A Model of Protests, Revolution, and Information

Salvador Barberà
Matthew O. Jackson

This version: October 2019

(January 2017)

Barcelona GSE Working Paper Series

Working Paper nº 951

A Model of Protests, Revolution, and Information

Salvador Barbera and Matthew O. Jackson *

Draft: October 2019

Abstract

A collective action or revolt succeeds only if sufficiently many people participate. We study how potential revolutionaries' ability to coordinate is affected by what they learn from different sources. We first examine how people learn about the likelihood of a revolution's success by talking to those around themselves, which can either work in favor or against the success of an uprising, depending on the prior beliefs of the agents, the homogeneity of preferences in the population, and the number of contacts. We extend the analysis by examining the effects of homophily on learning: people are more likely to meet others who have similar preferences, undercutting learning. We introduce variants of our model to discuss other ways of learning about the support for a revolution. We discuss why holding mass protests before a revolt provides more informative signals of peoples willingness to actively participate than other less costly forms of communication (e.g., via social media). We also show how outcomes of revolutions in one region can inform citizens of another region and thus trigger (or discourage) neighboring revolutions. We also discuss the role of governments in avoiding revolutions and learning about their citizens' concerns; in particular, by observing the strength of protests and counter-protests.

Keywords: Revolution, demonstration, protests, revolt, rebellion, strike, Arab Spring, homophily

JEL Classification Codes: D74, D72, D71, D83, C72

---

*Barbera is at MOVE, Universitat Autònoma de Barcelona and Barcelona GSE. He acknowledges financial support from the Spanish Ministry of Economy, Industry and Competitiveness, through the Severo Ochoa Program for Centers of Excellence in R and D (SEV-2015- 0563) and grant ECO2017-83534-P and FEDER, and from the Generalitat de Catalunya, through grant 2017SGR711. Jackson is at the Department of Economics, Stanford University, Stanford, California 94305-6072 USA, and is an external faculty member at the Santa Fe Institute and a member of CIFAR. Emails: and jacksonm@stanford.edu. He gratefully acknowledges financial support under ARO MURI Award No. W911NF-12-1-0509. We thank Ernesto Dal Bo and Santiago Oliveros for very helpful discussions and suggestions. We also thank Daron Acemoglu, Avi Acharya, Dan Bernhardt, Ethan Bueno de Mesquita, Raphael Boleslavsky, Steve Callendar, Ying Chen, Georgy Egerov, Jon Eguia, Hulya Eraslan, Navid Hassanpour, Navin Kartik, Dániel Pass, Arturas Rozenas, Mehdi Shadmehr, Scott Tyson, participants of various seminars and conferences, and a co-editor and referees for helpful comments.

When enough people agree that change is desirable, but is not being addressed by a government or other organization, then there can be room to force it. However, revolts and protests are risky, with substantial costs to participants if their actions fail - not just in terms of lost time and effort, but in some cases involving imprisonment, exile, or even death. There are also potential personal advantages, both moral and material, from having participated in successful revolutions. Hence, whether one unfolds hinges on how confident supporters of change are about the number of others who are also willing to take action. Because of that, it can be difficult to collectively act, even when a majority of the population supports change.

In this paper, we study various ways in which people learn from each other about the level of support for change, and how this affects the chances that people show up for a revolution and the chance that it succeeds. We study how people learn from interacting with others around them, as well as learning from things like a buildup of protests prior to a revolt. We also discuss how, in spite of the improvements in social media and communication, demonstrations and protests remain differentiated methods of signaling the intensity of preferences, and the conditions under which agents are willing to take risks in favor of change.

In what follows, we refer to the collective action as a "revolt", even though it should be understood that our model encompasses different types of actions under this generic name. In some extreme cases, success is the overthrow of a government, but in others, it may be a significant change in a political or business scenario, enough to produce a desired change in policy, or even just gathering media attention to change a company's policy. In some cases, the revolt involves violence while in others it may remain peaceful.

Collective action has been analyzed from many angles and the importance of beliefs and learning in enabling revolutions has been analyzed before, usually by means of specialized models (see below for more discussion). We answer different questions, providing new insights, and using a model that allows a broader and more unified analysis. Indeed, the first main contribution of our paper is to provide a simple and versatile model that sheds light on the many ways in which people can learn about the potential success of a revolt. People have a type which characterizes how much they gain from taking part in a revolt supporting change, taking into account the fact that a failed revolt involves costs. People with high positive types are willing to revolt even believing that there is a low probability that enough others will also join for the revolt to be successful. People with lower positive values are willing to revolt, but only if they are sufficiently confident that the revolt will be successful. People with negative values prefer the status quo and never participate. By meeting other people, a potential revolutionary learns about the distribution of types in the population. A main insight of our analysis is that this has countervailing effects: some supporters of the revolt will be discouraged as they will meet partisans of the status quo. Others become encouraged as they meet other supporters and become more confident that a revolt will be successful. In cases in which people were initially confident enough to participate in a revolt without any information, meeting others can cause too many become discouraged and disable a revolt. One important aspect of this is that when it becomes clear that too many will be discouraged, then even strong supporters who are confident that a majority prefers change will know that the revolt is doomed since they know that sufficiently many other supporters will become discouraged and not show up. Thus, small amounts of information can actually make revolts impossible - while no information or large amounts of information would lead to successful revolts; and we show how this depends on the prior beliefs and the correlation of preferences across members of the society.

In contrast, in cases in which people were initially too uncertain to revolt, the increased confidence from meeting other supporters can make it possible to have a successful revolution. Which effect dominates depends in intuitive ways on nicely quantifiable factors: prior beliefs, the relative benefit of success compared to the cost of failure, and how correlated preferences are across the population. For example, with high levels of correlation of preferences (more homogeneous populations) the encouragement effect dominates, while with lower levels of correlation (more heterogeneous populations) the discouragement effect dominates.

With this basic understanding in hand, we turn to studying how homophily impacts people's ability to learn about the potential success of a revolt or collective action. Homophily refers to people being biased to meeting other people with the same preferences. For instance, in a society stratified by social and economic classes, people may mostly only talk to others who have very similar backgrounds and circumstances, and thus similar preferences. This effect is often cited as to why some people were surprised at the outcome of the Brexit vote.[1] In the extreme, if people only interact with others who are very similar to themselves, they learn nothing about the overall prevalence of different preferences in the society. By reducing information content, homophily makes it harder to hold revolts in cases in which learning was necessary to enable a revolt, but it makes it easier to hold revolts in cases in which learning would otherwise unravel them. Thus, homophily provides a new and consequential angle on when collective action may succeed.

The above results are all framed within a model whose flexibility we then build upon to investigate other ways in which people learn about how a revolution might turn out.

Social media have been very critical in helping coordinate protests, but they cannot substitute for protests, since they are 'cheap-talk' and do not involve the costly signaling that protests provide.[2] A natural setting is one in which there are many people who would prefer change, but also in which many of them are not willing to pay the personal costs of being an active part of a revolt. They may communicate their support, but fail to turn out when action is needed. Holding a costly protest is a filtering device, which can signal whether there are sufficient numbers of people who are willing to act for change, not just cheer it on. Thus, holding a protest before a revolt can be a necessary step to enabling the revolt.

[1] For example, see "How to check if youre in a news echo chamber and what to do about it" by Tom Stafford, December 12, 2016, *The Conversation*.

[2] See Little (2016), Kiss, Rodriguez-Lara and Rosa-Garcia (2017), and Christensen and Garfias (2018) for discussions of how improved technology has changed agents' knowledge of others' preferences, and also enables better coordination regarding where and when to hold protests.

We also examine how extremism can undermine a revolt. When potential protesters meet, or sequences of protests are held, or a population learns about outcomes in similar countries, people may observe not only about how much support for change exists, but also from which constituencies that support emerges. This helps potential revolutionaries forecast what will happen after the revolt. If there are many extremists whose new agenda might not be preferable to the status quo, learning about their numbers can lead moderates to back away from a revolt that they might have otherwise supported.

Finally, we also provide a picture of how the gradual build up of mass protests can provide detailed information about preferences for a revolution, and how this extends to the case of contagion across regions or countries with similar political structures. For example, a government, or their partisans, may wish to hold counter-protests to signal the strength of support for the status quo. In some settings, if an initial protest in favor of change leaves some doubts as to the size of the support, counter-protests can become important in fully revealing the preferences of the population and can inhibit an eventual revolt. Governments can also respond by manipulating beliefs and sowing doubt via propaganda, increasing the costs of protests, or buying off some of the disenchanted.

Relation to the Literature

As we noted above, collective action has been analyzed from various angles, both theoretically and empirically, and in reference to different countries and circumstances. Although the topics we cover have been discussed in the literature, our results have not. It is not possible to survey the literature on the subject here, but let us discuss some key references, with more in the text as we proceed.

An early precursor on coordination games and thresholds for action is Granovetter (1978) - a theme explored in detail by Hassanpour (2017) who examines how network position and connectivity can affect leadership in collective action. Other important studies of collective action and mobilization build upon the herding literature of Banerjee (1993) and Bikhchandani, Hirshleifer, and Welch (1993). Some of these papers examine sequential observations and how these affect voting, a politician's decision, or a collective action (e.g., see Chwe (1999), Lohmann (1993, 1994ab, 2000), Bueno de Mesquita (2010), Kricheli, Livne, and Magaloni (2011), Loeper, Steiner, and Stewart (2014), Little (2016), Battaglini (2016), Shadmehr and Bernhardt (2016ab,2017)). The importance of information is central to all of these papers. At a high level, there is a common theme that there can be inefficiencies in outcomes due to imperfect information aggregation, and that learning can affect the ability to coordinate.

For example, in a voting setting, Lohmann (1993) examines costly political action (i.e., signalling) prior to voting, when voters are trying to estimate a state variable about which alternative is best to vote for, and her effects are based on the fact that only agents who have extreme preferences take political action, which does not provide full information about the state and may actually confound it. Kricheli, Livne, and Magaloni (2011) analyze a two-period model in which the first period turnout informs second period activists about whether they should try a revolt. Our focus is instead on how individuals learn from other individuals, and how homophily affects that information and incentives to attempt a revolution. Here, our effects are from direct meetings that all agents experience, and a main effect is a strategic one: agents know that some supporters will be discouraged and hence will not participate, and then through the strategic complementarity of the revolt discourages even those with strong information from participating.

Shadmehr and Bernhardt (2017) examine a game that is similar to ours, except that it has common values, and just two agents who both need to revolt in order to be successful. In that model communication can hurt the possibility of revolt when the status quo is bad, but help it when the status quo is good. The reasoning is that when the status quo is really bad, then without information people expect that there will be gains from a revolution and so are willing to undertake it, while with communication it is possible that agents learn that the revolution will not bring much improvement. In contrast, when the status quo is good, people generally don't expect the benefits to be high, and so would not be inclined to both revolt, unless they can communicate and coordinate. Thus communication can help prevent revolution for a very bad regime and encourage it for a good regime. Our analysis is complementary to that of Shadmehr and Bernhardt (2017) on several dimensions. We examine private values and many agents, and see a different reason as to why communication can help or hinder a revolution. In our setting, agents do not all communicate together, and so separate binary meetings can lead some agents to be encouraged while others are discouraged. It is this interplay that is at the heart of our analysis. Agents who would prefer to revolt and are fairly confident of the state can still expect the revolution to be doomed, simply because they know that too many other supporters of change will be discouraged about the chances of success they meet others who do not support change. In addition, our private values setting allows us to investigate how homophily affects collective action, along with some other issues about who sees what when.

From this expansive literature two things are clear: crowds can get things wrong, and learning from others is important and can change the possibility of collective action for good or bad. The simplicity of our model allows us to get a broader picture of learning from others that goes beyond previous analyses in several regards. First, our formulation allows us to see what happens when individuals learn from their neighbors and not from some collective communication or common history. This plays out differently from the cascades approach in which each person observes the previous ones and the usual herds or cascades can form, or the global games approach in which agents make deductions about the potential behaviors of others, or in which all agents communicate or see some common signals. Instead, in our model people are *simultaneously* learning from each other, but each with a *different* sample of the population. Such a model has not been analyzed before, and matches how many people learn. We show that this leads to countervailing effects that we are able to relate to prior beliefs and the correlation in preferences across the population. Second, the simple formulation allows us to introduce homophily, an important phenomenon that has not been studied in the context of collective action. We show how homophily impedes learning and how its effects depend upon the prior beliefs and the correlation in preferences across the population Third this is done within a unified framework that allows us to discuss a variety of other topics that have been analyzed piecemeal in the literature and in more complicated models, such as the role of protests, counter-protests. We also show how these interplay with the size of an extremist population, and government actions. Our model also enables a graphical analysis, which we emphasize throughout by the heavy use of graphics in our presentation.

1 A Static Model as a Building Block

We begin by describing a one-shot model in which a population must simultaneously decide whether to participate in a revolt (or rebel, protest, strike, etc.) in ignorance of other agents' types. We present an analysis of this model of collective action first, since it is a useful benchmark and building block for our results.

The Appendix provides a general form of payoffs as a base for future study, but for the main body of the paper we specialize the model to a simple case for much of our main analysis. In particular, we consider correlated private values as the relevant base case, but much of the analysis extends to common values, and the description of the more general model is in the Appendix.

The Players

A continuum of citizens of mass 1 are indexed by $i \in [0, 1]$.

They have a choice to participate in a revolt.

Again, we use the term 'revolt' but the model obviously has many applications.

The revolt is successful if at least a fraction $q \in (0, 1]$ of the population participates. If fewer than q participate, then the revolt fails.[3]

Payoffs

We consider the case in which an agent i's payoff is described by:

$$\begin{array}{lcc} & Success & Failure \\ Participate & \theta_i & -C \\ NotParticipate & 0 & 0 \end{array}$$

where $\theta_i \in I\!R$ is the payoff type of agent i, which is the private information of that agent.[4]

[3]This is a sharp model of collective action with a discontinuous change above or below a threshold. In reality there may be grey areas close to the threshold, but it does not seem that complicating the model in that way would add much insight. Uncertainty on the part of agents can already be used to smooth equilibrium behavior, and this version of the model is very tractable.

[4]This is a special case of a more general model that we describe in the Appendix.

The structure of this game is similar to that in the expressive voting literature (e.g., see Feddersen (2004) for a review). Here, the chance that any single agent will be pivotal in terms of making the revolt succeed rather than fail, in a large population and a nontrivial threshold, is negligible. In the continuum model it is in fact 0. Thus, we follow the usual approach from the voting literature and have θ_i be the marginal utility for having participated, conditional upon success. *So this is not the payoff that an agent gets in terms of expecting to be pivotal, but instead the utility they get from knowing or being able to say that they participated in a revolt that was successful:* from having been one of those who stormed the Bastille, or protests in Tunisia, etc. It will almost never be the case that one more or one fewer revolutionary would make the difference. People react to some more basic utility from the action itself, which can be motivated in many different ways, just as discussed in the voting context (again, see Feddersen (2004) for discussion and references). Clearly, someone who favors change will generally have a positive θ_i - which is the warm glow from having been there and supported the successful cause when it was needed. This is the positive feeling that someone who marched alongside Ghandi in the Salt March, or alongside Martin Luther King during the civil rights movement, feels for having been part of a history-changing protest. Generally, someone against change would have a negative θ_i as they would feel guilt or shame for being part of the revolt that overthrew a regime that they supported - for instance, a racist might have gotten a negative payoff from participating in the March on Washington for Jobs and Freedom in 1963.[5]

Uncertainty and a Base Model

If we let b_i denote i's belief about the probability that at least a fraction q of the other agents will participate, then the expected payoff to participation is $b_i\theta_i - (1 - b_i)C$ and the payoff from non-participation is 0. Thus, i is willing to participate if and only if $\theta_i \geq 0$ and

$$b_i \geq \frac{C}{\theta_i + C}. \tag{1}$$

As a base model, we consider a case in which there are some payoff types $\theta_L < 0$, called L types, for whom it is strictly dominant never to participate. Their beliefs become irrelevant. The other agents have a payoff type of $\theta_H > 0$, called H types. These types are willing to revolt providing they are sufficiently convinced that the fraction of others in the population that support change (and will act) is above q. Thus, the other critical thing to track is what an H agent (with payoff type θ_H) believes about the fraction of the population willing to revolt.

Agents with payoff type θ_H may have different beliefs b_i based on their information about the state; for instance, updating based on their own type and what they have learned about

[5]In some cases, people who favor the status quo might want to be seen as supporters of change if the status quo is overturned, so there are situations in which even some supporters of the status quo end up having a positive V_i. That is fine for the model, as we are agnostic on what drives V_i and just analyze its consequences.

other people's types. Let $b^* = \frac{C}{\theta_H + C}$ be the critical value of beliefs such that having a fraction of more than q agents who have with payoff type θ_H and $b_i \geq b^*$ is necessary and sufficient for there to exist an equilibrium with a successful revolt.

In terms of the base model's relative fraction of different payoff types, suppose that either $z > q \geq 1/2$ of the population are H types which happens with probability π, which we call the "High" state; or $1 - z < q$ of the population are H types, which happens with probability $1-\pi$, and we call the "Low" state. (It is direct to see that the symmetry between the fraction being z and $1 - z$ simplifies the expressions that appear below, but does not alter the intuition behind our results.) This is pictured in Figure 1 The parameter z captures

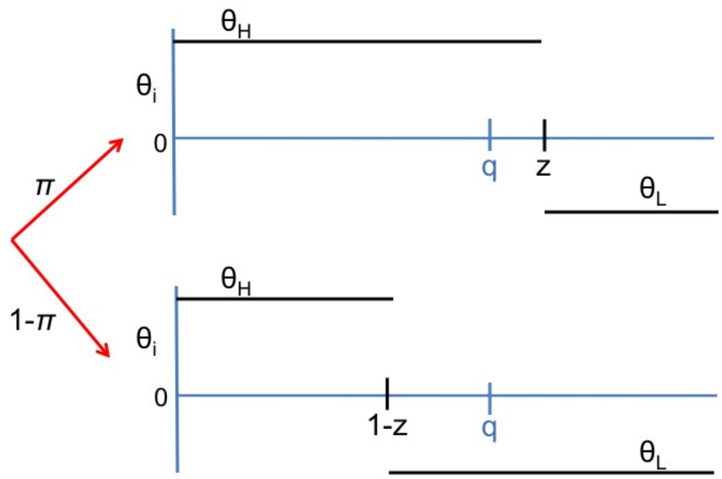


Figure 1: Two states, with the High state having probability π. The High state has more of the θ_H types (a fraction $z > q$) and the Low state has more of the θ_L types.

how homogeneous the society is. If z is very high, then the society is either almost all H or almost all L types, while if z is closer to $1/2$ then the society is more evenly balanced and the majority and minority groups are over more comparable sizes.

Let us assume that types are i.i.d. conditional upon the state of H or L.[6] Then, by Bayes' Rule, an agent i who is a θ_H payoff-type assigns a conditional probability of

$$p_i = \frac{\pi z}{\pi z + (1-\pi)(1-z)} \tag{2}$$

to the state being high. Thus, parameter z that describes homogeneity thus captures how correlated people's preferences are: A high level of z gives someone a high confidence that if they are type H then most other people are as well; while if z is closer to $1/2$, then knowing one's own type does not tell one much about other people's types.

[6]The usual caveat about the impossibility of having a continuum of independent random variables applies, but the model is easily approximated by standard techniques, so we simply work at the limit. See the discussion in the Appendix.

Note that an agent's beliefs about the high state do not necessarily correspond to beliefs that at least a fraction q of other agents will participate in a revolt, since that also depends on other agents' beliefs and their strategies, as we explore further below.

Equilibria

A strategy for player i is a function $\sigma_i : \mathbb{R}^2 \to \Delta(\{0,1\})$ that specifies a probability of participating as a (Lebesgue measurable) function of the agent's payoff type and beliefs about the state: $\sigma_i(\theta_i, p_i) \in [0,1]$. Let σ denote the profile of strategies.[7]

We examine Bayesian equilibria of the game. Later in the paper, when we consider dynamic versions of model, we examine weak perfect Bayesian equilibria.

As this is a simple coordination game, equilibria exist and in fact there is often a multiplicity. For instance, nobody participating is always a strict equilibrium: if none of the other agents participate then the revolt will surely fail and so it is a best response not to participate. However, in many cases there also exist participatory equilibria.

Analysis of the Base Game and the Possibility of a Revolt

In our base game in which agents only know their own payoff type, all H types' beliefs about the probability of the High state are given by equation (2). Equilibria come in three flavors.

First, there is always an equilibrium in which nobody participates. This is straightforward, and it is clear that this is a strict equilibrium. If nobody else participates, then an agent's payoff is strictly negative from participating as they pay the cost for certain.

Second, it is straightforward to see that if (and only if) the beliefs defined by (2) are above b^*:

$$\frac{\pi z}{\pi z + (1-\pi)(1-z)} \geq \frac{C}{\theta_H + C}, \tag{3}$$

then there also exist participatory equilibria. There is one in which all of the H types participate. This is (generically in the parameter space) also a strict equilibrium when it exists.

Third, when there exists a participatory equilibrium, there also exist mixed strategy equilibria in which H types are exactly indifferent between participating or not.[8] The mixed strategy equilibria are unstable, as slight perturbations of the actions lead best replies to converge either to the all participate or no participate equilibria. Thus, we focus our attention

[7]We work with strategies that are also Lebesgue measurable as a function of the agents' labels. Generally, the equilibria naturally depend only on agents' payoff types and information and not their labels, and so this is not really a restriction.

[8]With a continuum of agents, there exist a variety of such equilibria in which different agents use different mixtures - but the aggregate leads to exactly q agents showing up. Only some of those equilibria involve measurable strategies. In the world of a continuum, in order to get those sorts of equilibria to exist, one also has to put in place a rule that indicates the exact probability of the revolt being successful if exactly q agents show up (see Jackson, Simon, Swinkels, and Zame (2002)).

on the pure strategy equilibria which are strict and stable (in the sense that best replies converge back to the equilibrium under a perturbation of strategies).

In the base game, the equilibrium structure is then quite simple: the only stable and strict equilibria are the ones in which nobody participates and the ones in which all the H's participate. We do not take a stand on which of these two equilibria is more natural. In cases in which the unique equilibrium is non-participation, then the outcome is clear. In cases in which there exists a participatory equilibrium, one can find some conditions under which it is picked out via global games arguments. Since those arguments appear elsewhere,[9] we do not repeat them here.

The existence of an equilibrium in which the H types have a high enough belief that they expect a positive payoff from showing up, inequality (3), is pictured in Figure 2, as a function of π and z.

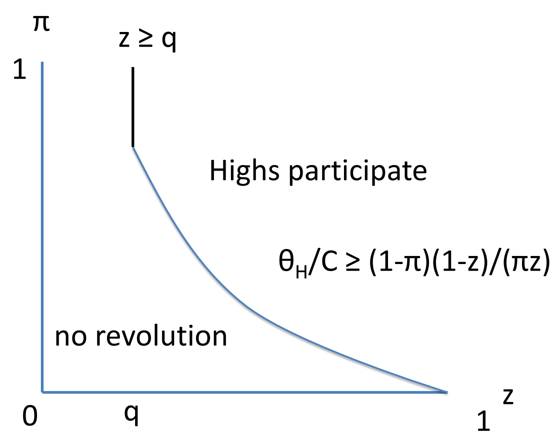


Figure 2: There is an equilibrium in which the H types participate if and only if the prior π and the correlation z are high enough.

We emphasize that there are two requirements for the existence of an equilibrium in which H types all participate:

- it must be that $z \geq q$, as otherwise even in the High state there would not be enough H types to be successful even if they were sure of the state, and
- it must be that beliefs of the H types put a large enough weight on the chance of success so that they are willing to participate, which is true if and only if $\pi z \theta_H \geq (1-\pi)(1-z)C$.

[9] That multiplicity has been extensively studied in the global games literature (e.g., see Angeletos, Hellwig and Pavan (2007)) and in the protest literature (Bueno de Mesquita (2010)). See also Weinstein and Yildiz (2007) for an argument that one could select either of the strict equilibria in our setting could be made the unique equilibrium with different introductions of slight noise into the model, depending on how that noise is modeled.

The first constraint is that z lies to the right of the vertical segment at $z = q$ and the second constraint is that π and z are above the level curve at which $\theta_H/C = \frac{(1-\pi)(1-z)}{\pi z}$. If and only if both of these are satisfied does there exists an equilibrium in which H types participate. There always exists an equilibrium in which nobody participates.

The model produces some intuitive comparative statics that follow directly from equation (3) and are pictured in Figure 3. We see that the range of values of π and z for which there is a revolutionary equilibrium shrinks as we decrease θ_H and/or increase C. This is consistent with evidence discussed by Kricheli, Livne, and Magaloni (2013) showing that increased costs lead to fewer protests, but then ones that are more likely to be successful when they occur.

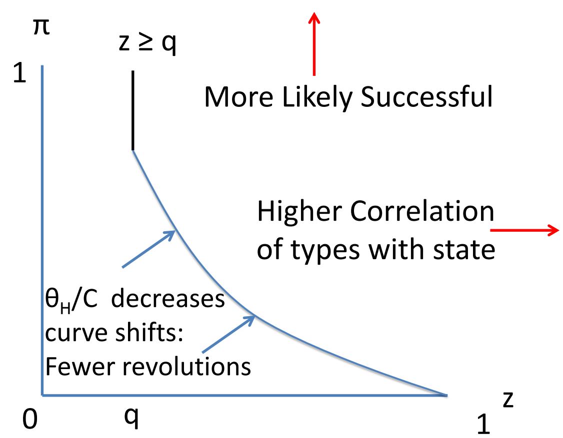


Figure 3: The range of values of the prior belief on the High state, π, and the correlation between types and the state, z, shrinks as the cost of the revolt increases or the value to H types from participating decreases. Also, as π increases, the likelihood of success increases, and as z increases there is a better match of the H types with the state.

There are H types in either state, and they act based on their beliefs conditional on the fact that they are a H type. So, they know that they still face a chance of failure as it is possible that it is the Low state and there are just not enough H types to succeed.[10] So, H types participate but the revolt still fails whenever the state is Low; and thus the likelihood of success increases as the likelihood of the High state, π, increases. Also, as z increases there is a higher correlation of the H types with the state: there are more H types who show up in the High state when the revolt is successful, and fewer who show up in the Low state when the revolt fails.

Note that the change from no revolt to a revolt is discontinuous: as π, z, and θ_H/C pass a threshold we can go to a regime that experiences no revolts to one that can have (large) ones.

[10]This is provided $z < 1$, as otherwise (if $z = 1$) types are fully correlated with the state and fully revealing and the analysis becomes trivial.

2 Learning from others Prior to a Revolt, and the Impact of Homophily

With the base model in hand, we can now turn to analyze how having agents learn additional information their own types affects the possibility of having revolts.

We begin with the question of what happens when people get to see some information about others' preferences, and how that depends on how homophilistic the society is.

2.1 Learning from Seeing One Other Agent's Type, Uniformly at Random

We first consider what happens if each agent gets to meet another agent and to see that other agent's type, where that agent is chosen uniformly at random – so that there is no homophily in the meeting process.

Each agent gets to talk to one other agent in the society and learn that agent's type.[11] This provides additional information to the agents, since now they have two signals about the state rather than just one.

If a H type sees another agent of a H type, then by Bayes' Rule, the agent's belief that the state is 'High' is

$$\frac{\pi z^2}{\pi z^2 + (1-\pi)(1-z)^2}.$$

If a H type sees that the other agent is a L type, then by Bayes' Rule, the agent's belief that the state is 'High' is

$$\frac{\pi z(1-z)}{\pi z(1-z) + (1-\pi)z(1-z)} = \pi.$$

So, agents with payoff type θ_H are now broken into two different types in terms of their beliefs, which can be denoted HH and HL.

The evolution of beliefs is pictured in Figure 4.[12]

Let's examine how this impacts an equilibrium.

Since $\theta_L < 0$, we only have to analyze the H type's incentives in order to characterize equilibria, since L types never participate regardless of their beliefs. As we see from Figure 4, some of the H types meet other H's and become more sure that it is the High state,

[11]Note that it is incentive compatible for agents to tell each other their types, and so it is without loss of generality to simply assume that types are observed when two agents meet. This does depend on the payoff normalization in our model. If we allowed the θ_L types to still prefer the revolt to be successful, but not want to participate, then that would induce them to lie. Alternatively, as long as people can observe each others' types as depending on some demographic variables (e.g., employment, income, etc.) then the types would be at least partly observable.

[12]We thank Santiago Oliveros and Ernesto Dal Bo for suggesting Figures 4 and 5.

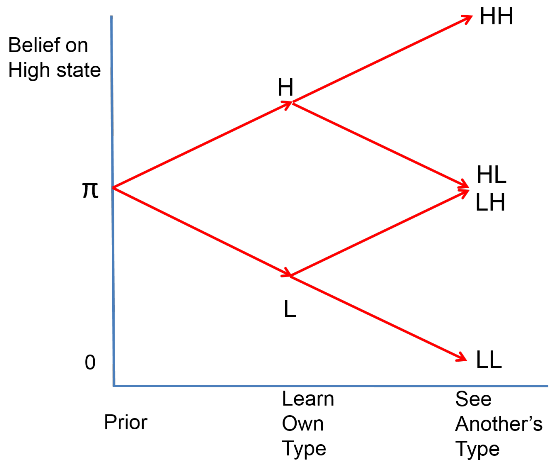
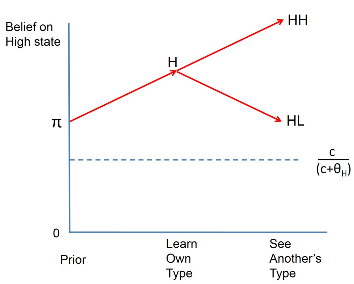
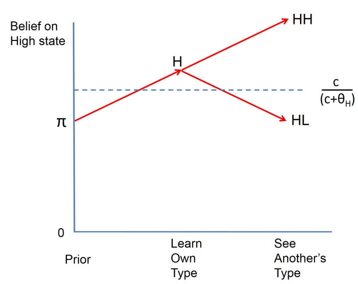


Figure 4: How an agent's beliefs evolve from the prior, to learning his or her own type, to meeting another agent and learning their type. For instance, HL indicates that an H type agent meets an L type agent.

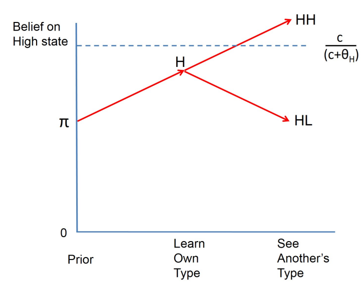
while others meet L's and put lower probability on the High state. Which effect dominates depends on how high the prior is and what the relative payoff from success is compared to the cost of failure. There are several possibilities, pictured in Figure 5.



(a) Meeting another makes no difference: Beliefs are high enough regardless of whom an H meets .



(b) HLs are discouraged. Without meeting another they would have showed up, but now will not. This disables a revolt if $z^2 < q$.



(c) HHs are encouraged. Without meetings they would not have showed up, but now will. This enables a revolt if $z^2 > q$. .

Figure 5: **The Effects of Meeting Another Agent**: Depending on the prior and the relative payoffs from the revolt, meeting another agent may either enable or disable a revolt.

First, as in panel (a) of Figure 5, the prior belief could be so high that even if an H type meets an L type, the H type is still convinced enough of the High state that the agent is willing to go to the revolt (beliefs are still above b^*). In particular, since an HL type's beliefs are given by π, there exists an equilibrium where all the H types show up *regardless of whom they meet* if and only if:

$$\pi\theta_H \geq (1-\pi)C. \tag{4}$$

This condition is more demanding than our previous equilibrium in the absence of meeting another agent, since it is asking whether H types go even when having lower beliefs induced by meeting an L and getting discouraged.

Second, as in panel (b) of Figure 5, it could be that there would have been a revolt if agents did not meet one another, but now the H types who meet L's are discouraged and no longer have enough confidence in the state to participate. The other H types know that some H's will be meeting L's and will no longer show up. If $z^2 < q$, then the H types who meet other H types will not be numerous enough to have a successful revolt on their own, and so even though HH types are confident in the state and believe it to be High with probability above b^*, they also know that the HL will not participate and so the revolt is doomed to failure. Knowing this, the only equilibrium is no participation.

Third, as in panel (c) of Figure 5, it could be that there would *not* have been a revolt if agents did not meet one another, but now the HH types (who meet another H type) are encouraged enough and numerous enough, to hold a revolt by themselves. For there to exist an equilibrium in which the HH types show up when they see another H type, two things are necessary: one is that they are sufficiently convinced of the High state that they are willing to show up, which requires that

$$\frac{\theta_H}{C} \geq \frac{(1-\pi)(1-z)^2}{\pi z^2}. \tag{5}$$

The second requirement is that there have to be enough of these HH types (in the High state) for the revolt to be successful, which requires that

$$z^2 \geq q.$$

If we view how this works in terms of the space of parameters π and z, Figure 6 shows that there are three different regions.

Comparing this to the no information case, Figure 7 shows the difference in equilibrium structures for the two settings:

In Figure 8 we see that the information from meeting another person helps enable the revolt when π (the prior prob of the High state) is low and when types are sufficiently correlated with the state and so seeing another H type is very informative. In contrast, it inhibits the revolt when the correlation between types and the state is low – in which case many people meet others who have low signals and thus become discouraged and the revolt unravels, since even those who meet others with high signals know that too few people will show up for the revolt to be successful.

So, to summarize, we see four different possibilities:

- With a high enough prior on the High state, there exists an equilibrium in which the H types to show up regardless of what they observe, in which case it would have been an equilibrium for them to show up without seeing another agent's type. Here the

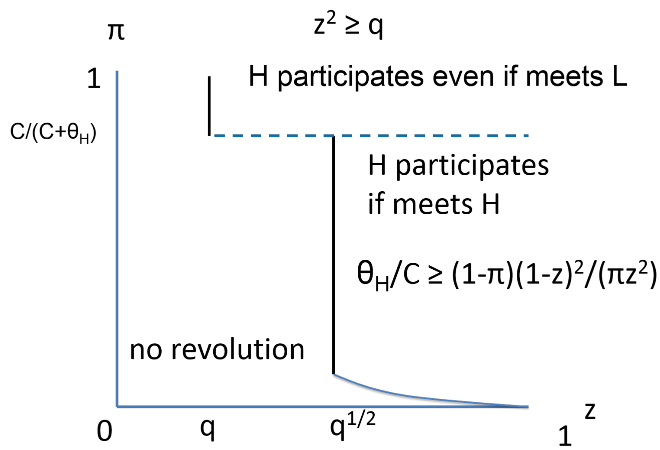


Figure 6: **Three Regions of Equilibria:** priors are so high that all H types show up regardless of whom they meet; or priors are in a region such that H types only show up if they meet an H type, and there are enough HH types to be successful ($z > \sqrt{q}$); and otherwise there is too little confidence or too few confident HH types to hold a revolt and there is no revolution.

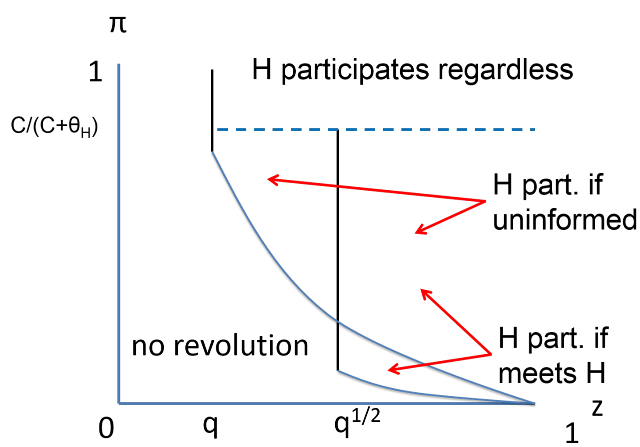


Figure 7: Five different regions: no revolt, always a revolt regardless of what info is, only a revolution if don't see signals, only a revolt if see signals and both are high, revolt if see signals or not - but only H types that see another H type show up.

equilibrium is the same as not observing anything, as the prior is strong enough so that information does not influence the agents' decisions. This happens if (4) holds (and $z \geq q$). In this case, it also would have been an equilibrium for all H types to show up without any information, and so there is no change in equilibrium structure in this parameter region.

- Next, there is a region in which there was an equilibrium for H types to show up

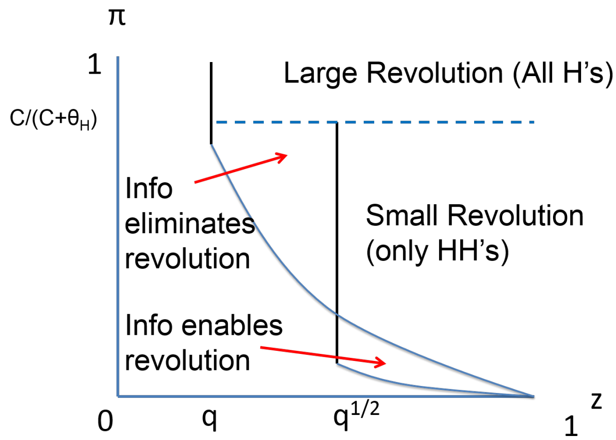


Figure 8: Sometimes information aids the revolt and other times it blocks it

without any information, but with information it is no longer an equilibrium for the H types to show up *even if they see another H type*. Here the equilibrium fails, not because those who see two H types are not convinced enough about the High state, but instead because they know that they are too small a fraction of the society to be successful. Here information is damaging for the H types as it would have been an equilibrium for them to show up if they did not see another agent! In this region the H types are ex ante worse off and the L types are better off. This happens if $z^2 < q < z$ while (3) holds.

- Finally, there is a region in which there is an equilibrium in which the H types show up if and only if they see that the other agent is an H type. This breaks into two pieces.
  - One part of this region is where it would also have been an equilibrium for them to show up without seeing anything. Here the equilibrium is now changed, as fewer H types show up in both states, but the revolt is still successful in the High and not the Low state. The H types are better off ex ante, and the L types are indifferent. Ex post, some H types are better off and others worse off in this setting than in the no information case, and overall they are better off ex ante. This happens if $z^2 > q$ and (5) holds, as does (3), while (4) do not.
  - The other part of this region in where it is an equilibrium for the H types to show up if and only if they see that the other agent is of the H type, but it would not have been an equilibrium for them to show up without seeing anything. Here the equilibrium is now changed, as seeing the other type enables H types to show up as they are now surer of the state, while without the information they would not have been able to have a revolt. Again, the H types are better off ex ante, and the L types are worse off. This happens if $z^2 > q$ and (5) holds, while (3) does

not.

We should point out that the basic intuition that having some information can disrupt a revolt, as it will inevitably discourage some higher types, extends to more general payoffs. For instance, the same result holds in a common values version of the model, as well as hybrids. Basically, seeing a low type lowers the high type's beliefs about the state regardless of the specifics of private versus common values, and so makes her more pessimistic. Knowing that some high types will be discouraged then means that even the more optimistic agents now know that their numbers are reduced.

We also note that the result easily extends to settings with more types. For instance, one can split the high types into many sub-types, who have similar but slightly different payoffs, and split the low types into many different types, all of whom prefer not to join the revolution. The result would remain unchanged. In fact, it is also easy to see that the result can also be extended to a full continuum model: what really matters is that some agents who would go to a revolution without any information can meet low types and be discouraged, and this can unravel the revolution. We describe the foundations of such a model in an online appendix.

2.2 How Homophily Dampens Learning

The previous analysis considers a case in which an individual meets another person chosen uniformly at random from the population. However, as we know, in many contexts people that we talk with are those around us in our networks and local communities. People are substantially more likely to interact with others who are similar to each other, not only in some base characteristic, but also in preferences and political views.[13]

To capture this, let us consider a variation on the above setting in which we incorporate homophily. Some of the meetings are biased towards own type. In particular, a direct way to model this (effectively, without loss of generality) is to have a fraction $h \in [0,1]$ of matches between highs and lows under uniform random matching that are instead re-mapped to have highs matched to highs and lows to lows. Thus, if $h = 0$ then there is no homophily and matching is uniformly random, while if $h = 1$ then highs always meet highs and lows always meet lows, and in between h regulates the bias towards meeting own types.

In terms of information, when $h = 1$, there is no information in a partner's type as it is then the same as the agent's own type regardless of the agent's type. The informativeness of the signal is highest when $h = 0$. However, given the non-monotonicities in equilibrium, the effect of homophily on equilibrium can be ambiguous, as we now show.

Homophily makes people more likely to meet their own type. Thus, they update their beliefs less from meeting own types as homophily increases. Meeting other types decreases

[13] For background on this empirical observation, termed "homophily", see McPherson, Smith-Lovin and Cook (2001) and Jackson (2008,2019).

under homophily, but the rate at which it decreases is the same across states and so the updating conditional on meeting an opposite type is unaffected. This is as pictured in Figure 9. Thus, since homophily's impact is to dampen the updating conditional on meeting one's own type, and also to reduce the chance that other types are met, it thus reduces learning all around.

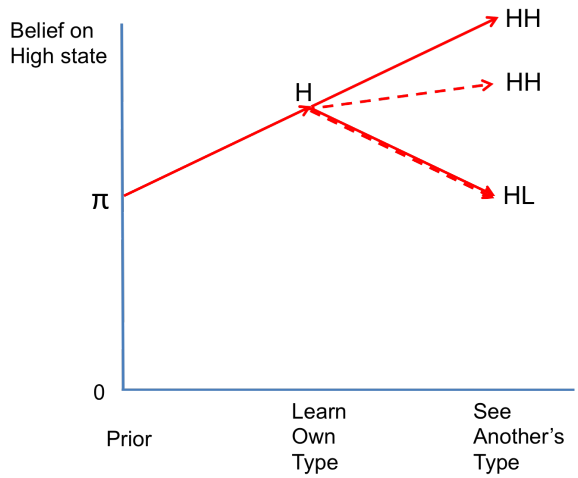


Figure 9: Homophily lowers the updating of beliefs from an H meeting an H, but does not affect the updating of an H meeting an L (the *relative* odds of that happening across states are not affected by homophily). However it does lower the frequency with which H's meet L's regardless of the state

In particular, the probability of an H type seeing another H type with homophily $h \in [0, 1]$ is $z^2 + z(1 - z)h$ in the High state and $(1 - z)^2 + z(1 - z)h$ in the Low state.

This leads to a new constraint for the equilibrium in which a H type is willing to participate if and only if seeing another H type. These is a variation on the previous analysis, using Bayes' rule. It is summarized in the following proposition.

Proposition 1 *If $\pi > \frac{C}{C+\theta_H}$, there is no effect of homophily on the existence of an equilibrium in which there is a revolt: H types participate regardless of whom they meet.*[14] *In contrast, if $\pi < \frac{C}{C+\theta_H}$, then there exists an equilibrium in which a revolt occurs if and only if:*

$$\frac{\theta_H}{C} \geq \frac{(1-\pi)[(1-z)^2 + z(1-z)h]}{\pi[z^2 + z(1-z)h]},$$

and

$$z^2 + z(1-z)h \geq q;$$

and the participants in the revolt are the H types who meet other H types.

[14] At the exact equality there exist a continuum of equilibria in which the H types that meet L types mix on participating as they are exactly indifferent.

We have already provided the proof in the derivation from the previous sections and above.

Note that the second inequality gets easier to satisfy as h increases, while the first one gets harder to satisfy as h increases: this is the tradeoff as homophily is increased. Homophily decreases information, making the individual incentive to participate harder to satisfy, but also leads to fewer agents who are discouraged by meeting L types. Which effect dominates depends, again, on the relative prior and correlation of types with the state.

This leads to the adjustment in the equilibrium structure as pictured in Figure 10.

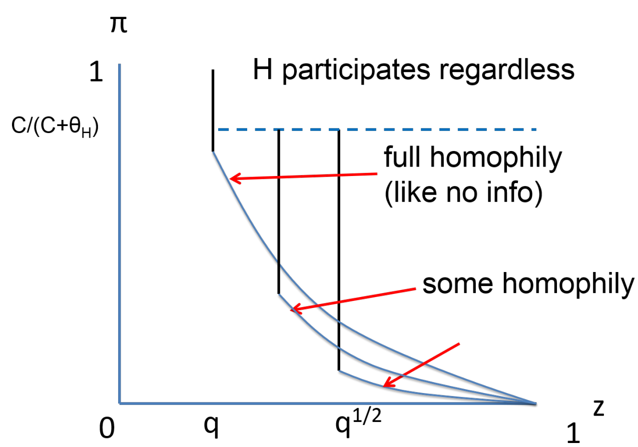


Figure 10: Homophily (assortativity in meetings) changes the equilibrium structure.

We see that higher homophily increases the region of having a revolt if the prior is in a middle range, since more highs then see high signals and be willing to join, but higher homophily reduces the region for low priors and high z since it decreases the information contained in a meeting, which otherwise could have given more confidence to H types who meet other H types.

Under-Estimating Homophily

The above discussion presumes that agents are aware of how much homophily there is in their society. In many settings people may understand that they are more likely to meet others who are similar to themselves, but have no idea of how extreme homophily tends to be.

Under-estimating homophily can inflate the confidence that agents derive from meeting others. With high levels of homophily, agents are very likely to meet others who have similar types. For instance, without knowing the extent of homophily most H types would meet other H types (even in the *Low* state) and could become over-confident that it is the *High* state. For example, in situations in which a revolt would not take place without meeting anyone, as in panel (c) of Figure 5, H types would become more confident and be willing to take part (provided $z^2 > q$), if they meet another H type without correcting for potential extreme homophily and so end up over-participating in the *Low* state.

Learning from Seeing Many Other Agents' Types

Next, we consider what happens in the same setting when agents get to see several or even many other agents' types, say some number $m \geq 1$, presuming there is some chance of meeting the other types of agents – so that homophily is not complete. For instance, for each meeting: with probability h, H types are matched with other H's and L's with L's, and with probability $1-h$ agents are matched at random so that they expect to meet other types in proportion to their weight in the population. This can be done independently over matches, so that if a person meets a sequence of people then that sequence is biased towards own type (at a rate h), but each meeting is independent of the other meetings.

PROPOSITION **2** *For any $z > q$, π, $h < 1$, and θ_H/C, there exists a number of signals, $\overline{m}$, above which (if $m \geq \overline{m}$) there is an equilibrium which involves an H type participating in a revolt conditional upon a sufficient fraction of H types being observed. Moreover, as the number of others observed increases, the fraction of H types participating in the* High *state converges to 1 almost surely and the fraction of H types participating in the* Low *state converges to 0 almost surely: the revolt is perfectly effective in the limit.*

We can also see directly how the speed of convergence of beliefs is affected by the presence of homophily. The usual rate of convergence (expected error of the posterior) would be proportional to $1/\sqrt{m}$, where m is the number of observations of draws observed from the distribution. That would be the rate with no homophily, by the Central Limit Theorem. Under homophily, however, the effective number of real observations is expected to be $1+(1-h)m$, or roughly $(1-h)m$ (on average). Thus, the rate of convergence is roughly proportional to $1/\sqrt{(1-h)m}$ - which gets slower as h increases. So, homophily slows the rate of learning at a rate $1/\sqrt{1-h}$, which becomes infinite as homophily tends to 1.

We sketch the key idea behind the proof of the proposition.

As long as homophily is not complete ($h < 1$), the posterior belief on the High state converges to 0 or 1 and is correct almost surely. This can be derived from Levy's Zero-One Law: The fraction of other agents that an H type meets in the High state tends to differ from that in the *Low* state. The H type would expect to meet a fraction p_H of H types in the High state in expectation, and some p_L in the Low state. As long as homophily is not complete ($h < 1$), $p_L < p_H$. An event that the H type meets a fraction of more than $(p_L + p_H)/2$ over an infinite sequence of meetings is a tail event, and its probability is then 0 or 1 (depending on the state). Thus, as an agent sees a sequence of meetings, the posterior that they have will converge to either 0 or 1 in the limit as m grows, almost surely (by Doob's Martingale Convergence Theorem). Thus, as the number of people agents meet becomes large, then the fraction of H types in the population whose posterior has high certainty on the correct state tends to 1. With enough observations, H agents can be sure that there will be a fraction above q who have sufficient belief on the High state to make the revolt successful with high probability in the High state.

Thus, even with homophily the expected long stream of people that an agent will meet will differ by the state, and so agents will eventually be sure of the state, and so there exists an equilibrium in which agents who are H types show up whenever their posterior is above a high enough threshold, and in the limit they are almost always successful. For any finite number some agents will make mistakes, but that proportion vanishes in the limit.

The interesting aspect, putting this result together with our analysis of just seeing one other agent's type, is that information can have non-monotonic effects: small amounts of information can be disruptive, while large enough amounts of information are always enhancing. As a numerical example, let $z = 2/3$ and $q = 1/2$ $\theta_H \pi/[C(1-\pi)] = .8$. Here the non-montonicity of information is clear: we have an equilibrium with H types participating if $m = 0$ or if $m = 2$, but *not* if $m = 1$.

As such, we might expect that technological advances, including social media and cell phones, that allow agents to learn about the opinions of greater numbers of others to eventually lead to more accurate protests. As people learn about greater number of others the correlation of the size of the protest with the state will increase. This is consistent with empirical background on this sort of effect.[15]

Protests of nontrivial size may become more or less frequent depending on the parameter region, but then much more likely to be successful when of large size.

We can solve for some aspects of the equilibrium in more detail. An individual now gets to see m random other individuals' types. We now can see how many signals they must see before they are willing to participate.

There are two constraints that need to be satisfied in order to have an equilibrium where some people participate. One is that the threshold must be high enough so that at least some agents are sufficiently convinced that it is the High state so that they would be willing to revolt (presuming that revolt will be successful in the High state). This requires that the threshold exceed some lower bound, $\underline{t}(m)$. The other constraint is that in the High state, not too many of the H types end up failing to pass the threshold, as otherwise there will be insufficient participation for a successful revolution. This puts an upper bound on the threshold $\bar{t}(m)$. So, collective action is feasible only if the lower bound is below the upper bound, and then agents who see enough to be sufficiently convinced that it is the High state, will also be sufficient in number to succeed in the High state.

Let us examine first the lower bound $\underline{t}(m)$. If a player is of type θ_H and sees t out of m other H types, then the conditional probability on the state that z of the population are of the H type is

$$\frac{\pi b(t+1, m+1, z)}{\pi b(t+1, m+1, z) + (1-\pi) b(t+1, m+1, 1-z)}$$

[15] See Breuer, Landman and Farquhar (2012) and Farrell (2012), as well as Kiss, Rodriguez-Lara, and Rosa-Garcia (2017), Manacorda and Tesei (2016), Pierskalla and Hollenback (2013), and Steinert-Threlkeld, Mocanu, Vespignani and Fowler (2015).

where $b(t, m, z)$ is the binomial probability of seeing t positives out of m trials that are positive with probability z. So to get an agent to act (presuming the agent expects success conditional upon the High state) requires:

$$\theta_H/C \geq \frac{1-p_i}{p_i} = \frac{(1-\pi)b(t+1, m+1, 1-z)}{\pi b(t+1, m+1, z)} = \frac{(1-\pi)(1-z)^{2t+1-m}}{\pi z^{2t+1-m}}.$$

Solving this with equality allows us to deduce $\underline{t}(m)$

$$\underline{t}(m) = \frac{m-1}{2} + \frac{\log\left(\frac{\theta_H \pi}{C(1-\pi)}\right)}{2\log\left(\frac{1-z}{z}\right)}.$$

Next, we solve for $\bar{t}(m)$, the threshold such that if that were used, then the fraction of agents who show up would be at least q conditional upon the High state. This would be the largest t for which

$$(1 - B(t-1, m, z))z \geq q,$$

where $B(t-1, m, z)$ is the c.d.f. of the binomial distribution (so the probability that there are $t-1$ or fewer other H types out of the m observed when drawn with probability z). Thus,[16]

$$\bar{t}(m) = B^{-1}_{m,z}\left(1 - \frac{q}{z}\right) + 1.$$

In order to have agents be sufficiently confident of the High state, and also that there will be enough others also confident of the High state in the High state for the revolt to succeed, it must be that $\bar{t}(m)$ be at least as high as $\underline{t}(m)$. In the limit, $\bar{t}(m) \to zm$, while $\underline{t}(m) \to m/2$, and so eventually $\bar{t}(m) > \underline{t}(m)$.

Note that if we add homophily, then this slows the rate of informativeness of our signals. If you get to meet a hundred people, but more than ninety percent of them are similar to you in terms of always having the same political views, then that is almost like meeting just ten people. Thus, with substantial homophily the interaction rates before people really learn about the world might need to be high. Also, it is worth noting that, to the extent to which people do not fully understand the homophily around them, then that can make H types more confident that the state is High, regardless of the true state. So homophily slows the learning of fully rational people, but can lead naive people to overestimate the chance that the state matches their type.

3 Extensions of the Model

We now illustrate a variety of other issues whose main features and subtleties can be addressed via simple extensions of the base model.

[16] Here, the inverse of B is rounded downwards, so it is the largest value of t for which $B(t-1, m, z) < 1 - \frac{q}{z}$, which then assures that t is the smallest values for which the chance that at least t H types are observed is at least q/z.

3.1 Learning in a Dynamic Version of the Model

So far we have examined information revelation as agents meet some other people from the population. Another important informational channel is having mass protests in which agents protest before a revolt. These can be important precursors to a strike or revolt as they signal information in a much broader and more revealing way than people just seeing the preferences of a few friends.

Learning from a Protest before a Revolt To study the role of protests, we enrich the model so that there are two periods and three types. The types are $\theta_L, \theta_M, \theta_H$. Now the values are purely private, and the highest value types simply are more disadvantaged by the current government, and the moderate types would also prefer to overthrow the government if it is possible, but are harder to convince to join the revolt since they are not as dissatisfied as the higher types.

There are two states. In the High state $1-z$ of the population are θ_L and $z/2$ are θ_M and $z/2$ are θ_H, while in the Low state z of the population are θ_L and $(1-z)/2$ are θ_M and $(1-z)/2$ are θ_H.

So, this is exactly the same as our first model, except that we have split the H types equally into moderates and highs. This allows us to see the value of having protests before the revolt. This is pictured in Figure 11.

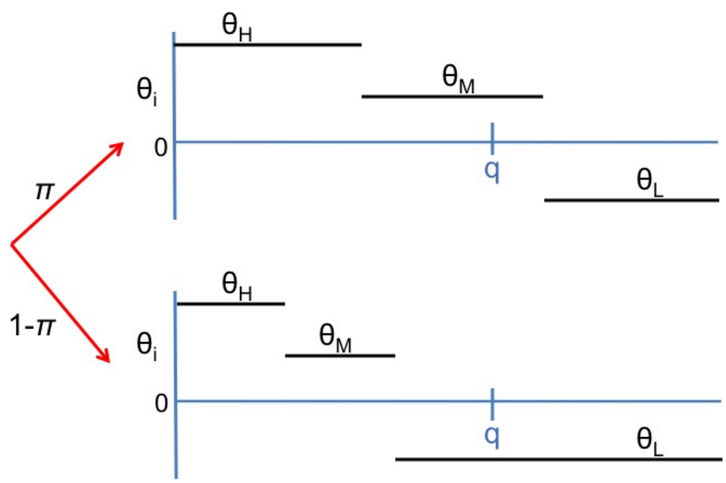


Figure 11: Two possible states. By seeing how many θ_H types turn out at a protest, the state is revealed, which can enable a revolt.

So, there is a first period in which the population can hold a protest, and then a second period in which they can hold the revolt. They can skip the first period if they wish, but it signals information about the state.

We let the cost of having participated in a protest or revolt if the revolt is not ultimately successful depend on the period, as participating in a protest may have less at stake then a revolt. For the first period let the cost be c and for the second period denote it by C.

Let us consider a case in which

$$\frac{\theta_M}{C} < \frac{(1-\pi)(1-z)}{\pi z},$$

but $z \geq q$.

So, without any additional information, the moderates are too frightened/pessimistic to participate in the revolt.

However, note that if

$$\frac{\theta_H}{c} \geq \frac{(1-\pi)(1-z)}{\pi z},$$

then it is possible to have the revolt.

The highs are willing to protest in the first period. If $z/2$ of them show up, then the moderates learn that it is the High state and the revolt takes place in the second period. If only $(1-z)/2$ of them show up in the first period, then the protest is a failure and there is no revolt in the second period.

This illustrates the possibility of having successive protests, where people learn about how many people are dissatisfied by observing the size of the turnout, and more extreme individuals protest earlier, enabling more moderate types to assess the state and join later if things look strong enough.

It should be clear that with richer heterogeneity one could build richer versions in which protests gradually escalate over time, and which several successive protests are needed, over time and/or geography, before sufficient certainty is reached to hold a successful revolt.

Homophily and Protests Our discussion about the role of protests focused on situations in which the prior beliefs of the moderates were insufficient to revolt without the information gained from a protest. Other situations in which a protest helps are ones in which moderates meet others but are not willing to revolt. For instance, one case is such that there is sufficient homophily such that moderate beliefs remain to low to prompt willingness to participate even if they meet other supporters. Thus, protests can help overcome homophily since people see the actions of many people, and they are sure of what others have seen.

The Difference Between Learning from Polls and Protests One question that we have not yet addressed, but is important, is why one needs protests at all in a world where people can hear about how others feel via polls and/or social media. In the above example, why do they still need to turn out at a protest in order to convince the population to revolt rather than just expressing their preferences in a poll or on some social platform?

The answer is that protests involve costs - and so agents must be sufficiently willing to participate to overcome those costs. Having many agents willing to pay those costs can signal to others that there is enough of the population willing to take costly action, that the revolt has a chance of succeeding. In contrast, polls and social media may involve much lower costs, and so agents simply saying that they support change does not indicate that they would be willing to act if needed. This is illustrated in the following example.

Suppose that payoffs are:

$$\begin{array}{lcc} & Success & Failure \\ Participate & \theta_i & -C \\ NotParticipate & a_i & 0 \end{array}$$

Here, agents who have $a_i > 0$ and $\theta_i > -C$ would like to see the revolt succeed. However, those who have $a_i > \theta_i$ have a dominant strategy not to participate. These are non-activist people who prefer to have others participate, but still would like to see change.

In this sort of setting, if one holds a poll to see who favors change, the agents who have $a_i > \theta_i > -C$ and $a_i > 0$ will say that they favor change. However these people cannot be counted upon to show up for the revolt when it is needed. Thus, the poll does not differentiate between people who favor change and those who support it enough to do something about it. In contrast, a protest can be costly to show up for, and so can screen out the non-activists and give a more accurate assessment of agents who are willing to act for change.

Thus, protests can be essential for successful further action and change in ways that polls and other sorts of media posting and cheap-talk might not.[17]

The Arab Spring Another variation on the above example is one in which there are not two periods, but instead two correlated countries. If one country has a large enough turnout in its revolt, then other country's population may learn about their own state and revolt as well.

Let us consider our original setting, but the only difference is that there are now two countries. The have the same probability of a High state, designated by π, but differ in the value and costs to H types, and the correlation of types with the state. We use the obvious notation: $(\theta_{H1}, C_1, z_1, q_1), (\theta_{H2}, C_2, z_2, q_2)$

The states of the two countries are correlated, with the correlation in High states being $\rho \geq 0$. In particular, the probability of the High or Low states for the respective countries are given by:

$$\begin{array}{rcc} & High_2 & Low_2 \\ High_1 & \pi^2 + \rho\pi(1-\pi) & \pi(1-\pi)(1-\rho) \\ Low_1 & \pi(1-\pi)(1-\rho) & (1-\pi)^2 + \rho\pi(1-\pi) \end{array}$$

Let us suppose also that $z_1 \geq q_1$ and $z_2 \geq q_2$, so that both countries can have successful revolts in their respective High states.

[17]Note that in a very repressive regime - that penalizes people who even say they support change - then it would be possible for that to provide a costly signal. However, that would only work if sufficiently many people are able to express their opinions, and such very repressive regimes may also censor information about any opposition.

Suppose that

$$\frac{\theta_{H1}}{C_1} \geq \frac{(1-\pi)(1-z_1)}{\pi z_1}$$

but

$$\frac{\theta_{H2}}{C_2} < \frac{(1-\pi)(1-z_2)}{\pi z_2}.$$

This is a world in which members of country 1 are sufficiently unhappy, or convinced of the High state, that a revolt is possible for that country on its own, while country 2 fails to satisfy that constraint, and so its members would only be willing to revolt if they are sufficiently convinced.

This is consistent with the data on the Arab Spring collected by Brummitt, Barnett, and D'Souza (2014), who find a significant correlation between the unemployment rate in countries and the date of first protest (e.g., Tunisia had higher unemployment than Egypt than Syria, and the first date of protests occurred in that order - and they analyze fifteen countries in total).

In this case, if country 1 holds its protest/revolt, then country 2 can learn about the state, provided there is sufficient correlation.

In particular, some direct calculations of the posterior conditional on success in country 1 (together with the appropriate variation of (3)) show that if

$$\rho \geq \frac{\frac{C_2(1-z_2)}{z_2\theta_{H2}} - \frac{\pi}{1-\pi}}{\frac{C_2(1-z_2)}{z_2\theta_{H2}} + 1},$$

then there is an equilibrium with contagion.

Beyond the Arab-Spring example, a careful analysis of the Swing Riots of the 1830s by Aidt, Leon, and Satchell (2017) provides convincing evidence of a contagion of riots; again much along the lines predicted here.

3.2 Learning when there are Extremists: Forecasting the Post-Revolution World

The analysis so far has been on situations in which the forecast of what might happen after a revolution does not depend on the state. The state determines whether the revolt succeeds or not, but if it is successful, then the forecast of what will happen was not state-dependent.

In many situations, however, participation may depend on what people expect to happen after a revolution – which involves their expectations of what a new government will be like.[18]

[18] See Shadmehr (2015) for an analysis of an endogenous agenda as part of a revolution. Our example here presumes that there is no ability to commit to what will happen after the revolution.

Such an enrichment of the model can add to the analysis of all of the situations we have discussed so far: meeting and learning types of other agents, observing protests, and observing the outcomes from other countries. Any of these information revelations can include not only information about number of dissenters and likelihood of success, but also about the size of relative factions of potential revolutionaries and who might emerge in power after a revolution.

To explore how potential conflict after a revolution can affect the revolt, let us consider the following variation on our basic model.

Suppose that there are now three types: θ_L support the government and never want to participate, θ_M are moderates who will support a revolt, but only if they are the majority of the revolutionaries and get to impose a moderate government after a successful revolt; and extreme types θ_E, who want a revolt whenever it would be successful regardless of the next government.[19] In particular, participating moderate agents get θ_M if the revolt is successful and there are more moderates than extremists, and get $-C$ otherwise. Extremists get θ_E if the revolt is successful and they outnumber moderates, $\alpha\theta_E$ if the revolt is successful and moderates outnumber extremists, and $-C$ if it fails.

In particular, moderate types prefer to participate in the revolt only if the fraction of moderate and extreme types exceeds q, but *also* only if the fraction of moderates exceeds the fraction of extreme types.

A state ω of the world is now a list, $\omega(\theta_L), \omega(\theta_M), \omega(\theta_E)$, of the fractions of the population that are of the corresponding types.

There are three states $\omega \in \{\omega^L, \omega^M, \omega^E\}$:

- Low state: $\omega^L(\theta_M)+\omega^L(\theta_E) < q$, so the revolt will fail even if moderates and extremists participate.

- Moderate state: $\omega^M(\theta_M) + \omega^M(\theta_E)$, but $\omega^M(\theta_M) < q$ and $\omega^M(\theta_E) < q$ (so the revolt will succeed if and only if both moderates and extremists participate), and moderates outnumber extremists, $\omega^M(\theta_M) > \omega^M(\theta_E)$.

- Extreme state: $\omega^E(\theta_M) + \omega^E(\theta_E) \geq q$, but $\omega^E(\theta_M) < q$ and $\omega^E(\theta_E) < q$ (so the revolt will succeed if and only if both moderates and extremists participate), and extremists outnumber moderates $\omega^E(\theta_M) < \omega^E(\theta_E)$.

There are different equilibrium possibilities depending on the prior probabilities of the states, π^L, π^M, π^E. Here we focus on the case without communication, although the extension to communication is straightforward and parallels that above.

In order for a revolt to be possible, it must be that the moderates place a high enough probability on the moderate state (conditional on being a moderate), while the extremists

[19] See Acemoglu, Hassan, and Tahoun (2015) for a description of conflicts between different revolutionary groups during the Arab Spring.

place a high enough probability on both the moderate and the extreme state. In particular, it is straightforward to check that the necessary conditions for having a revolt are that

$$\theta_M/C \geq \frac{\pi^L\omega^L(\theta_M) + \pi^E\omega^E(\theta_M)}{\pi^M\omega^M(\theta_M)}.$$

and

$$\theta_E/C \geq \frac{\pi^L\omega^L(\theta_E)}{\alpha\pi^M\omega^M(\theta_E) + \pi^E\omega^E(\theta_E)}.$$

This can allow a revolt to take place in both the moderate and extreme case, provided the prior on the moderate state is high enough relative to the extreme state for the moderates.

It is easy to see how this then enhances the analysis of meeting others, seeing protests before a revolt, and seeing the outcome in other countries. If any of those processes reveal sufficient likelihood that it is the extreme state (or that the revolt would fail), then the moderates would no longer participate. Thus, the conditions for the revolt to succeed require sufficiently high prior information, or revelation of a high likelihood, of it being the moderate state. Again, information could be either encouraging or disruptive to the revolt, depending on the state and prior probabilities.

For example, extending the analysis from above in which two stages of protests can enable a revolt, we could also view that example's high types as the extremists. The composition of extremists versus moderates in the High state then matters. We could split that state into two sub-states: one in which the extremist high types are in the majority of those who favor change, and the other in which the moderates are in the majority of those who favor change. This makes for interesting dynamics, as if the first period protest shows that there are too many extremist high types, then the revolt would fail, as the moderates would prefer to avoid an extremist state.[20] The equilibrium thus then only successful in the second period if enough people - but not too many extremists - show up in the first period protest. Similarly, if a revolt in a correlated country turns too extreme, it may discourage a nearby population from revolting.

How Counter-Protests can Enhance Learning

Let us discuss the role of counter-protests.

Protests can be useful in signalling to the government the level of support for a policy change, and counter-protests can be useful in signalling the level of support for keeping the current policy.

We make this point in the context of a setting in which after a first stage of protests in which those supporting change choose whether to show up, there is a second stage in which those who support the status quo, after seeing the turnout in the first stage, can choose whether to show up to a "counter-protest" that shows support for the government.

[20]For an illuminating but different discussion of how information revelation by a government about potential counter-policies can affect revolutions, see Shadmehr and Bernhardt (2017).

The population consists of three equal-sized groups that can differ in their preferred policies, since three groups is just enough to allow for variation in which is the most preferred policy and also to allow different sizes of protests and counter-protests to non-trivially signal the state. The groups can either support change or no-change, which we denote by $\mathcal{C}$ and $\mathcal{N}$. The groups can also either be strong supporters or weak supporters - in terms of how much they prefer their choice to the opposite choice. We denote these by $\mathcal{S}$ and $\mathcal{W}$.

In terms of preference parameters:

- $\theta_{\mathcal{CS}} > \theta_{\mathcal{CW}} > 0$
- $\theta_{\mathcal{NS}} = -\theta_{\mathcal{CS}}$, and
- $\theta_{\mathcal{NW}} = -\theta_{\mathcal{CW}}$.

So, a group's preference is one of four types: $\mathcal{CS}, \mathcal{CW}, \mathcal{NW}, \mathcal{NS}$. A state is listed as a triple of each group's preference type. For instance, $(\mathcal{CS}, \mathcal{CS}, \mathcal{NS})$ indicates that the first two groups both strongly prefer change while the third group strongly prefers no change.

With three groups and four types for each group, this leads to 64 possible states. To simplify the exposition, we focus on just four states - which capture the main ideas. Obviously, the analysis extends to including all 64 states depending on the prior probability on the various states, provided there is some uncertainty after a first protest, and sufficient likelihood that a counter-protest will resolve that uncertainty when it arises. The main point that counter-protests can be useful for learning holds in the more complicated setting, but then specifying all of the possible priors for which this holds becomes intractable, so we just illustrate the point for one possible prior that has weight on four possibilities.

In particular, we presume that one group prefers change, one group prefers no-change, and the remaining group is the only one that could be on either side - so there are two "partisan" groups whose direction of preference is known, just not their intensity, and one "pivotal" group which could have any preference and whose preference always determines the direction of the majority preference. We focus on 4 key states.

State 1 $(\mathcal{CS}, \mathcal{CS}, \mathcal{NS})$

State 2 $(\mathcal{CS}, \mathcal{CW}, \mathcal{NW})$

State 3 $(\mathcal{CW}, \mathcal{NW}, \mathcal{NS})$

State 4 $(\mathcal{CS}, \mathcal{NS}, \mathcal{NS})$

This is pictured in Figure 12

Let these four states be equally likely.

The optimal policy (in terms of a utilitarian goal of maximizing total welfare) is change in states 1 and 2, and no change in states 3 and 4. In states 1 and 4, all have strong preferences but either have 2/3 of the population in favor of change or in favor of no change. In states

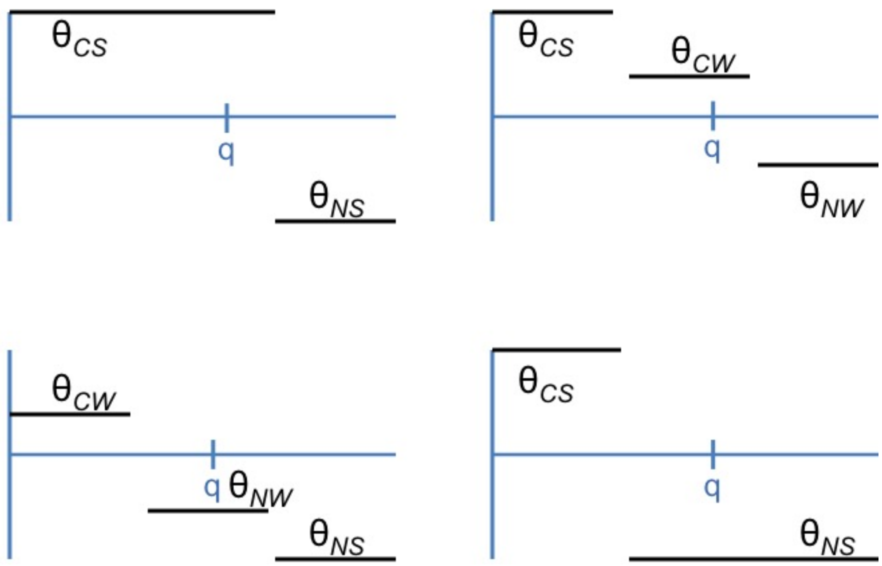


Figure 12: The four states that we focus on. In the two on the top, change is overall utility maximizing, while in the two on the bottom no-change is overall utility maximizing.

2 and 3, there is a mixture of weak and strong preferences, but the strong preferences are always on the side of a majority, and so the preference on the majority side is stronger than the minority.

Proposition **3 (Counter-Protests)** *Suppose that $\theta_{\mathcal{CS}} > C/2$. Then there exists an equilibrium in which:*

- *a protest is held by all $\mathcal{CS}$ types.*
- *a counter-protest is held by $\mathcal{NS}$ types if there is a protest in which only 1/3 of the population shows up.*
- *weak types never show up to a protest or counter-protest.*
- *there is a successful revolution (or the government voluntarily enacts change) if either 2/3 of the population shows up at the original protest, or if there is a counter-protest and nobody shows up to that. Otherwise, they do not make any change.*

The four resulting cases are pictured in Figure 13.

The reasoning behind the proposition is straightforward and so we simply explain it here. The possible outcomes under the prescribed strategies are:

- If 2/3 show up, then it must be state 1 and change is enacted. There is no use for a counter-protest.
- If 1/3 show up, it could be either state 2 or 4. After the counter-protest:

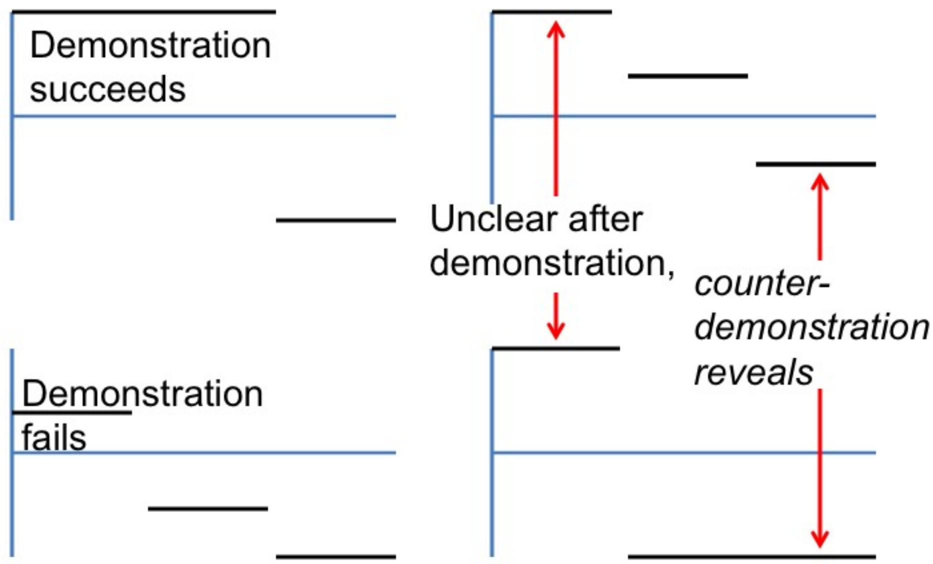


Figure 13: The counter-protests are needed in the two states on the right in which there is a middle level of turnout at the original protest. That makes it clear that it is one of the two states on the right, but does not distinguish the state. The counter-protest then reveals whether there is a large support for no change, and so distinguishes the two states. The combination of the protest and counter-protest fully distinguishes among the four states.

  - If 2/3 show up to counter-protest then it must be state 4, and there is no change.
  - If 0 show up to counter-protest then it must be state 2, and change will be enacted either via a revolution or via the government.

- if 0 show up, then it must be state 3 and no change is enacted. There is no use for a counter-protest.

The incentives for the groups to protest or counter-protest are clear:

The first group, whenever it has strong preferences would like to protest since it has a two-thirds chance of eventual success. The necessary and sufficient condition for it to want to protest in equilibrium is that $\theta_{\mathcal{CS}} > C/2$.

The third group clearly wants to counter-protest they are $\mathcal{NS}$ types, since they know it is state 4 and they will be successful. They do not want to counter-protest when they are $\mathcal{NW}$ types since then they know it is state 2 and they will fail.

The second group always gets its most preferred outcome by showing up to a protest or counter-protest when they are strong but not weak, and so they have no reason to change their strategy.

This example shows how counter-protests can reveal a state and be useful in learning the state.

Note also that the example is fully symmetric - which group holds the first protest and which counter-protests could also be reversed. In this case, since a natural status-quo is no change, it seems more natural to have the group supporting change be the one to hold the first protest and to bear some risk in doing so. But the example works in either way.

4 Concluding remarks

Summary We have provided a very tractable model that serves as a basis for the investigation of how information and learning affect the possibility of having successful revolutions or collective action.

We have shown four ways in which information can be either enabling or disruptive: (i) by encouraging some but discouraging others from participating, (ii) in settings with homophily, by weakening the content of information, (iii) by gaining information about the number of extremists in a society who might replace the status quo with an undesirable policy, and (iv) by triggering counter-protests that reveal support for the status quo.

We have shown that there are non-monotonicities so that small amounts of information can actually discourage enough of the population to make success impossible. We have also shown how protests can provide important information, both within and across countries, that can help make revolts possible, and increase the likelihood of their success.

Our model is deliberately simple, which makes many intuitions very clear and allows us to analyse a number of questions within one model - providing a more holistic view of what is needed for collective action to succeed, and should provide a basis for further studies of collective action.

Testable Implications Although our results provide for different possibilities – sometimes revolts are precluded and other times they are not – those results are tied to specific parameters. Thus, the results are far from "anything can happen". For example, it is for medium levels of z that meeting others can disable/unravel a revolt. The variable z is easily interpreted: How homogeneous is the society in question? In a country in which most agents are likely to feel the same impact of the government (e.g., most are in similar economic conditions, such as in Tunisia prior to its revolution, or France prior to its, etc.), then it is much easier to sustain a revolt even in the face of small numbers of meetings of others than in a country where there many different social classes who might have diverging preferences.

The comparative statics in terms of homophily are also testable. Levels of homophily, both in networks of interactions and in which media people pay attention to, are measurable and differ across societies. For instance, the caste structure in India leads to very high levels of homophily compared to other countries, in terms of interactions, political affiliations, and the news sources that people access.

Our results generally suggest that higher levels of homogeneity and of homophily, given bad enough conditions (so high enough payoffs) will lead to revolts whereas lower levels of homogeneity or homophily could preclude a revolt. Our results also suggest that high homogeneity, bad conditions, and very high levels of communication, would make revolts easier to have, all else held equal.

Our results also suggest that revolutions or successful collective actions are more likely to happen after relatively larger protests, and less likely after small first protests or large counter-protests; again all else held equal.

As finding a data set that would allow one to test such hypotheses, especially causally, may be difficult, another option would be to first test many of our model's predictions in a laboratory setting in which agents' information, interactions, and payoffs can be observed and controlled.

In an online appendix, we also provide a few additional thoughts on the implications of the model for how a government might act, as well as other topics that can be studied in further detail in future research.

References

Acemoglu, Daron; Georgy Egorov, and Konstantin Sonin. (2012). "Dynamics and Stability of Constitutions, Coalitions, and Clubs," *The American Economic Review*, 102(4), 1446-1476.

Acemoglu, Daron; Tarek A. Hassan and Ahmed Tahoun (2014). "The Power of the Street: Evidence from Egypt's Arab Spring". Mimeo.

Aghion, Philippe; Alberto Alesina and Francesco Trebbi (2004). "Endogenous Political Institutions," *The Quarterly Journal of Economics*, 119(2), 565-611.

Aidt, Toke, Gabriel Leon, and Max Satchell (2017) "The social dynamics of riots: Evidence from the Captain Swing Riots, 183031," mimeo: University of Cambridge.

Angeletos, George-Marios; Christian Hellwig and Alessandro Pavan (2007). "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks". *Econometrica*, 75(3), 711-756.

Banerjee, Abhijit V. (1992). "A Simple Model of Herd Behavior". *The Quarterly Journal of Economics*, 107(3), 797-817.

Barbera, Salvador, and Matthew O. Jackson. (2004). "Choosing How to Choose: Self-Stable Majority Rules," *The Quarterly Journal of Economics*, 119(3), 1011-1048.

Battaglini, Marco (2016). "Public Protests and Policy Making". forthcoming in the *Quarterly Journal of Economics*.

Bikhchandani, Sushil; David Hirshleifer and Ivo Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". *The Journal of Political Economy*, 100(5), 992-1026.

Bikhchandani, Sushil; David Hirshleifer and Ivo Welch (1998). "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades". *The Journal of Economic Perspectives*, 12(3), 151-170.

Breuer, Anita; Todd Landman and Dorothea Farquhar (2015). "Social Media and protest Mobilization: Evidence from the Tunisian Revolution". *Democratization*, 22(4), 764-792.

Brummitt, Charles D.; George Barnett and Raissa M. D'Souza (2015). "Coupled Catastrophes: Sudden Shifts Cascade and Hop among Interdependent Systems". *Journal of the Royal Society Interface*, 12(112), 1-12.

Bueno de Mesquita, Ethan (2010). "Regime Change and Revolutionary Entrepreneurs". *American Political Science Review*, 104(3), 446-466.

Cantoni, Davide; David Y. Yang, Noam Yuchtman, Y. Jane Zhang, (2017). "Are Protests Games of Strategic Complements or Substitutes? Experimental Evidence from Hong Kongs Democracy Movement" mimeo.

Chen, Ying and Hulya Eraslan (2014). "Rhetoric in legislative bargaining with asymmetric information". *Theoretical Economics*, 9, 483-513.

Chwe, Michael Suk-Young (2000). "Communication and Coordination in Social Networks". *Review of Economic Studies*, 67, 1-16.

Christensen, Darin, and Francisco Garfias (2018) "Can You Hear Me Now? How Communication Technology Affects Protest and Repression," *Quarterly Journal of Political Science* 13.1, 89-117.

Duffie, Darrell, Gaston Giroux, and Gustavo Manso (2010) "Information percolation." American Economic Journal: Microeconomics 2.1: 100-111.

Edmond, Chris (2013). "Information Manipulation, Coordination, and Regime Change". *Review of Economic Studies*, 80 (4): 1422-1458

Edmond, Chris (2016). "Information Manipulation and Social Coordination", mimeo.

Egorov, Georgy and Konstantin Sonin (2017). "Elections in Non-Democracies". mimeo Northwestern University and University of Chicago.

Egorov, Georgy; Sergei Guriev and Konstantin Sonin (2009). "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data". *American Political Science Review*, 103(4), 645-668.

Farrell, Henry (2012). "The Consequences of the Internet for Politics". *Annual Review of Political Science*, 15, 35-52.

Feddersen, Timothy J. (2004) "Rational choice theory and the paradox of not voting." The Journal of Economic Perspectives 18.1, 99-112.

Feldman, Mark and Christian Gilles (1985). "An Expository Note on Individual Risk without Aggregate Uncertainty". *Journal of Economic Theory*, 35, 26-32.

Frick, Mira, Ryota Iijima, and Yuhta Ishii. (2018) "Dispersed Behavior and Perceptions in Assortative Societies." working paper. Yale University.

González-Torres, Ada and Elena Espisito (2017) "Epidemics and Conflict: Evidence from the Ebola outbreak in Western Africa," mimeo: U.C. Berkeley.

Granovetter, Mark (1978). "Threshold Models of Collective Behavior". *The American Journal of Sociology*, 83(6), 1420-1443.

Hassanpour, Navid (2017)*Leading from the periphery and network collective action*, Cambridge University Press.

Jackson, Matthew O. (2008). *Social and Economic Networks*, Princeton, NJ: Princeton University Press.

Jackson, Matthew O. (2014) "Networks in the Understanding of Economic Behaviors," *Journal of Economic Perspectives*, 28(4), 322.

Jackson, Matthew O. (2018) "The friendship paradox and systematic biases in perceptions and social norms.," *Journal of Political Economy*, forthcoming.

Jackson, Matthew O. (2019) *The Human Network*, Pantheon Books - Knopf Doubleday.

Jackson, Matthew O., Leo K. Simon, Jeroen M. Swinkels, and William R. Zame. (2002). "Communication and equilibrium in discontinuous games of incomplete information." *Econometrica*, 70(5): 1711-1740.

Judd, Kenneth L. (1985). "The Law of Large Numbers with a Continuum of IID Random Variables". *Journal of Economic Theory*, 35, 19-25.

King, Gary; Jennifer Pan and Margaret E. Roberts (2013). "How Censorship in China Allows Government Criticism but Silences Collective Expression". *American Political Science Review*, 107(2), 326-343.

Kiss, Hubert Janos, Ismael Rodriguez-Lara, and Alfonso Rosa-Garcia (2017) "Overthrowing the dictator: a game-theoretic approach to revolutions and media," *Social Choice and Welfare*, 49: 329, 329-355.

Kricheli, Ruth; Yair Livne and Beatriz Magaloni (2011). "Taking to the Streets. Theory and Evidence on Protests under Authoritarianism". Mimeo, Stanford University.

Little, Andrew T. (2012). "Elections, Fraud, and Election Monitoring in the Shadow of Revolution". *Quarterly Journal of Political Science* 7, 249-283.

Little, Andrew T. (2016). "Communication Technology and Protest," *Journal of Politics*, 78(1), 152-166.

Little, Andrew T. (2016b). "Propaganda and Credulity," mimeo.

Loeper, Antoine; Jakub Steiner, Colin Stewart (2014) "Influential Opinion Leaders," *The Economic Journal*, 124, (581), 1147-1167.

Lohmann, Susanne (1993). "A Signaling Model of Informative and Manipulative Political Action". *The American Political Science Review* 87(2), 319-333.

Lohmann, Susanne (1994a). "Information Aggregation Through Costly Political Action". *The American Economic Review*, 84(3), 518-530.

Lohmann, Susanne (1994b). "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91". *World Politics*, 47(1), 42-101.

Lohmann, Susanne (2000). "Collective Action Cascades: An Informational Rationale for the Power in Numbers". *Journal of Economic Surveys*, 14(5), 655-684 .

Luo, Zhaotian and Arturas Rozenas (2016). "Ruling the Ruling Coalition: Information Control and Autocratic Governance," mimeo, NYU.

Manacorda, Marco and Andrea Tesei (2016). "Liberation Technology: Mobile Phones and Political Mobilization in Africa". Mimeo.

McPherson, Miller; Smith-Lovin, Lynn and James M. Cook (2001). "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, 27, 415-444.

Milgrom, Paul (1981). "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380 - 391.

Myatt, David P. (2016). "A theory of protest voting." *The Economic Journal*, 1-41.

Pierskalla, Jan H. and Florian M. Hollenbach (2013). "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa". *American Political Science Review*, 107(2), 207-224.

Shadmehr, Mehdi (2015). "Extremism in Revolutionary Movements". *Games and Economic Behavior*, 94, 97-121.

Shadmehr, Mehdi and Bernhardt, Dan (2016a). "Information Aggregation and Coordination". mimeo.

Shadmehr, Mehdi and Bernhardt, Dan (2016b). "Vanguards and Revolution". mimeo.

Shadmehr, Mehdi and Bernhardt, Dan (2017). "When Can Citizen Communication Hinder Successful Revolution?" *Quarterly Journal of Political Science* 12: 301-23.

Shadmehr, Mehdi and Boleslavsky, Raphael (2016). "Institutions, Repression and the Spread of Protest". mimeo.

Song, Yangbo and Mofei Zhao "Fickle for Survival: Bayesian Persuasion by a Strategic Regime," mimeo Chinese University of Hong Kong.

Steinert-Threlkeld, Zachary C.; Delia Mocanu; Alessandro Vespignani and James Fowler (2015). "Online Social Networks and Offline Protest". *EPJ Data Science*, 4(1), 1-9.

Weinstein, Jonathan, and Muhamet Yildiz (2007). "A structure theorem for rationalizability with application to robust predictions of refinements." *Econometrica* 75(2): 365-400.

Online Appendix: Generalizations of the Model

We present a more general version of the model and an existence result.

Uncertainty

$\omega \in I\!R$ is the state of the world, which can encode information about the value of the revolt and what fraction of the population would gain from the revolt, and so forth.

There is a prior distribution over ω, denoted G - and agents do not directly observe ω.

$\theta_i \in I\!R$ is the type of agent i, which is the private information of that agent.[21]

The distribution over types depends on the state of the world and is denoted $F(\theta_i|\omega)$. We treat these as if they are independent across agents conditional upon the state, which is technically convenient but has some measurability issues that are easily handled as the limit of a finite model.[22]

We assume the standard ordering property on information:[23] conditional upon θ_i, the distribution on ω and others' types are both increasing in θ_i in the sense of strict first order stochastic dominance. Thus, higher types of an agent lead that agent to expect higher types of other agents.

Payoffs

An agent gets a value from the revolt as a function of whether it is successful or not and whether the agent participates or not. All of these payoffs can be type and state dependent,

[21] We could allow the states and types to be multidimensional and more complicated. The advantage of one dimension is that what we ultimately care about is whether an agent is sufficiently unhappy with the government would revolt. More dimensions would involve partial orders, but the story would basically be the same - some people are unhappy enough to revolt and others are not, and the agents are trying to learn about the relative fractions and potential for success.

[22] For a discussion of the issues of a continuum of agents having independent observations see Feldman and Gilles (1985) and Judd (1985). In our model, the independence is not really needed, and so a very easy way of formalizing the signals for our purposes is as follows. Uniformly at random, draw i_0 from $[0,1]$ - this will be the agent who gets the lowest signal in society. Then let $\theta_i = F^{-1}(i - i_0|\omega)$, where $F^{-1}(\cdot|\omega)$ is the inverse of $F(\theta_i|\omega)$, and we take $i - i_0$ modulo 1, so that if $i < i_0$, then we set $i - i_0 \equiv i + 1 - i_0$. So, we randomly pick an agent to have the lowest signal, and then just distribute the signals then in a nondecreasing way for the rest of the agents with higher labels, and then wrap around beginning again at 0. This results in the right distribution of types without any measurability issues and the independence of types is not needed for our results, as agents only care about the population behavior rather than any particular agent's behavior.

[23] See Milgrom (1981).

and are given by the following table.

	$Success$	$Failure$
$Participate$	$a(\theta_i,\omega)+V_i(\theta_i,\omega)$	$b(\theta_i,\omega)-C_i(\theta_i,\omega)$
$NotParticipate$	$a(\theta_i,\omega)$	$b(\theta_i,\omega)$

Here, $a(\theta_i,\omega)$ is the value that an agent gets if the revolt is successful, regardless of whether the agent participates or not, and this can depend on the agent's type and the state. Similarly, $b(\theta_i,\omega)$ is the value that an agent gets if the revolt fails, regardless of whether the agent participates or not, and this can depend on the agent's type and the state. The values, $V_i(\theta_i,\omega)$ and $C_i(\theta_i,\omega)$ then are the additional value and cost that an agent gets from participating in the revolt as a function of whether it is successful or fails.

Generally, C_i will be positive ($-C_i$ is negative), which represents the personal cost to a person of being caught in an unsuccessful revolt - for instance, being jailed, fined, executed, etc. On the other side, V_i captures the personal pleasure or pain that a person would feel from participating in a successful revolt, as discussed in the paper. Again, as discussed above, the structure of this game is similar to that in the expressive voting literature (e.g., see Feddersen (2004) for a review).

Note that this is strategically equivalent to the following payoff matrix:

	$Success$	$Failure$
$Participate$	$V_i(\theta_i,\omega)$	$-C_i(\theta_i,\omega)$
$NotParticipate$	0	0

The strategic equivalence is due to the fact that the only thing that motivates an agent to participate is the difference that they experience from participating or not, as a function of whether the revolt is successful or not.

Since V_i can already encode relevant heterogeneity in the population via θ_i, from a strategic perspective only V_i/C_i matters and so it is without loss of generality for the strategic analysis to normalize the model so that $C_i = C > 0$ for all i. We still keep C as a variable, as we wish to consider cases in which a government adjusts the penalties for participating in a failed revolt.

We presume that V_i is symmetric across agents - depending on their identity only via their type and thus drop the subscript i. We take V be nondecreasing in θ_i,ω, and increasing in at least one of the two arguments.

Thus, we consider games of the form:

	$Success$	$Failure$
$Participate$	$V(\theta_i,\omega)$	$-C$
$NotParticipate$	0	0

Let us mention two canonical cases:

(Correlated) Private Values

One case of interest is that of "private-values" so that $V(\theta_i, \omega)$ depends only on θ_i. In this case it is without loss of generality (adjusting distributions) to set $V(\theta_i, \omega) = \theta_i$, and so payoffs are

	Success	*Failure*
Participate	θ_i	$-C$
NotParticipate	0	0

An interpretation of this case is that each citizen knows how unhappy he or she is with the government - which is the θ_i. Here, the state of the world ω captures how unhappy the overall population is via the distribution of θ_i's. Agents, via Bayes' rule, can infer how unhappy the rest of the world is by inference given that higher states, ω's, lead to a higher distribution over θ_i's. So, if an agent is very unhappy, then she infers that it is likely that ω is high and so it is then likely that other agents are unhappy too.

Common Values

Another case of interest is where $V(\theta_i, \omega)$ depends only on ω. In this case, if preferences are symmetric, then it is without loss of generality (adjusting distributions) to set $V(\theta_i, \omega) = \omega$, and so payoffs are

	Success	*Failure*
Participate	ω	$-C$
NotParticipate	0	0

This case is one in which agents do not really know whether they would like to have a successful revolt – that is governed by a state ω. For instance, agents might not know how competent or corrupt the government really is, or what might replace it. Each agent has a signal θ_i which is some noisy information about the state, and so they must infer ω via Bayes' rule from their own types.

For our purposes, it is not really important which formulation we use as they all have similar effects: agents with higher θ_i's are more optimistic that there is a high payoff from participation and that other agents feel the same. So, they all have the same basic structure of equilibria: agents with types or signals (θ_is) above some threshold participate, and others do not. Thus, we first state that general result, and then we specialize to the model with private values, for a clean and intuitive analysis.

Strategies and Best Responses

A strategy for player i is a function $\sigma_i : I\!R \to \Delta(\{0, 1\})$, which specifies a probability of participating, $\sigma_i(\theta_i) \in [0, 1]$, as a (Lebesgue measurable) function of an agent's type. Let σ

denote the profile of strategies.[24]

Let $p_\sigma(\theta_i)$ denote i's beliefs that at least a fraction q of the other agents will participate, conditional on other players playing according to σ and the agent seeing θ_i.

Given the continuum, an agent is never pivotal in determining whether there is a fraction of at least q of the population who participate, and so this is a straightforward calculation.

The expected payoff to participation is then

$$p_\sigma(\theta_i)E[V(\theta_i,\omega)|\theta_i] - (1-p_\sigma(\theta_i))C,$$

and the payoff from non-participation is 0, and so it is a best response to participate if and only if

$$\frac{E[V(\theta_i,\omega)|\theta_i]}{C} \geq \frac{1-p_\sigma(\theta_i)}{p_\sigma(\theta_i)} \quad \text{or, equivalently} \quad p_\sigma(\theta_i) \geq \frac{C}{E[V(\theta_i,\omega)|\theta_i]+C}. \tag{6}$$

Note that, given the ordering of types and preferences, $\frac{E[V(\theta_i,\omega)|\theta_i]}{C}$ is strictly increasing in θ_i.

Existence

As this is a coordination game, equilibria exist and in fact there are generally multiple equilibria. For instance, nobody participating is always a strict equilibrium: if none of the other agents participate then the revolt will surely fail and so it is a best response not to participate. However, in many cases there also exist participatory equilibria.

These games have equilibria in which agents play monotone strategies: their probability of participating is non-decreasing in θ_i, and for the cases that we examine those will be the only equilibria. Generally, given the increasing preferences and ordering on information, such equilibria always exist. Nonetheless, for some special cases there do exist other equilibria, although for generic distributions these will be the only equilibria.[25]

Proposition 4 *Equilibria exist, and in fact, symmetric and monotone equilibria exist. Each monotone equilibrium can be described by a single threshold t (the same for all agents), such that an agent participates if $\theta_i > t$ and not if $\theta_i < t$. Monotone equilibria are all symmetric up to the possible mixing that occurs at t. Monotone equilibria can be ordered by their thresholds, with ∞ always being an equilibrium threshold.*

[24] We work with strategies that are also Lebesgue measurable as a function of the agents' labels. Generally, the equilibria will naturally depend only on agents' types and not their labels, and so this is not really a restriction.

[25] For an example of a non-monotone equilibrium consider a common values setting with $\omega = 2, 3$ with equal probability and $C = 1$; and such that $\theta_i = \omega$ so that all agents know the state. In this case, regardless of q, there is always a 'best' equilibrium in which all agents participate, and there is a worst equilibrium in which no agents participate, in either state. However, there is also a non-monotone equilibrium in which all agents participate if $\theta_i = \omega = 2$ and none participate if $\theta_i = \omega = 3$.

This follows from an application of Tarski's fixed point theorem, which establishes that equilibria form a complete lattice, which here is just ordered in terms of the thresholds. Given that the proof is standard, we omit it. The symmetry is implied by the continuum of agents who have the same priors, and the fact that payoffs are monotone in types and states, so that higher types lead have higher expected payoffs from participation conditional on success.

So, we can represent monotone equilibria by thresholds t, such that an agent participates if $\theta_i > t$ and not if $\theta_i < t$. In cases with atoms in the distribution it is possible to have mixing at t.

Equilibria for Continuous Distributions Consider a canonical case in which θ_i distributed with mean ω plus some noise ε_i, where ε_i is distributed according to H (so, $F_\omega(\theta) = H(\theta - \omega)$).

$$\theta_i = \omega + \varepsilon_i.$$

In this case, the probability of success is $1 - G(t - H^{-1}(1-q))$. This follows since it must be that the fraction of people with $\omega + \varepsilon_i$ below t less than $1-q$. So, $H(t-\omega)$ must be at most $1-q$, and so ω must at least $t - H^{-1}(1-q)$. The probability of that is $1 - G(t - H^{-1}(1-q))$.

Thus, in the case of private values an equilibrium t satisfies (assuming no atoms in the distributions and an interior t):

$$t = \frac{G(t - H^{-1}(1-q)|\theta_i = t)}{1 - G(t - H^{-1}(1-q)|\theta_i = t)}.$$

Note that by Bayes' rule, if H and G have densities, h and g, then

$$G(\omega'|\theta) = \frac{\int_{-\infty}^{\omega'} h(\theta - \omega) g(\omega) d\omega}{\int_{-\infty}^{\infty} h(\theta - \omega) g(\omega) d\omega}.$$

Then a common values equilibrium is characterized by

$$\int \omega' dG(\omega'|t) = \frac{G(t - H^{-1}(1-q)|\theta_i = t)}{1 - G(t - H^{-1}(1-q)|\theta_i = t)}.$$

Some Further Thoughts and Comments

We close with a few additional thoughts on the implications of the model for how a government might act, as well as other topics that can be studied in further detail in future research.

4.1 Other Actions by Governments

A government can change the world from being one in which there is an equilibrium with a revolt to one in which there is not, by affecting the various parameters.[26,27] This presumes that the government would like to avoid a revolution and keep the status quo.

Let us examine some of those behaviors.

Costs Most directly, by increasing the cost to failed revolutionaries (increasing C), the government can make the conditions for a revolt harder to satisfy. For instance, in the base model, it is sufficient to raise C to a point at which

$$\theta_H/C < \frac{(1-\pi)(1-z)}{\pi z}$$

to avoid the revolution. Correspondingly, there are values of C that prevent revolution for different levels of information.[28]

Information Control, Seclusion, and Homophily The government can also suppress and censor information. As we saw, having only a few meetings with others, or if those meetings are mostly with own type then this can lessen the chance that people have to learn about the number of others who support change. By limiting information flows, especially across groups or geography, so that most interactions are limited and local, one could shift an equilibrium to preclude a revolt. As we have seen however, it could also work the other way in cases in which the prior beliefs are strong enough – by encouraging information exchange one could end up undercutting the support for a revolt and preclude it. Which policy a government would want to undertake would depend on the information structure.[29]

Our results on homophily also suggests that a revolutionary group might want to seclude its members. By allowing its members to possibly meet others who do not support the revolt, the group risks having its members doubt the possibility of success which could disrupt the revolt.[30]

Propaganda and Fake News The government could also bias information via propaganda.[31] Propaganda is interesting in that it does not have to convince all of the potential

[26]For important analyses of governments and propaganda as well as censoring and other informational distortions in models that are very different from ours, see Edmond (2013) as well as Egorov, Guriev, and Sonin (2009), Little (2012), and King, Pan and Roberts (2013).

[27]Events beyond a government's control that uncover its weaknesses can also change conditions, enhancing the possibility of a revolution - for instance, see González-Torres, Ada and Elena Espisito's (2017) discussion of Ebola and social unrest.

[28]For a model of repression, see Shadmehr and Boleslavsky (2016).

[29]See Luo and Rozenas (2016) for more discussion of informational control by a government.

[30] This applies quite generally, and military forces and paramilitary groups are at times discouraged from interacting with populations that might raise doubts about their mission or the support for it.

[31]For different views of information manipulation in the face of social coordination, see Edmond (2016), Little (2016b), and Song and Zhao (2018).

revolutionaries that revolt is a bad idea or that the state is Low, but instead it just needs to convince enough of them so that the remaining types know that they will no longer have sufficient numbers to be successful. For instance, if more than $z - q$ of the potential revolutionaries are convinced by the propaganda, then the revolt cannot succeed, regardless of whether the remaining H types are convinced or not.

Thus, propaganda can be disruptive even if it only convinces a small subset of the population that they should not take part in a revolt. This could happen by convincing people that they stand no chance of success, for instance, by inflating the estimates of how many θ_L types there are in the population; or by convincing people that they are better off than they are, or better off than what would happen after a revolution, etc.

Noisy news sources can have the same effect as propaganda, effectively lowering the confidence that individuals have in the information that they receive as well as increasing the chance that some others may be discouraged. It can work in the same way as overt propaganda in that if people are concerned that a small subset of supporters of change may be discouraged, then that can unravel the revolt. Of course, it could work in reverse, convincing a small group that might not otherwise participate which then enables a revolt.

Redistribution Finally, the government could also redistribute resources. Again, the government does not have to redistribute resources to all of the potential revolutionaries, they simply need to buy enough of them off to discourage the rest - so they just need to please $z - q$ of the H types. They can produce some very unhappy parts of the population, provided that they make the middle range sufficiently happy that they will no longer revolt.

Specifically, suppose that redistribution by the government is observable and that the government knows the state (so it knows the condition of the whole population). Thus, whenever the government does redistribute income, then the population knows it is the High state. So, it is clear that in that case they must pay at least θ_H to a fraction $z - q$ to avoid the revolt. The equilibrium must be one in mixed strategies. To see this note that if it were a pure strategy equilibrium, then it would be one in which the government only redistributed in the High state. But then when seeing no redistribution, agents would infer it is the Low state and not revolt. In that case, the government would not need to redistribute in order to avoid the revolt. Thus, the redistribution must be in mixed strategies. In order for this to make sense with a continuum of agents, we then allow agents to correlate their strategies, so that H types revolt with some probability p when not seeing redistribution. The probability of redistribution is then just enough to make agents indifferent conditional on seeing no redistribution, and the probability of revolt is just enough to keep the government indifferent between being overthrown and paying the redistribution.

4.2 Other topics

We have focused on the coordination issues and the role of information. There can also be public-good aspects and free-riding behavior in protests and revolutions that we have not modeled here and could be interesting to combine with the coordination issue.[32]

We have provided our analysis in the context of correlated private values, but a similar analysis applies to more general affiliated and common value settings. The analog of homophily is still that people with similar types are likely to meet each other – so that people who are in close contact are likely to be getting similar information, and thus do not learn as much from meeting each other as meeting a uniformly random draw from the population.

Our focus in terms of learning has been on the population learning how many others are willing to support a revolt. In some instances, it might be that the key learning that goes on is whether sufficient numbers of the military and/or police would also support a revolt, and whether they would fire upon the population. This could easily be added to the model by having different roles in the population, and could explain why some governments try to isolate some of their key military from the general population.

In our analysis we have taken the meeting and homophily structures as exogenous. Control of social media by a government, as well as rules that limit internal movement, could be used to control the interaction structure within a society, and could be interesting to explore as another extension of the model.

In our model agents have been Bayesians who understand the full model. One could also suppose that people fail to realize that their sample is biased. For instance, if they neglect homophily, then this could increase support for a revolt since people fail to realize that the population is more heterogeneous then the people that they meet. For a more on various sorts of correlation neglect see Frick, Iijima, and Ishii (2018) and Jackson (2018,2019).

Finally, the feedback between politics and protetsts is something that is deserving of much more study. This can fit into a more general study of the endogeneity of governments.[33]

[32]For an interesting paper on free-riding in protests, see Cantoni, Yang, Yuchtman, and Zhang (2017).

[33]E.g., see Aghion, Alesina, Trebbi (2004), Barbera and Jackson (2004), Acemoglu, Egorov, Sonin (2012), and Egorov and Sonin (2017).