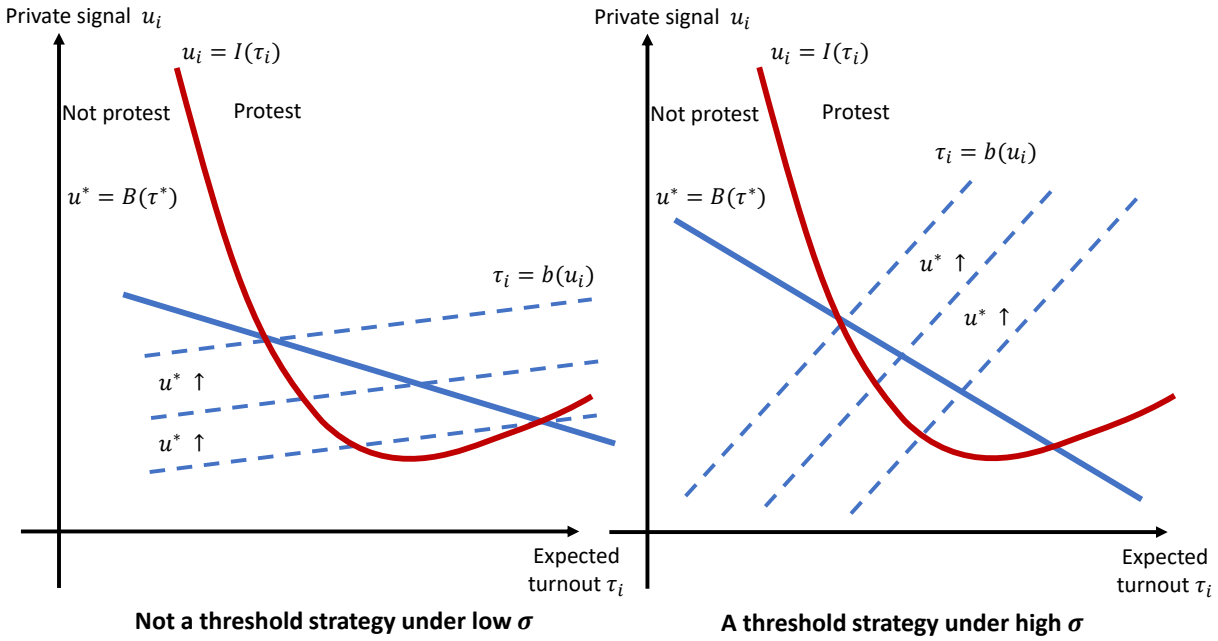


Figure 2: Protest strategy.

We simply refer to τ_i as citizen i's expected turnout, which increases with his private interest in protesting (u_i). This is because a higher private interest raises his posterior belief about θ_i, leading him to expect that his peers have similarly higher private interest in protesting. As a result, more of them are epxected to join the protest. Based on Lemma 1, citizen i chooses to protest in a threshold equilbrium iff (6) is satisfied, i.e.,

$$u_i > \mathcal{I}(\tau_i) \equiv (z_R - z_S)\sqrt{(1+\psi)(2+\psi)}\tau_i - \log\Phi(\tau_i) + \xi, \tag{16}$$

where

$$\xi \equiv q - \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right). \tag{17}$$

The log odds ratio of repression at the limit is $\log\left(P_i^R/P_i^S\right) = (z_R - z_S)\sqrt{(1+\psi)(2+\psi)}\tau_i - (1+\psi)\left(z_R^2 - z_S^2\right)/2$, which is increasing linearly in the expected turnout τ_i and is equal to zero when $\tau_i = (z_R + z_S)/\left[2\sqrt{1+(1+\psi)^{-1}}\right]$. The linearity of the log odds ratio of repression is the result of normal noises, ϵ and ε.

Figure 2 illustrates the individual decision to protest given the marginal interest u^*. $\mathcal{I}(\tau_i)$ represents the indifference curve: citizen i chooses to protest iff his private interest u_i and expected turnout τ_i are realized on the RHS to $\mathcal{I}(\tau_i)$, not to protest on the LHS of $\mathcal{I}(\tau_i)$, and indifferent on $\mathcal{I}(\tau_i)$.

However, not all combinations of u_i and τ_i are rational; citizens update their expected turnout τ_i with respect to u_i according (2). Combining (2) and (15), citizen i's private interest u_i and

expected turnout τ_i are realized on the locus given by

$$\tau_i = b\left(u_i; u^*\right) \equiv \frac{u_i + \psi\bar{\theta} - (1+\psi)\, u^*}{\sigma\sqrt{(1+\psi)\,(2+\psi)}}, \tag{18}$$

It is referred to as the individual expectation condition, illustrated by the dashed lines in Figure 2. According to (16), citizens with u_i and τ_i on the individual expectation curve, and to the RHS (LHS) of the indifference curve will (not) protest. A higher marginal interest u^* or a lower prior $\bar{\theta}$ shifts the individual expectation curve upward, while a higher dispersion of private interest σ steepens the curve.

What remains to be verified is the threshold strategy. It is not obvious because u_i enters both the LHS of (16) (directly as he has higher benefit from protest) and the RHS (indirectly via updating his belief from u_i in $\mathcal{I}\left(\tau_i\right)$). Citizen i would not follow a threshold strategy if the individual expectation curve (18) crosses the indifference curve (16) more than once. Nevertheless, the adoption of the threshold strategy is guaranteed by the following lemma.

Lemma 2 *Under a threshold equilibrium, citizen i follows a threshold strategy iff*

$$z_S \geq z_R - \sigma. \tag{19}$$

The adoption of the threshold strategy requires that the private interest is sufficiently dispersed (i.e. σ is high). If σ is low, the individual expectation curve (18) in Figure (2) becomes very flat such that it intersects the bending portion of the indifference curve (16) more than once. In the rest of this paper (except in Section IV.), we focus on analyzing the threshold equilibrium and assume that (19) is always satisfied.

Having characterized the individual choice to protest given the marginal interest in protest, we are now ready to solve for the threshold equilibria. Obviously, there always exists a trivial threshold equilibrium with $u^* = \infty$, as not protesting is self-fulfilling. We solve for the non-trivial threshold equilibrium as the fixed point of u^*. The threshold u^* in the threhold equilibrium represents the private interest of the marginal citizen who is indifferent between protesting or not.

In the equilibrium, the marginal citizen evaulates his posterior (2) with $u_i = u^*$, his expected turnout $\tau_i = \tau^*$ according to (15) is given by the following *equilibrium expectation condition*

$$u^* = \mathcal{B}\left(\tau^*\right) = \bar{\theta} - \frac{\sigma}{\psi}\sqrt{(1+\psi)\,(2+\psi)}\tau^*. \tag{20}$$

In the threshold equilibrium, the marginal citizen is indifferent between protesting and not protesting, such that (16) holds at equality at τ^* and u^*, which is referred to as the *indifference condition*. A threshold equilibrium is the fixed point of

$$\begin{aligned} u^* &= \mathcal{I}\left(\tau^*\right); \\ u^* &= \mathcal{B}\left(\tau^*\right). \end{aligned}$$

The following proposition completely characterizes the typology of the threshold equilibrium.

Proposition 1 *Suppose (19) is satisfied. Define two conditions:*

$$z_S < z_R + \frac{\sigma}{\psi}, \tag{21}$$

$$q < q_c, \tag{22}$$

where q_c and τ_c are uniquely given by

$$q_c \equiv \bar{\theta} + \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right) - \phi\left(\tau_c\right)\tau_c + \log\Phi\left(\tau_c\right),$$

$$\phi\left(\tau_c\right) = \sqrt{\left(1+\psi\right)\left(2+\psi\right)}\left(z_R - z_S + \frac{\sigma}{\psi}\right),$$

where $\phi(z) \equiv \Phi'(z)/\Phi(z)$.

a. There always exists a trivial threshold equilibrium with $u^* = \infty$.

b. Suppose condition (21) is not satisfied. There exists a unique non-trivial threshold equilibrium.

Suppose condition (21) is satisfied.

c.i. If condition (22) is satisfied, then there exist two non-trivial threshold equilibria.

c.ii. If condition (22) is not satisfied with equality, then there exists a unique non-trivial threshold equilibrium.

c.iii If condition (22) is not satisfied with inequality, then there does not exist any non-trivial threshold equilibrium.

Figure 3 illustrates Proposition 1. The marginal citizen's indifference curve, $u^* = \mathcal{I}(\tau^*)$, is illustrated by the red line in Figure 3. The equilibrium expectation condition, $u^* = \mathcal{B}(\tau^*)$, is illustrated by the blue line in Figure 3. The left panel illustrates the equilibrium of Proposition 1.b. The right panel illustrates the equilibrium of Proposition 1.c.i. The situation of Proposition 1.c.ii is the case when the indifference curve (red line), $u^* = \mathcal{I}(\tau^*)$, does not intersect with the equilibrium expectation condition (blue line), $u^* = \mathcal{B}(\tau^*)$, in the right panel.

III..A. Endogenous Strategic Interaction

Do the equilibria found in Proposition 1 feature strategic complementarity or strategic substitutability? To address this, we first define the concept of strategic interaction. Formally, consider a scenario where every citizen expects turnouts higher than the equilibrium turnout by δ to $\tau^* + \delta$. As a result, the corresponding threshold for protest is lower to $u_\delta = \mathcal{B}(\tau^* + \delta) < u^*$. Expecting that citizens observing $u_j > u_\delta^*$ will protest, citizen i's expected turnout becomes $\tau_i^\delta = b(u_i; u_\delta)$. He will protest iff his private interest satisifies $u_i > \mathcal{I}\left(\tau_i^\delta\right)$, which is equivalent to $u_i > u_\delta^*$, where

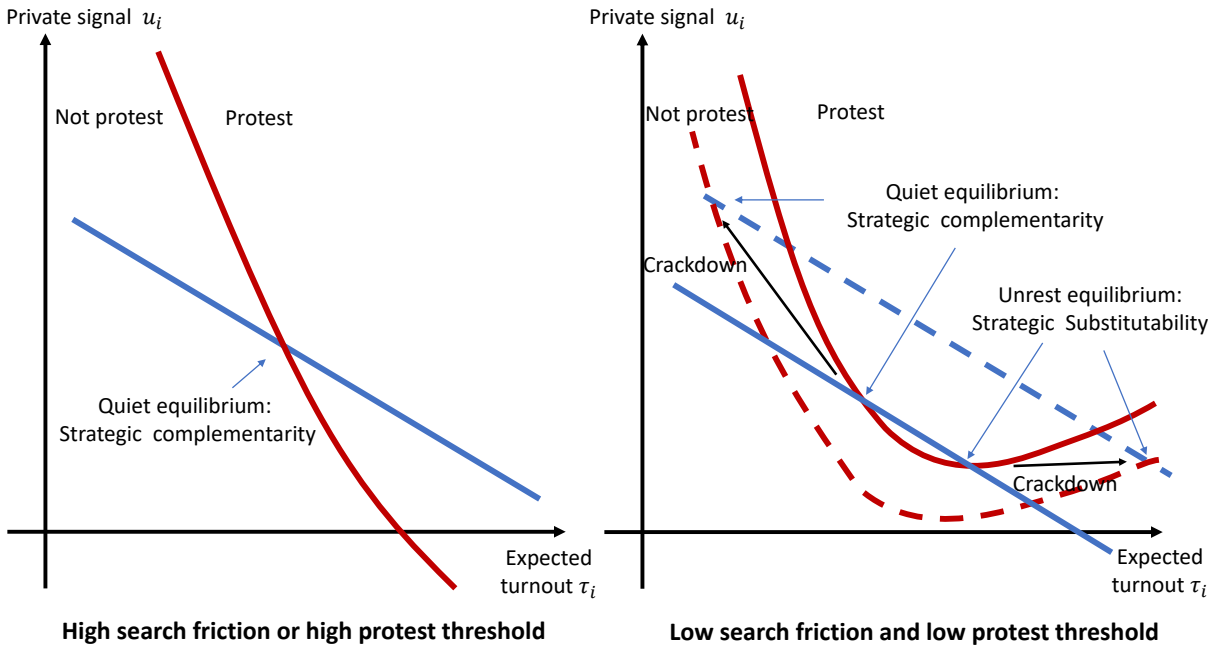


Figure 3: Left: unique equilibrium of stratergic complementarity. Right: multiple equilibria with stratergic substitutability.

u_δ^* solves $u_\delta^* = \mathcal{I}\left[b\left(u_\delta^*; u_\delta\right)\right]$. A threshold equilibrium features strategic complementarity if there is a higher fraction of protesting citizens than expected; strategic substitutability if there is a lower faction; and strategic neutrality if the faction is the same, which occurs when the protest threshold is lower, i.e., $u_\delta^* < u^*$. The above concept is formalized by the following definition.

Definition 1 *For all infinitestimal δ, define the following strategic interaction of given a threshold equilibrium*

$$\textit{strategic complementarity} \qquad \textit{if } u_\delta^* < u^*; \tag{23}$$

$$\textit{strategic substitutability} \qquad \textit{if } u_\delta^* > u^*; \tag{24}$$

$$\textit{strategic neutrality} \qquad \textit{if } u_\delta^* = u^*. \tag{25}$$

The following lemma simplifies the classification of strategic interaction by comparing the slopes of the indifference curve and the equilibrium expectation condition.

Lemma 3 *Strategic interaction of given a threshold equilibrium is equivalent to the following slope comparison:*

$$\textit{strategic complementarity} \quad : \quad \mathcal{B}'\left(\tau^*\right) > \mathcal{I}'\left(\tau^*\right); \tag{26}$$

$$\textit{strategic substitutability} \quad : \quad \mathcal{B}'\left(\tau^*\right) < \mathcal{I}'\left(\tau^*\right); \tag{27}$$

$$\textit{strategic neutrality} \quad : \quad \mathcal{B}'\left(\tau^*\right) = \mathcal{I}'\left(\tau^*\right). \tag{28}$$

Note that in equilibrium every combination of u^* and τ^* must lie on the locus of (20), $u^* = \mathcal{B}(\tau^*)$. When the marginal citizen expects slitghtly higher turnouts from other citizens, the equilibrium marginal interest of protest changes by $\mathcal{B}'(\tau^*)$, and the individual expectation curve shifts downward. Recall that, by following a threshold strategy, citizens strictly prefer to protest if their (u_i, τ_i) lies on the right of the indifference curve, such that $u_i > \mathcal{I}(\tau_i)$. Thus, after expecting higher turnouts from other citizens, more citizens on the individual expectation curve will find themselves in the region of $u_i > \mathcal{I}(\tau_i)$ if $\mathcal{B}'(\tau^*) > \mathcal{I}'(\tau^*)$, i.e. strategic complementarity. Conversely, this is the case of strategic substitutability.

Proposition 2 *a. The threshold equilibrium in Proposition 1b features strategic complementarity.*

b. The threshold equilibrium in Proposition 1ci with a higher expected turnout features strategic substitutability, while the other threshold equilibrium with a lower expected turnout features strategic complementarity.

c. The threshold equilibrium in Proposition 1cii features strategic neutrality.

Proposition 2b offers a novel resolution to Tullock's paradox. In this case, a larger turnout still occurs in the equilibrium featuring strategic substitutability. The intuition behind this is that, in our framework, strategic interaction is endogenous. Specifically, as citizen's turnout is large, the marginal benefit from encountering other protesting citizens in a decentralized mass movement diminishes due to the search-and-matching nature of the process. On the other hand, the marginal cost of protest, in terms of triggering repression, becomes higher. Thus, as turnout grows, the choice to protest become strategic substitution among citizens. In Tullock's paradox, the nature of strategic substitutability is exogeneously given, always leading to low turnouts because protesting citizens crowd each other out. In our model, however, there is a feedback casuality driven by the search-and-matching mechanism.

Proposition 3 *Condition (22) in Proposition 1ci, i.e., the necessary condition for multiple threshold equilibria is more likely to be satisfied when the search friction q is low, the goal of protest p_S is low, the prior interest $\bar{\theta}$ is high, or the information dispersion σ is low.*

The source of multiplicity in our model lies in the search-and-matching nature of the decentralized movement. As illustrated in Figure 3, multiple equilibria arise because the indifference curve $u^* = \mathcal{I}(\tau^*)$ bends backward and intersects the equilibrium expectation condition $u^* = \mathcal{B}(\tau^*)$. It is clearly illsutrated in the equation that the bending is due to the presence of $\log M_i$ in $\mathcal{I}(\tau_i)$. Without this factor, the indifference curve $u^* = \mathcal{I}(\tau^*)$ is monotonic and cuts the equilibrium expectation condition $u^* = \mathcal{B}(\tau^*)$ only once. The effect of search-and-matching is stronger when the search friction q is low. The mechanism by which search-and-matching creates multiplicity is

through strategic uncertainty: citizens are uncertain about the faction of protesting citizens they will encounter (M_i), but their payoff increases with that fraction. As a resut, they can coordinate in two different ways, either with high or low thresholds and turnouts. This effect is more pronounced when citizens are more certain about achieving the protest payoff without triggering repression, which is more likely when the goal of the protest p_S is low, the prior interest $\bar{\theta}$ is high, or the information dispersion σ is low.

III..B. Comparative Statics

To characterize the implications of different strategic interactions, we are interested in deriving the comparative statics across various threshold equilibria in the case of Propositon 1c-i.

Proposition 4 *The comparative statics for the threshold equilibria, featuring strategic substitutability and strategic complementarity, are given by*

		dq	$d\bar{\theta}$
Strategic	$d\tau^*$	$-$	$+$
Substitutability	du^*	$+$	$\pm$
Strategic	$d\tau^*$	$+$	$-$
Complementarity	du^*	$-$	$+$

As illustrated in Figure 3, a higher search friction q shifts the indifference curve $u^* = \mathcal{I}(\tau^*)$ downard, which reduces the expected turnout τ^* and increases the marginal interest u^* in the high-τ equilibrium (which also features strategic substitutability for the reason of Proposition 2) but increases the expected turnout τ^* and decreases the marginal interest u^* in the low-τ equilibrium (strategic complementarity).

A higher prior $\bar{\theta}$ shifts the equilibrium expectation condition $u^* = \mathcal{B}(\tau^*)$ in Figure 3 upward. It reduces the expected turnout τ^* and increases the marginal interest u^* in the low-τ equilibrium (strategic complementarity), but increases the expected turnout τ^* in the high-τ equilibrium (strategic substitutability). The comparative statics on the marginal interest u^* depends on the location of (τ^*, u^*) on the indifference curve. If (τ^*, u^*) is located on the upward (downward) sloping part of the indifference curve, then we have $du^*/d\bar{\theta} > (<)0$, as delineated in Figure 3.

III..C. Implications of Unpopular Crackdowns

Consider an unpopular crackdown that makes it more difficult for citizens to meet each other for a decentralized protest but also raises the prior interest in protest. Specifically, the search friction q increases with the intensity of an unpopular crackdown x, while the prior $\bar{\theta}$ increases even more with x, such that

$$\theta_x > q_x.$$

Proposition 5 *The comparative statics of a unpopular crackdown are given by*

	Own effects		Marginal effects		
		dx		dq	$d\bar{\theta}$
Strategic Substitutability	$d\tau^*$	$+$	$d\frac{d\tau^*}{dx}$	$+$	$-$
Strategic Complementarity	$d\tau^*$	$-$	$d\frac{d\tau^*}{dx}$	$-$	$+$

Proposition 5 states that the consequences of an unpopular crackdown are opposite in equilibria with different strategic interactions, as illustrated in Figure 3. Furthermore, an unpopular crackdown will polarize the two equilibria, as the large turnout in the strategic substitutability equilibrium will become even larger, while the small turnout in the strategic complementaritye equilibrium will become even smaller. This highlighs an important policy implication of distinguishing between these mechanisms, as otherwise opposite predictions could be made. Specifically, an unpopular crackdown actually increases protest turnout, although the marginal effects of the crackdown through reducing the search friction q and higher prior $\bar{\theta}$ are diminishing.

III..D. Strategic Uncertainty at the Limit

In this section we are interested in the equilibrium selection when strategic uncertainty goes to the limit. We ignore the uninteresting case $\sigma = 0$, since (19) and (21) imply $z_R \leq z_S$, hence there is always unique non-trivial threshold.

Proposition 6 *a. If $\sigma \to \infty$, then $\tau^* \to 0$ and $u^* \to q - \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right) + \log 2$. The threshold equilibrium features strategic substitutability.*

b. If $\psi \to \infty$, then $\tau^* \to \frac{z_R+z_S}{2}$ and $u^* \to \bar{\theta} - \frac{\sigma}{2}(z_R + z_S)$. The threshold equilibrium features strategic substitutability if $z_R < z_S$.

c. If $\psi = 0$, then $\tau^* \to 0$ and $u^* \to q - \frac{1}{2}\left(z_R^2 - z_S^2\right) + \log 2$. The threshold equilibrium features strategic substitutability.

When there are multiple threshold equilibrium and strategic uncertainty approaches its limit, the equilibrium featuring strategic substitutability is always selected (except in the unusual case $z_R < z_S$ in Proposition 6b). This result suggests that, theoretically, the equilibrium featuring strategic substitutability is more robust.

IV. INVERTED-THRESHOLD EQUILIBRIUM

So far we have maintained condition (19), ensuring the existence of a threshold equilibrium. To complete our analysis, in this section we solve for the equilibrium when condition (19) is violated.

In this case, the choice to protest follows an inverted threshold strategy, rather than a threshold strategy.

Consider an inverted threshold equilibrium (to be verified later) where a citizen j will protest iff his private interest is sufficiently *lower* than the marginal interest u^*:

$$\Delta_j > 0 \Leftrightarrow u_j < u^*. \tag{29}$$

Then the probability that a citizen protests given social interest becomes

$$P(\theta) = \Pr\{u_j < u^* | \theta\} = \Phi\left(\frac{u^* - \theta}{\sigma}\right). \tag{30}$$

Following the proof of Lemma 1, we can show that under an inverted threshold equilibrium, if exists, the log odds ratio of pivotal probabilities of citizen i at the limit is

$$\lim_{n\to\infty} \log\left(\frac{P_i^R}{P_i^S}\right) = -(z_R - z_S)\sqrt{(1+\psi)(2+\psi)}\tau_i - \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right), \tag{31}$$

where the expected turnout becomes

$$\tau_i = \frac{u^* - \theta_i}{\sigma\sqrt{1 + (1+\psi)^{-1}}}. \tag{32}$$

Citizen i's expected fraction of protesters in his social group is $M_i = \Phi(\tau_i)$.

Citizen i chooses to protest in an inverted threshold equilbrium iff

$$u_i > \tilde{I}(\tau_i) \equiv -(z_R - z_S)\sqrt{(1+\psi)(2+\psi)}\tau_i - \log\Phi(\tau_i) + \xi. \tag{33}$$

The following lemma verifies the the adoption of inverted threshold strategy in this case.

Lemma 4 *Under an inverted threshold equilibrium, citizen i follows an inverted threshold strategy iff (19) is violated.*

The equilibrium expectation condition *becomes*

$$u^* = \tilde{\mathcal{B}}(\tau^*) = \bar{\theta} + \frac{\sigma}{\psi}\sqrt{(1+\psi)(2+\psi)}\tau^*, \tag{34}$$

which becomes a *upward-sloping* curve. Like before, an inverted threshold equilibrium is the fixed point of $u^* = \tilde{I}(\tau^*) = \tilde{\mathcal{B}}(\tau^*)$. The following proposition completely characterizes the typology in the remaining region.

Proposition 7 *Suppose (19) is violated. Define the condition*

$$q < \tilde{q}_c, \tag{35}$$

where $\tilde{q}_c$ *and* $\tilde{\tau}_c$ *are uniquely given by*

$$\tilde{q}_c \equiv \bar{\theta} + \frac{1+\psi}{2}\left(z_R^2 - z_S^2\right) - \phi\left(\tilde{\tau}_c\right)\tilde{\tau}_c + \log\Phi\left(\tilde{\tau}_c\right),$$

$$\phi\left(\tilde{\tau}_c\right) = -\sqrt{\left(1+\psi\right)\left(2+\psi\right)}\left(z_R - z_S + \frac{\sigma}{\psi}\right).$$

a. There always exists a trivial inverted threshold equilibrium with $u^* = -\infty$.

b. Suppose condition (21) is satisfied. There exists a unique non-trivial inverted-threshold equilibrium, which features strategic complementarity.

Suppose condition (21) is not satisfied.

c.i. If condition (35) is satisfied, then there exist two non-trivial inverted-threshold equilibria. The one with a higher expected turnout features strategic substitutability; the other threshold equilibrium with a lower expected turnout features strategic complementarity.

c.ii. If condition (35) is not satisfied with equality, then there exists a unique non-trivial inverted-threshold equilibrium, which features strategic neutrality.

c.iii If condition (35) is not satisfied with inequality, then there does not exist any non-trivial inverted-threshold equilibrium.

V. EXTENSION WITH VOTING

In most situations, individual-level protest participation data is unavailable (with notable exceptions such as field experiments), but voting data is typically more accessible at the constituency level. In this section, we extend the model to incorporate the decision of voting against the incumbent, allowing the model to generate testable implications using voting data.

Voting. Now, a typical citizen must decide not only whether to vote against the incumbent but also whether to protest against him. The payoff of citizen i is

$$\mathcal{U}_i\left(a_i, b_i\right) = \mathbb{E}_i\left\{\begin{array}{l}\Pr_i\left\{\frac{B}{n} \leq p_W\right\} W_i + \Pr_i\left\{\frac{B}{n} > p_W\right\} L_i \\ +e^{-q}\Pr_i\left\{\frac{A}{n} > p_S\right\} M_i U_i - \Pr_i\left\{\frac{A}{n} > p_R\right\} C_i\end{array}\right\},$$

where citizen i votes against (for) the incumbent if $b_i = 1$ ($b_i = 0$). The incumbent wins the election if the number of protest votes $B \equiv \sum b_i$ is below the threshold $p_W \in (0,1)$, i.e., $B/n \leq p_W$. If the incumbent wins (loses), then citizen i enjoys the payoff W_i (L_i).

Assumptions. In this paper, we focus on the scenario where $W_i > L_i$, i.e., the incumbent's victory is beneficial to all citizens and should be considered optimal in the first-best outcome. However, preferences regarding protests differ between two types of citizens. A *conservative* citizen j prefers avoiding protests, i.e., $U_j < 0$. Therefore, a conservative citizen always chooses $a_j = b_j = 0$. On the other hand, an *activist* citizen i favors engaging in protests, i.e., $U_i > 0$. However, activist citizens, acting with consistent principles, can not simultaneously choose $a_i = 1$ and $b_i = 0$, i.e., they cannot protest on the street but vote for the incumbent. This behaviorial assumption

simiplifies the dimensionality of the action space that needs to be considered when solving for the threshold equilibrium. To further reduce the dimensionality of heterogeneity, we assume that the cost of repression is proportional to the differential benefit of electing the incumbent, i.e., $C_i = k^{-1}(W_i - L_i)$ for some constant $k > 0$. In the rest of this paper, we focus on the decisions of activist citizens. Denote π as the share of activist citizens. We assume that if all activists vote and protest against the incumbent, the protest will succeed and incumbent will lose, i.e.,

$$\max\{p_W, p_S\} < \pi. \tag{36}$$

Incorporating the choice of voting, modify citizen i's gain from protest (both in action and in the ballot) as

$$\Delta_i \equiv U_i(1,1) - U_i(0,0) = -P_i^W(W_i - L_i) - P_i^R C_i + e^{-q}\Pr_i\{a_j = 1\} P_i^S U_i, \tag{37}$$

where

$$P_i^W \equiv \Pr_i\{B_{-i} \leq \lfloor p_W n \rfloor\} - \Pr_i\{B_{-i} + 1 \leq \lfloor p_W n \rfloor\} = \Pr_i\{B_{-i} = \lfloor p_W n \rfloor\}. \tag{38}$$

We refer to P_i^W as citizen i's *pivotal* probability of electing the incumbent. In particular, P_i^W is the probability that citizen i is the marginal citizen whose vote will be critical to electing the incumbent. Arranging the terms, it is strictly optimal for citizen i to protest after observing u_i iff

$$\Delta_i > 0 \Leftrightarrow u_i > \log\left(\frac{kP_i^W + P_i^R}{P_i^S}\right) - \log \mathbb{E}_i P(\theta) + q. \tag{39}$$

$$\begin{aligned} u^* = \mathcal{I}(\tau^*) &= \log\{\exp[(1+\psi)(z_R - z_S)\tau^*] + \zeta_W \exp[(1+\psi)(z_W - z_S)\tau^*]\} && (40) \\ &\quad - \log\Phi\left(\frac{\tau^*}{\sqrt{1 + (1+\psi)^{-1}}}\right) + \xi && (41) \end{aligned}$$

Lemma 5 *Under a threshold equilibrium in the extended model, citizen i follows a threshold strategy iff*

$$z_S > \max\{z_R, z_W\} - \sigma. \tag{42}$$

Proposition 8 *In the extended model, suppose the following conditions are satisfied:*

$$z_S < \min\{z_R, z_W\} + \frac{\sigma}{\psi}, \tag{43}$$

$$q < q_c, \tag{44}$$

where q_c is given in Appendix A5. There exist two non-trivial threshold equilibria. Proposition 2b, Proposition 4 and Proposition 5 hold in the extended model.

VI. HYPOTHESES

Our theoretical results provide a number of testable implications. The equlibria of our model exhibit opposite mechanisms. Formally, to test the equilibrium strategic interaction predicted by the model, we propose the following hypotheses according to Proposition 4.

Hypothesis 1a A society features *strategic substitutability* only if the number of turnout *increases* when citizens are

i. easier to connect with each other;

ii. more sympathetic to protests.

Hypothesis 1b A society features *strategic complementarity* only if the number of turnout *decreases* when citizens are

i. easier to connect with one another;

ii. more sympathetic to protests.

To address multiple equilibria in our model, we propose two procedures. One procedure is to assume that a particular equilibrium is always selected for the entire sample (although we do not know which one) and test Hypotheses 1a and 1b directly. Of course, if our model is correct but the data includes a mixture of two equilibria, insiginificant results may arise, as the comparative statics summarized in Proposition 4 differ between the two equilibria.

A more robust, albeit less powerful, procedure is to rely on quantile analysis, in the spirit of Echenique and Komunjer (2009), which we will describe in the next section. Intuitively, if we observe a larger protest, it is more likely that an equilibrium of strategic substitutability is in play, allowing us to test Hypothesis 1a with higher quantiles and vice versa for testing Hypothesis 1b. The model is rejected by the data if both Hypotheses 1a and 1b are rejected by either procedure. In this case, the data exhibits neither mechanisms of strategic substitutability nor of strategic complementarity as predicted by the model.

If either Hypotheses 1a or 1b is not rejected, we can further test the corresponding policy implication predicted by the model. For example, suppose the goverment adopts a policy involving unpopular crackdowns. The model predicts that such a crackdown will polarize the two equilibria. To formally test the policy implications predicted by the model, we propose the following hypotheses according to Proposition 5.

Hypothesis 2a Given a society featuring *strategic substitutability*, a unpoplar crackdown *increases* the number of turnout.

Hypothesis 2b Given a society featuring *strategic complementarity*, a unpoplar crackdown *decreases* the number of turnout.

If either Hypotheses 1a and 2a (for strategic substitutability), or Hypotheses 1b and 2b (for strategic complementarity) is not rejected, then we can finally test the corresponding marginal effects of the policy predicted by the model. Formally, we propose the following hypotheses according to second part of Proposition 5.

Hypothesis 3a Given a society featuring *strategic substitutability*, the effect of a unpoplar crackdown on the number of turnout is *diminishing* when citizens are

i. easier to connect with one another;

ii. more sympathetic to protests.

Hypothesis 3b Given a society featuring *strategic complementarity*, the effect of a unpoplar crackdown on the number of turnout is *amplifying* when citizens are

i. easier to connect with one another;

ii. more sympathetic to protests.

These implications are tested in our empirical analysis performed in the next section.

VII. DATA

The 2019 District Council election in Hong Kong provides a unique opportunity to study the effect of crackdown on collective actions. On February 13, 2019, the Hong Kong government introduced an extradition bill that sought to repeal the existing ban on surrendering fugitives to Mainland China.[3] Since June 2019, large-scale rallies and decentralized protests have proliferated. The police adopted a confrontational approach, including the use of force during the movement. Protests and repression remained active in the leap-up to the District Council election held in November 2019. [4] As election day approached, candidates were frequently asked to state their views on the protests, which influenced many people's voting decisions. This made the election outcomes a valuable test for our theory, as it helps rule out potential confounding factors that could explain the election outcomes.

[3] There has been no extradition agreement between Mainland China and Hong Kong. The extradition ban was established by the British colonial government before the 1997 handover, driven by concerns "about the consequences of an agreement which results in defendants being transferred from the transparent, impartial, legal system found in Hong Kong to the mainland's considerably less transparent system," according to the Tenth Report by the House of Commons Foreign Affairs Committee of the United Kingdom.

[4] All elected members serve a four-year term. District councils hold budgetary power over local projects within their constituencies. They manage public facilities, such as libraries and markets, and provide advice to the government on local policies, including transport and zoning.

VII..A. Tear Gas Deployment

To test our hypotheses, we hand-collected data on tear gas deployment across 452 constituencies (nested within 18 districts) in Hong Kong, from June 1, 2019 (the start of large scale demonstrations) to election day (November 23, 2019). We verified each incident by cross-checking over 22 news sources and media websites, as listed in Appendix **??**. The location of each incident was identified, sometimes using information from videos and photos. We avoided using existing online databases, which rely heavily on self-reports and are prone to significant Type I and Type II errors.[5] For example, we found that a popular online source missed 20% of the tear-gassed constituencies in our dataset, and about 21% of the tear-gassed constituencies in the onlien source could not be verified.[6]

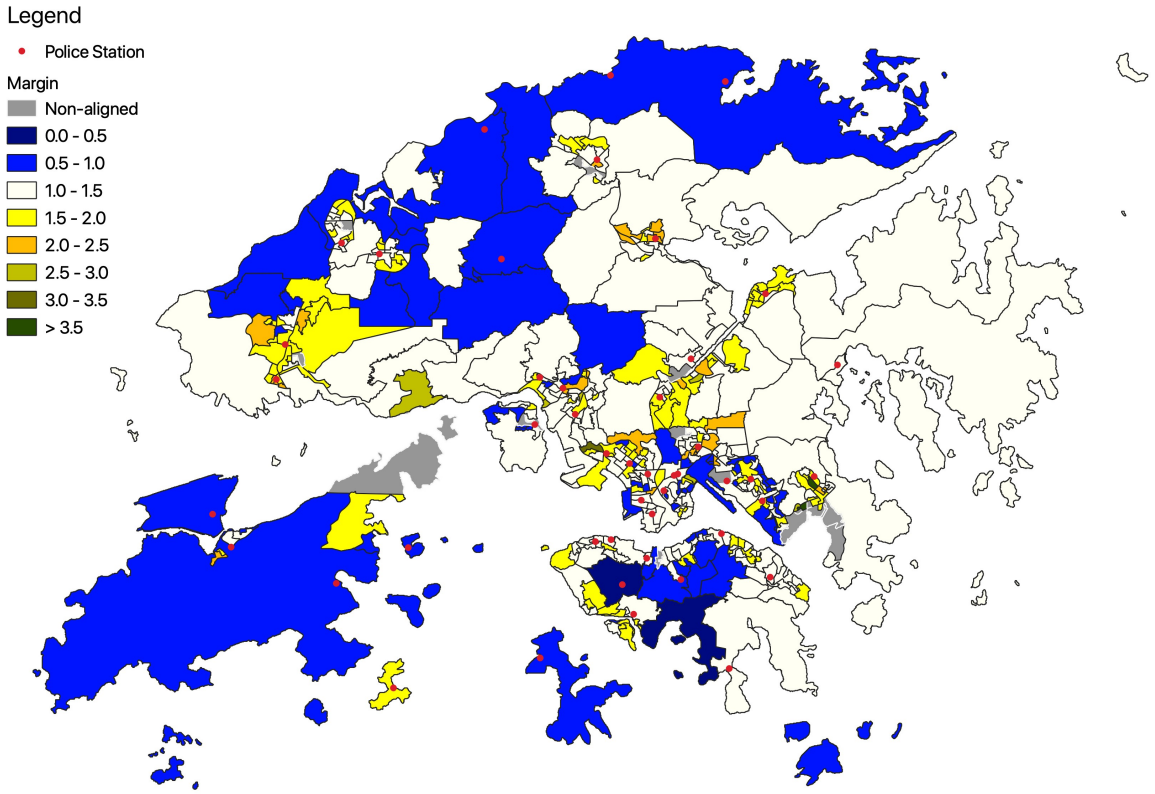
According to our data set, police deployed tear gas in over 25% of the constituencies. Rarely used in the history of Hong Kong, more than 15,000 rounds of tear gas were fired between June and November, right before the election. In contrast, fewer than 100 rounds were used during the Umbrella Movement in 2014.

The use of tear gas is linked to crackdowns in our model. [7] While our data set records tear gas deployment, other invasive crowd control weapons, such as rubber rounds, super socks and 40mm react rounds, were used less frequently. When deployed, they were often used in combination with tear gas. From June to November 2019, the correlation coefficients between tear gas and these crowd control weapons were 97%, 95% and 93% respectively, according to data released by the Hong Kong Police Force. Tear gas deployment is easier to identified, even for those far from the protest sites, as it is visible and leaves a long-lasting smell. It is also easily recognized and widely reported in the the media during the movement. Since tear gas had rarely deployed before the 2019 movement in Hong Kong (with the exception of the Umbrella Movement), media reports are expected to capture geographical variations sufficiently and accurately.

[5] We did use information from groups on platofrms like Telegram to cross-validate some events. However, unless an incident was reported by at least one other independent source, and accompanied by a photo or a video that allows us to identify the location using Google Maps and Google Street View, the event is excluded.

[6] For example, HKMap.live, a popular online mapping of tear gas deployment and protests in Hong Kong, missed 26 out of 130 treated constituencies. Self-reporting tends to amplify Type I errors. According to our dataset, 28 constituencies contained false positives. For example, tear gas was reportedly smelled in a constituency in Tuen Mun but none was fired at the reported location. Sometimes, location names were abbreviated in reporets shared on information groups. For instance, Tuen Mun Tai Hing could refer to Tai Hing Estate, Tai Hing Operations Base, Tai Hing Sports Center, all of which are located in different constituencies. While we found that tear gas was reported to have been fired in only one of these constituencies, other locations were marked in some online databases. Finally, some reports confused tear gas warnings (black flag) with actual firings.

[7] Although the name “tear gas" may sound harmless, its hazardous effects on human health are well documented. Tear gas has been used in high-density residential areas and enclosed space, such as subways and shopping malls, where it exposes people to high concentrations of irritants for prolonged periods. During the Arab Spring, police in Cairo fired tear gas near a van, leading to 37 deaths, according to Reuters. Scientific research reports deaths and permanent disabilities, such as blindness, traumatic brain injury resulting in vegetative states, amputations, loss of limb function, and chronic asthma (Chan et al. 2019). Moreover, its genetic consequences, carcinogenicity, and environmental impacts remain largely unknown. Given its health impacts, the Chemical Weapons Convention explicitly prohibits its use in warfare.



Notes. Margin is defined as the ratio of votes for pro-protest candidates to votes for the pro- government candidates.

Figure 4: Voting Pattern in the 2019 Hong Kong District Council Election.

VII..B. Votes for the Pro-protest Candidates

The variable is defined as the log difference between the votes for pro-protest candidates and those for pro-government candidates. Figure 4 displays the election results of the election. Political alignments are determined based on the party affiliations of the candidates, who may be pro-protest, pro-government or independent. For candidates who identify themselves as independent, their alignments are classified based on their speeches and/or manifestoes.[8] No constituency was uncontested in the election. For candidates without a clear alighment, votes for them are categorized as non-aligned. As a result, when only one political camp has candidates in a constituency, the data for that constituency is treated as missing value. Election results and the number of electors are based on administrative data provided by the Registration and Electoral Office.

[8] Weeks before the election, civil society organizations and media outlets created websites that listed the political alignments of candidates based on past speeches and social media posts. Many voters used these sites as a guide to identify the political alignment of a candidate. We incorporate this information when coding the variable.

VII..C. Social Media

Our model predicts that the effects of repression depend on people's awareness of others' opinions regarding the event. The Chinese University of Hong Kong conducted a series of surveys at protest sites (CCPOS, 2020), asking protesters how they obtained and shared information about the events. Approximately 80% to 90% of respondents reported using Facebook as a source of information.[9] To measure how easily citizens can access information about the protests and learn about others' views, we use the estimated number of Facebook and Instagram users in each constituency. Based on data from January 2019 data from Facebook, Hootsuite estimated the percentage of Facebook and Instagram users in Hong Kong by their gender and age groups. We then estimated the (logged) number of Facebook and Instagramer users in each constituency using these estimates along with demographic statistics from the Census and Statistics Deparment of Hong Kong.

VII..D. Other Data

To control for heterogeneities, we include a number of control variables from various sources in our cross-sectional data set.[10] These control variables include: (1) the (logged) number of college graduates in a constituency, (2) a binary variable indicating whether a constituency is pro-government, which takes the value of 1 if a pro-government candidate (based on their political party affiliations) has won in all four previous district council elections, and 0 otherwise, (3) a similar binary variable for pro-democracy constituencies, with the same coding rule applied; and (4) the income group that a constituency falls under (high, middle, or low), based on the tertile of median income in the constituency. According to a public opinion survey (CCPOS 2020), individuals with higher levels of education and those in higher income groups were more sympathetic towards the protesters. All data were obtained from the Census and Statistics Deparment of Hong Kong and the Registration and Electoral Office websites. We use data from the 2016 population by-census to code variables describing constituency characteristics.[11] Table 1 reports the summary statistics of the variables used in our econometric analysis.

[9] Another popular platform, LIHKG, became increasingly important as the event progressed. The disucssion forum was particularly popular among younger people, who were the main supporters of the movement.

[10] Ideally, if elections were held frequently and coincided with tear gas deployment, a difference-in-differences approach could be a viable option. However, the previous district council election was held in 2015, and given the rapidly changing political landscape in Hong Kong, the parallel trend assumption is likely violated, making difference-in-differences approach unsuitable.

[11] There were boundary changes between 2016 and 2019. Since most of the changes involve dividing more populated constituencies and recombining them with part of nearby smaller constituencies, we use the values of the preceding constituencies as the estimates for the newly established constituencies. As demographic variables such as median income tend to be stable in the short run, we do not expect significant distortions from this treatment.

	Pro-protest candidates won		Tear gas was deployed		
	Yes	No	Yes	No	All
Number of electors	9372.04	7705.19	8690.19	9309.66	9139.72
Number of pro-protest votes	3907.98	2461.21	3563.30	3760.10	3706.33
Number of pro-government votes	2718.60	2791.83	2600.95	2775.89	2728.65
Median income (US dollars)	27226.89	30437.36	27546.15	27704.05	27660.83
Average age	43.94	43.81	43.50	44.09	43.93
Number of colleage graduates	3400.49	3231.84	3683.65	3253.30	3371.10
Number of social media users	1.46	1.45	1.44	1.47	1.46
Observations	389	63	124	328	452

Notes. Each observation is a district council constituency. Each column is the mean in each subgroup.
Source: The incidences of tear gas comes from authors' collection from 27 news sources listed in the Appendix. The distributions of votes and electors are obtained by the Registration and Electoral Office. The demographic information in each constituency comes from the Census and Statistics Department. The data of local broadband speed comes from speedtest.net.

Table 1: Summary statistics

VIII. IDENTIFICATION STRATEGY

The outcomes in the theoretical model are operationalized in two ways. First, the turnout in Hypotheses 1-3 is captured by $log\theta^*(g_j)$, which is measured as the difference between the logarithms of the pro-protest votes and the pro-government votes in constituency j. The model abstracts from the counter-collective action by citizens, and we normalize the pro-protest votes by the pro-government votes. However, not every constituency has a pro-government candidates standing for election. Therefore, alternatively, we measure $I_j \equiv 1_{A(g_j)p_j \geq \theta_j}$ as an indicator of whether a pro-protest candidate won the election in constituency j. Although the two cases are isomorphic in the model, in practice they are not equivalent: pro-protest candidates can receive more votes than their pro-government competitors but still lose the election due to votes being diverted to non-affilated candidates. Setting up these two alternatives provide a more robust test of our theory using the available data.

The specification is given by the following potential-outcome framework:

$$\log \theta^*(g_j) = 1\{g_j = H\} y_{H,j} + 1\{g_j = L\} y_{L,j}, \text{ where} \tag{45}$$

$$\begin{aligned} y_{H,j} &= \tau + control_j \times \beta_H + \varepsilon_{H,j}, \\ y_{L,j} &= control_j \times \beta_L + \varepsilon_{L,j}, \\ I_j &= \begin{cases} 1, \text{ if } I_j^* \geq 0, \\ 0, \text{ if } I_j^* < 0, \end{cases} \end{aligned}$$

where $1\{g_j = H\} = 1$ if tear gas was deployed in unit j and $control_j$ consists of the logarithms of the number of college graduates and the number of social media users, the income level of unit j,

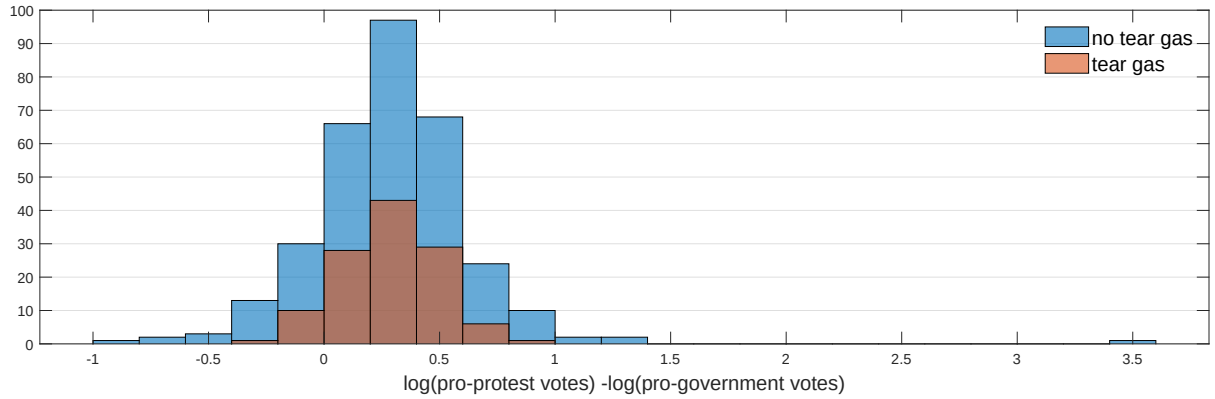


Figure 5: Conditional distributions of pro-protest votes.

historical political allignments, and district fixed effects. The average treatment effect of crackdown is thus given by $\mathbb{E}_j\left(y_{H,j} - y_{L,j}\right)$.

VIII..A. Instrument

Tear gas deployment is likely to be endogenous. For instance, a constituency with a higher proportion of pro-government citizens is less likely to elect pro-protest candidates, but it may be more prone to violent clashes with protesters. In such casese, ε_j and $1\{g_j = H\}$ are correlated, leading to biased OLS estimates that underestimate the true positive treatment effect. This potential endogeneity may explain why the distributions of pro-protest votes between the treatment and control constitutencies indistinguishable, as shown in Figure 5.

To address this concern, we instrument tear gas deployment using an exogeneous variable: the driving distance between the population center of a constituency and its nearest police station. Both of these are pre-determined before the protests and are unrelated to the political alignment of the constituency. [12] The distance were calculated using Google Maps.

A valid instrument must be both relevant and exogeneous. Our instrument is relevant because it captures variations in the endogenous variable. To maintain public order, police stations are typically located near regional population centers. Given that many protests were organized in population centers to maximize participation, the instrument is correlated with tear gas treatment. During the protest period, it was not uncommon for people to demonstrate outside politice statations to express their frustration with the police. (This occurred at about 15 out of 46 police stations, according to our data set.) In most cases, the police dispersed the crowds with tear gas. These factors make our instrument sensitive to the treatment status. The closer a police station is to a population center, the more likely that a constituency was treated. A weak but relevant

[12] There are a few outliners at the right tail, mostly in the rural districts. The distance measure is also subject to a non-neligible measurement errors due to potential misidentifications of the population center, traffic direction, and road conditions in urban areas.

instrument can lead to biased estimates and misleading test results. As we will show later, the first-stage F-statistics are sufficiently large to avoid the weak instrument problem.

Our instrument also satisfies the exogeneity condition. The location of police stations is unlikley to be correlated with voters' political preferences during the 2019 election. Many of the police stations in use were established up by the British colonial government before 1997, making them unlikely to be linked to political alignment of the population. The location of these stations was influenced by various factors. Some police stations, still in operation today, were built before the Second World War (WWII) to combat piracy, triads and other criminal activities. These police stations were built along the coast, near local settlements, or on higher ground for defensive purposes (Hong Kong Police Force n.d.). Other police stations are located in new towns, which were built to accommodate Hong Kong's booming population. To maintain public order, police stations were generally established in densely populated areas. As the city expanded and evolved, new police stations were built primarily for administrative and economic efficiency, land availability, and urban planning. Givan that the location of police stations is unlikely to be correlated with political alignment after the protests, and that people with certain polical preferences are unlikely to consciously choose to live closer to or farther from a police station, our instrument generates the necessary exogenous variation to explain tear gas deployment and can be used to estimate the impact of crackdowns on voter turnout.

IX. RESULTS

In this section we present our findings from testing the hypotheses derived in Section VI.. The Centre for Communication and Public Opinion Survey (CCPOS) at the Chinese University of Hong Kong conducted six waves of surveys on public opinions regarding the anti-extradition bill movement in Hong Kong (CCPOS 2020). The survey results indicate that citizens with higher levels of education and income tend to be more supportive of the protest. As a result, in addition to testing Hypotheses 1 to 3, we also expect constituencies with higher average income levels and more educated residents to have a greater number of citizens sympathetic to the protests. Finally, following research on the impact of communication technologies on protest participation (e.g. Enikolopov et al. 2020), we hypothesize that constituencies with mroe social media users will experience lower search friction, making it easier for citizens to access information about the protests.

IX..A. Testing Hypothesis 1

Table 2 reports the estimates from OLS and three quantile regressions. In constituencies where the pro-protest margin is higher (e.g. above the 90th percent quantile), the level of education significantly predicts voting patterns and in the expected direction. A similar pattern is observed for

our proxy of search frictions. Areas with more social media users saw more votes for pro-protest candidates. Additionally, we found that constituencies with a larger population from the democracy camp experienced a greater margin between the pro-protest and pro-government candidates. Overall, we find significant evidence supporting the equilibrium featuring strategic substitutability. Consequently, Hypothesis 1b is rejected in favor of Hypothesis 1a.

Table 2: EFFECTS OF SEARCH FRICTION AND OTHER COVARIATES

	Net log pro-protest votes			
	mean	50th	75th	90th
Not protest-sympathetic				
Middle income constituency	0.017	0.016	-0.056	-0.012
	(0.037)	(0.030)	(0.046)	(0.029)
High income constituency	-0.038	-0.023	-0.060	-0.048
	(0.060)	(0.047)	(0.065)	(0.057)
Solidly pro-gov't	-0.137****	-0.104****	-0.127****	-0.155****
	(0.027)	(0.022)	(0.030)	(0.027)
Protest-sympathetic				
Solidly pro-democracy	0.234****	0.152****	0.216***	0.190*
	(0.068)	(0.042)	(0.072)	(0.111)
Education	0.055	0.034	0.071	0.080*
	(0.048)	(0.035)	(0.049)	(0.045)
Search frictions				
Social media users	0.138	0.074	0.068	0.211**
	(0.099)	(0.073)	(0.102)	(0.101)
District FE	✓	✓	✓	✓
(Pseudo) R^2	0.288	0.153	0.144	0.200
N	437	437	437	437

Notes: The first column reports the OLS results. The second to the last columns report quantile regression results at the 50th, 75th and 90th quantiles. District fixed effects for 18 administrative districts are included in all models. Robust standard errors are reported in parentheses. $*$: p $<$ 0:10, $**$: p $<$ 0:05, $***$: p $<$0.01, $****$: p $<$ 0:001.

IX..B. Testing Hypotheses 2 and 3

Table 3 reports the findings under various model specifications. We can reject Hypothesis 1b, which posits an equilibrium characterized by strategic complementarity, and find significant results that support the equilibrium characterized by strategic substitutability.

Table 3: AVERAGE TREATMENT EFFECTS OF TEAR GAS DEPLOYMENT

	Net log pro-protest votes		
	IV	IV homo	OLS
Average Treatment Effect	0.411**	0.249***	0.027
	(0.168)	(0.076)	(0.029)
Marginal treatment effect of tear gas			
Not protest-sympathetic			
Middel income constituency	-0.040		-0.028
	(0.082)		(0.077)
High income constituency	-0.171		-0.148
	(0.114)		(0.113)
Solidly pro-gov't	-0.107		-0.120*
	(0.067)		(0.065)
Protest-sympathetic			
Solidly pro-democracy	0.176*		0.172
	(0.103)		(0.111)
Education	0.054		0.026
	(0.099)		(0.094)
Search frictions			
Social media users	0.121*		0.068
	(0.068)		(0.059)
Effective F statistic of IV	31.98	31.98	-
District FE	✓	✓	✓
N	437	437	437

Notes: Testing the average treatment effects and the marginal treatment effect on the net protest votes. The instrument is time distance between the population center of constituency and its nearest police station. Robust standard errors clustered at district level are reported in parentheses. $*$: p $<$ 0:10, $**$: p $<$ 0:05, $***$: p $<$0.01, $****$: p $<$ 0:001.

Table 3 reports the findings under different specifications. Columns 3 presents the plain vanilla regressions without the instrument. The treatment effect of repression is either not statistically sigificant (p-value = 0.338) or economically insignificant (0.027). The weak treatment effect may be due to confounders, as illustrated in Figure 5 .

Columns 1 and 2 of Table 3 present our key results based on instrumental variable estimation. To reject the null of a weak instrument, a standard F-statistic should be at least 10. However, a weak instrument problem may arise when the endogeneous variable (tear gas deployment) is binary and the excluded exogenous variable (distance from the nearest police station) is continuous. Following

Xu (2021), we also compute the effective F-statistic. The critical value for a 5% test with a 10% worst-case bias is 23.1, and the effective F-statistic for our instrument is 31.98, indicating that the instrument is not weak. Using the location of police station as an instrument, we find that repression has a statistically and economically significant positive effect on support for pro-protest candidates. Specifically, after tear gas deployment, the gap between pro-protest and pro-government votes widens by 0.36 log points on average, implying that pro-protest candidates are 23% more likely to win. This result is consistent with Hypothesis 2a where an unpopular crackdown polarizes the society and increases protest turnout in an equilibrium featuring strategic substitutability.

The marginal treatment effects, reported in the remaining rows of Columns 1 and 2, align with Hypothesis 3a, which posits that the effect of an unpopular crackdown diminishes in an equilibrium with strategic substitutability. In summary, our findings provide empirical support for the model's implications and strong evidence of an equilibrium characterized by strategic substitutability.

IX..C. Robustness Checks

IX..C..1 Placebo Test

We also conduct a placebo test to assess the robustness of our findings. This test examines whether tear gas deployment predicts election outcomes in the 2015 District Council Election. Since tear gas was not deployed prior to the 2015 election, thre treatment variable should have no predictive power on the election outcomes for this year. If it does, this would suggest that tear gas deployment is non-random and endogenous, potentially driven by a confounder that simultaneously influences tear gas locations and election outcomes.

To perform this test, we re-estimate the models specified in Table 3, using the election result from 2015 to code the dependent variable. The results of the placebo test are reported in Table 4. Across all specifications, the treatment variable does not predict election outcomes in 2015. These findings imply that the voting patterns observed in 2019 are driven by the use of tear gas but not pre-existing factors.

IX..C..2 Binary Outcome

As explained earlier, candidates from a particular political camp may secure more votes than they otherwise would but still lose the election if a significant portion of votes are diverted to independent candidates. This is especially relevant for inexperienced candidates running for office. Their victories may often be interprted as protest votes against the pro-government regime. Therefore, instead of focusing solely on vote counts, we also examine whether inexperienced, pro-democracy candidates are more likely to win constituencies where tear gas was deployed. Table 5 reports these findings. Similar to the linear model results, while a naive probit estimate suggests that tear gas

Table 4: PLACEBO TESTS

	Net log pro-democracy votes		
	IV	IV-homo	OLS
Average Treatment Effect	0.021	0.247	0.045
	(0.353)	(0.231)	(0.051)
Control variables	✓	✓	✓
District FE	✓	✓	✓
N	300	300	300

Notes: The dependent variable is net log pro-democracy votes in the 2015 District Council Election. Estimations are based on instrumental variable regressions. Control variables include income group, education, whether a constituency were solidly pro-government or pro-democracy, and numbers of social media users. Robust standard errors clustered at district level are reported in parentheses. $*$: p $<$ 0:10, $**$: p $<$ 0:05, $***$: p $<$0.01, $****$: p $<$ 0:001.

has no significant impacts, our IV estimates reveal a positive and significant effect after accounting for potential endogeneity. Constituencies where tear gas was deployed see an approximately 40 percentage point increase in the likelihood of pro-democracy winning. This substantial marginal effect is partly attributable to the overwhelming success of pro-democracy candidates, who secured over 85 per cent of the seats in the 2019 election.

Table 5: CONTROL FOR CANDIDATES' EXPERIENCE

	Inexperienced candidates win	
	IV	Probit
Average Treatment Effect	0.414***	0.123
	(0.070)	(0.134)
Control variables	✓	✓
District FE	✓	✓
N	452	452

Notes: The dependent variable is a binary variable taking value 1 if an experienced candidate win in the 2019 District Council Election. Estimations are based on instrumental variable regressions. Control variables include income group, education, whether a constituency were solidly pro-government or pro-democracy, and numbers of social media users. Robust standard errors clustered at district level are reported in parentheses. $*$: p $<$ 0:10, $**$: p $<$ 0:05, $***$: p $<$0.01, $****$: p $<$ 0:001.

X. OTHER EVIDENCE

The previous sections demonstrates that tear gas deployment increases the wining margins of pro-protest candidates. In this section, we present additional evidence from other sources to substantiate our model assumptions and supplement our econometric results.

X..A. Social Networks and Payoff

Our payoff function addresses the importance of social factors in influencing individuals' participation decisions. In our model, payoff from participating increases when individuals' expected effective fraction of protesters M increases, and M_i is a function of social networks G_{ij}^n. According to a survey conducted by the Chinese University of Hong Kong at the protest sites in June 2019 (CCPOS 2020), many participants used private communication softwares such as WhatsApp and Telegram to distribute information about, and probably their views on the protests.[13] If only few people use them, one may question the behavioral assumption behind the payoff function. Figure **??** shows that WhatsApp was commonly used among protesters to distribute information about the protests. About 60% of respondents indicated that they used it quite often or very often. Telegram was used less often, only about 35% people used that quiet often or very often. This is possible because Telegram was less popular than WhatsApp back then. And Telegram was preferred by more radical protests who used the self-destruct function of Telegram for their actions.This partly explains why about 10% of respondents refused to answer questions related to their use of Telegram during the survey.

[13] The survey was commissioned by the Independent Police Complaints Council (IPCC), a civilian body of the Hong Kong Government and adminstered by the Centre for Communication and Public Opinion Survey (CCPOS) at the Chinese University of Hong Kong. The cited report was based on data collected from a series of surveys conducted at protest sites between June 2019 and December 2019. These surveys covered six major demonstrations led by the Civil Human Rights Front (CHRF) and 10 other larger gatherings during the period.