The infrastructure of revolts: internet, game theory and complex networks in the Arab Spring

Brais Álvarez-Pereira[1] and Martín Portos-García[2]

February 24, 2014

Abstract

Combining game theory and complex networks, we study the flows of controversial political information in non-democracies. Our model explains how, under certain circumstances, by facilitating the communication between potential protesters the internet can mean the difference between massive protests happening or not happening. The key finding of our study is that a relatively small increase in the proportion of the population having internet access might imply a big difference in the resulting equilibrium, by making it feasible for challengers to estimate their potential support, a requirement for mass protests to take place. As the perceived support for a potential protest increases, more agents will find optimal joining it. This model is particularly oriented to seek an explanation for the fast spread of the self-organized protests that have recently shaken the Arab world.

Keywords: Internet, social networks, game theory, social movements, Arab Spring.

1, Department of Economics, European University Institute.
2, Department of Political and Social Sciences, European University Institute.
Contact: brais.alvarez@eui.eu, martin.portos@eui.eu.

1 Introduction.

1.1 Thinking collective action rationally?

Although embedded to some extent within political process, traditional rational choice accounts are far from mainstream in the study of contentious politics. With some remarkable exceptions (e.g. Marwell and Oliver 1993; Oberschall 1994; Lichbach 1995; Opp 2009), they have rather played a secondary role. Many scholars have cast rational choice insights aside too readily, which may have thrown the baby out with the bath water. Moreover, these criticisms have often been based on the narrower versions of rational choice models, which build on Mancur Olson's legacy (Finkel 2008; Opp 2013), when not simply on outright clichés associated to rational choice theory (RCT, hereafter). Unlike structuralist or culturalist metanarratives, RCTs place individuals' rational choices and actions at their core. This is precisely their virtue: behavioural theories of rational choice provide mechanisms linking structural dynamics and individuals solving social dilemmas – i.e. agency (see Ostrom 1998:2). This is not to argue that only RCTs monopolize the supply of bridgers. For instance, many causal mechanisms-based contributions have proved –and still proof – useful in overcoming the structure-action binomial. However, we argue, RCT can be helpful in this regard in many cases, too.

The Olsonian logic of individual (non)-involvement in collective action rests upon two axes (Olson 1965): individuals' participation will have insignificant marginal impact on the collective enterprise being successful and collective goods (e.g. regime change) are normally non-excludable. As protesting is costly and risky, rational individuals who respond to incentives and stimuli tend to not participate in collective actions (i.e. they free-ride). Dealing with the free-rider dilemma has consumed most time and effort of rational choice scholars within our field (see Lichbach 1995; Opp 2009).

The most resolute attempt to date in this regard has been the work by Opp and colleagues (Finkel et al 1989; Opp et al 1995; Opp 1989, 2009, 2013): their *wide* rational choice model of collective action (RCM, hereafter). It differs from the Olsonian logic (the *narrow* RCM) in two crucial aspects: 1. Individuals do act on their subjectively perceived influence, which does not necessarily equal zero. 2. Relevant private incentives or disincentives for individual participation could be material, but also social or moral (Finkel 2008:23-24). These principles, along with the basic propositions of RCT substantiate our approach: human behaviour is goal-directed and seeks to maximize individuals' net utility, cooperation and coordination are often required to obtain a collective good, and de facto goal-attainment depends on behavioural opportunities and constraints, or costs-benefits in rational choice terms (Opp 2013:1; Oberschall 1994:79).

More in depth, the relationahip between social beliefs and collective action is crucial throughout. Following a rational choice approach, Yang Lu et al (2013) build a global game in which they assume agents are homogeneous in preferences but receive different signals about the regime's strength, and share this information with each other when they are randomly paired. They use this model to study how different rumors might be more or less credible among the population, and how this credibility will determine the size of the mass of attackers in a revolution, which always happens but will only succeed if the amount of active revolutionaries is large enough. With a different focus, social beliefs and their tendency to consensus and wisdom are a topic widely studied, especially in economics − see for example Golub and Jackson (2010) and Acemoglu et al (2010)−, DeGroot (1973) and Olfati-Saber et al (2007) on network theory, and the social and natural sciences (Sueur et al 2012).

1.2 The state of the art: internet and the protest domain

A fruitful academic debate has emerged regarding the changes internet brings about mean for political mobilization and activism. Not in vain, digital technologies in general and social media in particular are inherent to most contentious developments nowadays, as the Arab Spring reflects.

We can distinguish two main views on the contribution of social media to the organization of protests (see Micó and Casero-Ripollés 2013:3-4). On the one hand, the *reinforcement* tradition states that the Web reinforces organization (strengthening rather than creating new ties), transnationalization and mobilization of traditional collective action (Diani 2000; Van Aelst and Walgrave 2002). On the other hand, according to the *innovation* perspective, internet generates new forms of activism and collective action, which are social media driven (Mercea 2012; Bennett and Segerberg 2013).

With regards to the organization of activism, internet and especially social networks promote the creation of communities based on shared interests, fostering collective identities and providing the infrastructure to generate a critical mass (Harlow 2013; cited in Micó and Casero-Ripollés 2013:4). In addition, especially in non-democracies, the internet is expected to decrease the influence of established media organizations (mostly controlled by the regime) on the political agenda and to open more political information to public scrutiny, which would imply lower dependence on institutional structures (Bimber 1998).

Broadly speaking, we assume that the internet sharply reduces costs for creating, organizing, and participating in protests and decreases the need for activists to be physically together in order to act together, as so do Earl and Kimport (2011). It does not mean that participation automatically increases with increasing internet access. It depends on how people use and appropriate it (Loader and Mercea 2011). As our analysis shows, virtual political activism is not restricted to online boundaries.To be sure, virtual interactions do not replace face-to-face ones (Diani 2000). However, internet access and flows of political information, we argue, have been crucial to account for the development of (protest) events in non-democratic regimes, as those that experienced the Arab Spring wave.

Why do we focus on the latter? Despite internal variability and heterogeneity across Arab countries where mobilizations took place −with varying degrees of intensity− (see Donker 2013), these protests are all together distinctive in certain regards; more importantly for our purposes: 1. They consisted of technologically mediated protest events where the internet played a key role. 2.The cross-country spread of protests followed the logic of "example modularity", based on the emulation of others prior successful example (Beissinger 2007). 3. Except from the triggering case (Tunisia), the other protests within the wave were initially driven by digital technologies, which later had an impact on the offline sphere, as we will show. 4. All of them took place in the Middle East/North Africa, which has historically been the least free region in the world (Freedom House 2013), with some key recurring and resonating claims: pro-employment, human rights protection, regime change and political freedom/elections, anti-corruption, etc.

But how have these different regimes coped with internet, online activism and the threat it may pose for regime stability? Some scholars had already stressed the importance of internet prior to the Arab Spring, for instance, to create an International network in support of Mexican Zapatistas in order to avoid large-scale repression in 1994, to overcome barriers to news flows during Milosevic's mandate in Serbian boundaries and to bring Malaysian anti-regime groups together in 1998 (Ferdinand 2000). However, contrary to more optimistic approaches, we would like to add a word of caution: internet can also be used as a tool for propaganda; moreover, it may also help monitoring and repressing online and offline political activism, especially under dictatorial regimes (Howard et al 2011).

Not in vain, internet censorship is nowadays a main element of state repression. Secret state affairs have been made public knowledge through the network, state repression of protests and conflicts in any part of the world are watched internationally in real time, facing a public opinion with increasing resources at hand. As a consequence, most authoritarian and autocratic regimes find increasingly difficult controlling the internet without totally shutting the network down, as evidenced during the Arab Spring by the Egyptian, Syrian and Libyan cases.

As aforementioned, this paper follows a rational choice approach, combining a game theoretical model (based on the n-person assurance game) with complex networks to explain how access to internet can help to improve the transmission of political information across individuals, especially in societies under authoritarian or autocratic regimes. These flows of information refer to the transmission of facts provided or learned about the political sphere. Besides, we show how under certain conditions the exchange of information through internet might contribute to trigger mass protests that challenge the political status quo, which would hardly have happened otherwise.

In section 2, we present a simple game-theoretical model based on the Assurance Game or Stag Hunt, in which cooperation pays off. Following R. Karklin and R. Petersen (1993), we study the aggregation of this game across society introducing the concept of an individual tipping point, given by a certain proportion of the population protesting, above which the corresponding agent will find it optimal to join the demonstration. Section 3 displays the process proposed to explain how internet can help to improve the communication of political information across society, and this is illustrated with several simulations on simple networks. Based on these results, in section 4 we develop the model that shows under which economic, social and political conditions we would expect internet to make a difference in affecting the individual utility of protesting. In section 5, for illustrative purposes, we apply our model to the Arab Spring. We stress some final remarks and conclude in section 6.

2 Triggering mass protests: the assurance game.

The assurance game[1] is our basic game-theoretical model to explain individual incentives to choose whether to mobilize or not. The structure of this simple game is very similar to a prisoner's dilemma, with the difference that the game has two pure-strategy Nash equilibria, including one under which the two players cooperate. The general form of this game is:

Table 1: The simple assurance game.

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$\underline{a} \setminus \underline{a}$	$c \setminus b$
	Defect	$b \setminus c$	$\underline{d} \setminus \underline{d}$

with payoffs characterized by,

$$a > b \geq d > c. \tag{1}$$

Amartya Sen (1967) shows the essential difference between these two games. In the prisoner's dilemma the best choice for the individual is always to defect (a strictly dominant strategy), even when he would like the other player to cooperate. When the game is generalized to n-players this is known as the isolation paradox, and it creates a free-rider problem which makes it necessary to enforce cooperation if we want an efficient Pareto optimal outcome.

[1]The original formulation of the assurance game comes from Rousseau (1754). See Skyrms (2004) for a study of cooperation and social action based on this game.

The assurance game has both players cooperating as the payoff dominant equilibrium, while both players defecting is the risk dominant equilibrium. This risk optimality comes from the lower variance of the possible payoffs when defecting, given uncertainty about the other player's strategy. While players are pushed towards defecting by the higher risk of cooperating, they are pushed towards cooperation by higher payoffs. In the hypothetical case in which a player knows that the other will cooperate, his or her optimal action will be to cooperate as well. This makes expectations and estimations about the other player's actions crucial to determine players' strategies.

In the context of the organization of a mass protest in an authoritarian or autocratic regime, internet might allow to improve the transmission of political information, potentially decreasing individual uncertainty about what others plan to do. Under certain conditions, this will push agents towards cooperation.

2.1 Mass protests, generalizing the assurance game to N players.

We adapt the N-players assurance game to the context of mass protests. In the original set-up defined by Sen (1967), the cooperative equilibrium holds if and only if each of the players expects all the other $(N-1)$ players to cooperate. We modify this game, using Sen's notation to formalize the approach of Karklin and Petersen[2] (1993), who introduce a tipping point into the model. So, $i(x^i)$ is defined as individual i, for $i = 1, 2, ..., n$, choosing strategy x^i, with $x^i \in \{A, B\}$, where $A =$ *protest* and $B =$ *do not protest*. ϕ^i is the utility for individual i's action given the set of $N-1$ actions chosen by the other players. Since we are studying the initial triggering of a mass protest, we assume that all the players choose their action simultaneously and so each of them will take his decision by comparing expected payoffs.

If the lack of information does not allow players to create an expectation of the potential actions of other players:

$$\phi^i(1(x^1), 2(x^2), ..., i(B), ..., n(x^n)) > \phi^i(1(x^1), 2(x^2), ..., i(A), ..., n(x^n)), \qquad (2)$$

the utility of non-demonstrating is larger, and so the risk dominant equilibrium will hold. Full uncertainty about other agents' choice makes demonstrating too risky.

When individuals have enough information as to create an expectation of other players' actions, for a given agent i:

$$\phi^i(E_i[1(x^1)], ..., i(A), ..., E_i[n(x^n)]) > \phi^i(E_i[1(x^1)], ..., i(B), ..., E_i[n(x^n)])$$

[2]They use this game to study the triggering of mass protests in Eastern Europe during 1989.

$$\iff \sum_{j=1}^{n} E_i \left[(j(x^j) = j(A))\right] \geq X_i, \tag{3}$$

an individual will find it optimal to choose $x^i = A$ if and only if he expects the number of other players also choosing to protest to be higher than a certain individual critical value X_i. The value of this threshold will depend on the utility the individual expects from both taking part and not taking part in the protest.

The perceived utility derived from protesting and the perceived utility derived from not protesting, are respectively increasing and decreasing functions of the number of people mobilizing. It is obvious that for a single individual to demonstrate alone under an authoritarian or autocratic regime is highly risky, while this risk decreases significantly if he is in the middle of a very large protest. Hence, we assume that utility increases (monotonically) as the proportion of people protesting increases. Denoting by X the total number of people participating in a protest,[3] this is interpreted as,

$$\frac{d\phi^i(i(A), X)}{dX} > 0, \tag{4}$$

for all $X \in \{1, 2, ..., N\}$.

As Karklin and Petersen (1993) argue, the utility derived from not protesting is assumed to be decreasing on the number of people mobilized, because not participating in the protest can be interpreted as to be a supporter of the current status quo. This can imply important social pressures and fear of future sanctions if the amount of people demonstrating is high. The effect is expected to be particularly strong in contexts in which the regime is seen as close to collapse. Then,

$$\frac{d\phi^i(i(B), X)}{dX} < 0, \tag{5}$$

for all $X \in \{1, 2, ..., N\}$, the utility derived from not protesting is monotonically decreasing on the amount of people protesting.

Figure 1 reproduces the graphic shown by Karklin and Peterson (1993), incorporating the individual critical point X_i.

X_i is determined by the position of these two utility curves. But what moves the lines in the diagrams? We follow a general micro-macro explanatory framework, adapted from Opp (2013; see figure 2).

[3]The term X allows us to write $\phi^i(1(x^1), ..., i(x^i), ..., n(x^n))$ simply as $\phi^i(i(x^i), X)$.

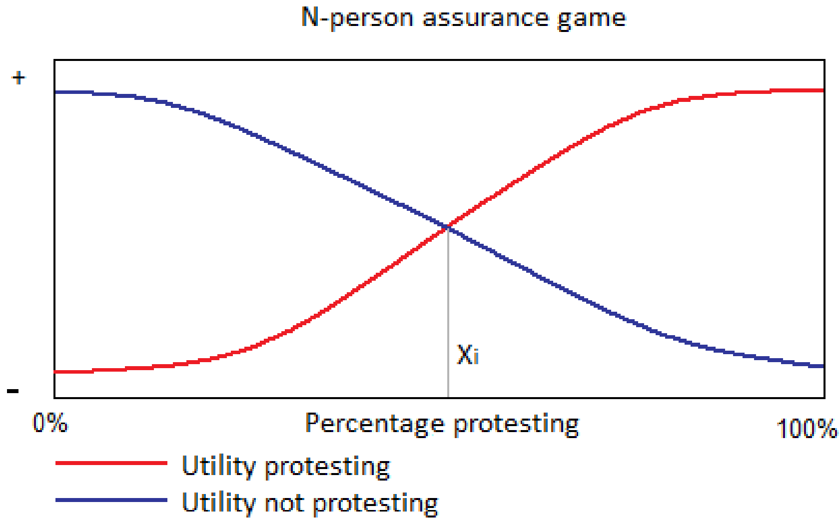

Figure 1: Simple N-person assurance game

The specific factors (incentives and disincentives) that affect the value of X_i are:

1. Changes in the subjective valuation of the situation under the current regime. An increase in food prices, the lost of a job, expropriation of personal property... Any change perceived as caused by the government and/or the regime at large that worsens the situation of the individual would lower the utility of not demonstrating, decreasing X_i. Outside RC accounts, these structural shocks and the consequences and perceptions they engender are often referred to as grievances (see Büchler 2004).

2. As aforementioned, many episodes of contention tend to be modular, which is a form of cross-group spread of collective action based on the emulation of others' prior successful example (Beissinger 2007). For example, protests succeeding in neighboring or similar regimes when the changes they are expected to bring are aligned with the individual's interests (i.e. when the claims resonate with the individual's) are important in pushing the utility curve for demonstrating upwards.

3. Norms. Breaking the rules in non-democracies is expected to derive in more or less severe punishment. For instance, individual's beliefs about the authorities' intentions to not repress certain events and actions will tend to move the utility curve for demonstrating up, lowering the critical value of X_i.

4. Perceived individual efficacy. Protests as a result of strong discontent are only likely to exist if they are expected to reduce these grievances. Thereby,

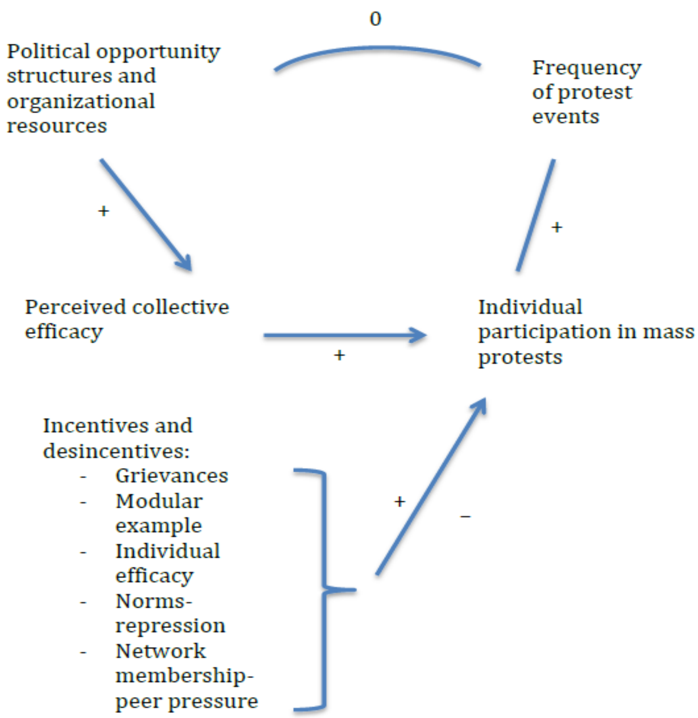

Figure 2: Example of a Micro-Macro explanation. Adapted from Opp (2013)

grievances matter for as long as people perceive they are (individually) influential. If the perception of individual influence is high, it increases the utility of protesting.

5. A further element (crucial in our model) that would affect the individual utility is the behaviour of friends or close neighbors, close ties in the network, as we would expect an individual to weight more the choice and social pressures coming from those people that are closer to him (this is called "peer pressure" throughout).

6. Macro-political structures (e.g. perceived openness of the system via availability of allies and divisions within elites) and other organizational resources (if time, location and an estimation of the number of demonstrators for a given protest are easier known) are expected to affect the individual utility of demonstrating. However, this effect is mediated by collective efficacy, which refers to perceiving joint action as efficacious insofar as helping to redress the situation. Although individual (not collective) perceived efficacy is what has a direct effect on individual participation in mass protests, it would have been dissonant to consider the latter does not play any role (Opp 2013).

The utility of protesting and not protesting $\phi^i(i(A), X)$ and $\phi^i(i(B), X)$ being monotonic on the number of people participating in the protest X, guarantee that if these two utility curves cross for individual i, they will do it only once. For every individual i for which they do in fact cross, i.e. for every individual who is willing to mobilize for some X, this implies that anything that pushes the utility curve for protesting upwards, will take the critical point X_i to a lower value, X_i', see Figure 3.

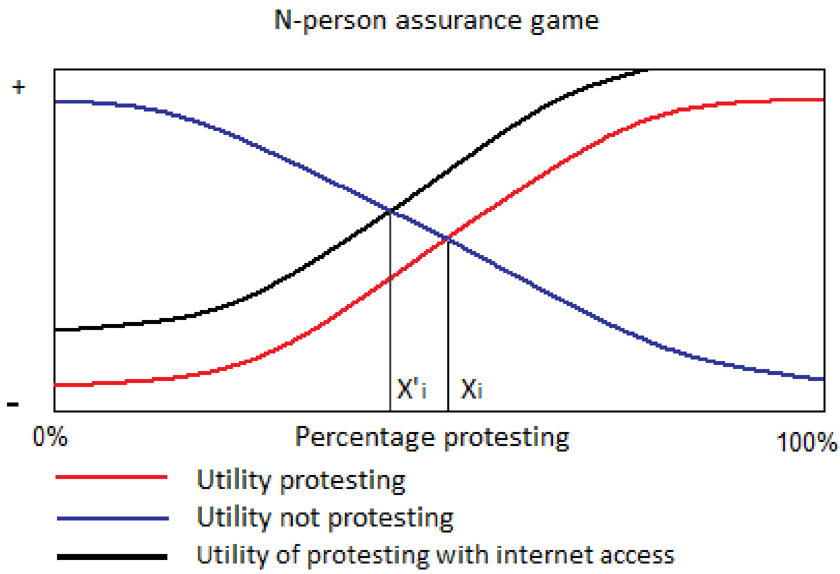

Figure 3: N-person assurance game with internet access.

A critical issue which to our knowledge has not been contemplated yet, is that even if a large proportion of the population has a relatively low critical value of X_i above which they would protest, they still need to be able to create an expectation of their number in order to be able to coordinate themselves. This is not always easy in countries under authoritarian or autocratic regimes.

The central aim of our analysis is to characterize the way in which reaching this expectation might or not be possible, and the conditions under which internet might facilitate the process.

3 Internet and the creation of small worlds.

We now assume that a significant part of the population in a given country under an authoritarian or autocratic regime would like to protest against the status quo, given their current economic situation and political environment. Formally,

$$\exists\, X : \sum_{i=1}^{n} i \ : \left\{ \phi^i(i(A), E_i(X)) > \phi^i(i(B), E_i(X)) \iff E_i(X) \geq X \right\} \geq X. \quad (6)$$

The number of individuals who would choose to protest if they knew that the number of protesters would be higher than X is larger than this critical amount. In this case, at least X individuals should coordinate in a protest.

We can easily think of situations under which this condition being satisfied is not enough for a mass protest to be successfully organized. Under authoritarian and autocratic regimes, the governmental control of communication media and the usual persecution of organized political opposition makes it often difficult for people to coordinate themselves in extra-institutional actions. Threat of more or less severe punishment frequently implies that individuals who trust each other enough to promote an action opposed to the interests of the government or regime tend to be socially close to each other. This might make exchanges of political information across different social groups relatively scarce, complicating or fully preventing the creation of an accurate expectation of potential supporters for a mass protest, a requirement for mass contentious performances to happen.

We take an undirected, regular ring lattice network as an appropriate tool to describe the flow of (potentially subversive) political information across a society in which this is exchanged only between individuals who are socially close to each other. In this lattice graph $L(N, E)$, a link is established between every two nodes within a lattice distance $k \in \mathbb{N}$ from each other. The node degree[4] of the network is given by $z = 2k$, and so for a given z, each node is connected to $z/2$ neighbors on each side. Figure 4 shows a small lattice with 12 nodes and degree $z = 4$.

For $k = 1$ the network coincides with the underlying lattice and for higher k the neighborhood[5] of each node i spans a larger fraction of the nodes closer to it. In the figure above we can see that for degree $z = 4$ each node is directly linked to those nodes at a lattice distance $k \leq 2$.

We argue that this kind of networks can be used to illustrate the characteristics of communication of political information across societies in which this happens mainly between individuals who are socially close, i.e, who belong to the same family, clan, group of friends or work together. Assuming information

[4]For a node i, its node degree z_i is the number of nodes with which it is directly linked.

[5]z^i, the set of neighbors, for a given node i. A node j is a neighbor of node i if they are directly connect, this is, if ij belongs to the set of links E.

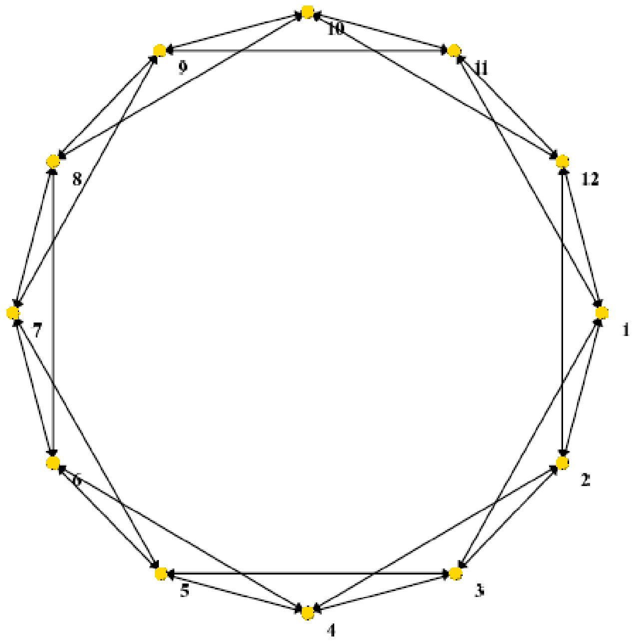

Figure 4: Lattice network, 12 nodes, z = 4.

needs time to flow through the network, long distances will make this process too slow, and it will be difficult − when not impossible− for potential protesters to create an expectation of their number. Also, the perception of little movement in favour of protesting, will make individuals less willing to declare themselves as part of the group of potential protesters. These two factors together will make it very difficult for the discontents to coordinate in a mass protest.

How slow information will flow in a regular ring lattice is easy to understand when we study the characteristic distances for this type of networks. With an average degree $z = 2k$, the diameter[6] of a regular lattice network is given by

$$\hat{d} = \frac{n-1}{2k} = \frac{n-1}{z},$$

and the average distance, or average path length[7] by,

$$\bar{d} = \frac{1}{2}\left(1 + \frac{n-1}{2}\right) = \frac{1}{2} + \frac{n-1}{2z}.$$

We can see that both $\hat{d}$ and $\bar{d}$ increase linearly with the number of nodes. This implies that for a relatively small social network of 1000 nodes and degree $z = 14$, we will have $\hat{d} \simeq 71$ and $\bar{d} \simeq 36$. Information leaving one node will in

[6]The maximum distance between any pair of nodes.

[7]The average number of steps along the shortest paths for all pairs of nodes in the network.

average need go through other 36 nodes before reaching another randomly chosen node in the network.

This accounts for the fact that in a society with serious communication constraints, the number of links that information needs to go through may be too big for the population to be able to build a correct estimation of the distribution of certain qualities or attitudes existing among themselves. Given the persecution of information contrary to the government's interests, knowing the proportion of the population that would support a mass protest should be particularly problematic.

3.1 The Watts and Strogatz Small Worlds Model.

Our model to study how internet might reduce the difficulty of coordinating a mass protest follows an structure similar to the one exposed by Watts and Strogatz (1998).[8] Their model departs from a basic one-dimensional lattice network as the one defined in the previous section, with $z = 2k$. Every link connecting each of the nodes $i = 1, 2, ..., n$ to any $j \in z^i$ breaks with a probability $\mu \in [0, 1]$, and randomly connects the corresponding i to any other link in the network. We call this process *rewiring a link*, or finding a new connection.

In our model we consider a link is rewired when an individual, who decides to look for political information farther away in the network, succeeds in his attempt of finding another agent in favor of protesting. If this happens, both of them will inform their friends (neighbors) about this new connection.

Importantly, together with the fast decrease in network distance caused by the first links being rewired, the risk of being caught by authorities,[9] very high for the first individuals beginning the process, decreases sharply as their number increases. This creates a very strong strategic complementarity, as defined by Topkis (1998), between player's strategies of trying to find new connections, which will happen to be crucial for the implications of our model.

Figure 5 is created from a regular lattice such as Figure 4, including a rewiring probability of 0.2 for every link. We can intuitively see that this relatively small change in the network (8 out of 50 links) has altered its properties significantly. Table 2 shows to which extent this is the case, comparing the value of our main parameters of interest for different rewiring probabilities, in a network with $z = 6$ and 300 nodes[10,11].

[8]For a review of network theory and concepts, see Vega-Redondo (2007) and Estrada(2011).

[9]This is related to the betweenness coefficient of each node, which can be interpreted as a measure of the proportion of information in the network that goes through that agent.

[10]Note we fulfill the condition given by Watts and Strogatz (1998): $n \gg z \gg ln(n) \gg 1$ to ensure that the network is connected.

[11]For a better layout we do not display here the results for $p \geq 0.7$. The change in the parameters after this probability is relatively small, as can be seen in the graphics.

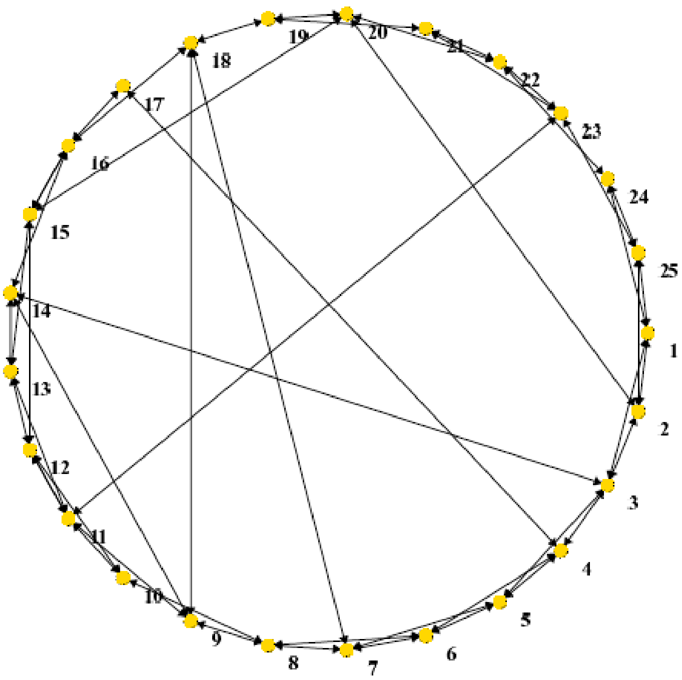

Figure 5: Small World, 25 nodes, $\mu = 0.2$.

Rew. Prob.	0	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6
Av. Distance.	25.41	8.85	6.29	4.64	4.11	3.85	3.59	3.5	3.46
Diameter	50	22	12	8	8	7	6	6	6
GCLC	0.3	0.60	0.55	0.45	0.38	0.31	0.22	0.19	0.16
GBC	0	0.45	0.21	0.13	0.06	0.05	0.03	0.043	0.04

Table 2: Parameter estimates for a Small World with n=300 and z=6

GCLC stands for 'Graph Clustering Coefficient', which takes a value of 0 when there are not cliques[12], and the value of 1 when each node and its neighborhood are complete cliques. The role of clustering is important to study communication in social networks, as it describes how typically my friends are friends among themselves. With a higher clustering coefficient, information will tend to be more redundant, since it is likely that the information a node gets from a neighbor has already been obtained −or it is very similar to that already obtained− from another neighbor, and so news and innovations tend to flow slower.

GBC stands for 'Graph Betweenness Coefficient'[13] which takes the value of 0 when all the nodes have the same betweenness coefficient and the value 1 when one node falls on all paths between the remaining $(N-1)$ nodes, as it is the case for a star graph.

With regards to the Watts and Strogatz algorithm of network formation, the evolution of these parameters as we change μ is qualitatively equivalent for networks with different length and degree. This is important because it allows us to assume that the evolution of our parameters of interest will be asymptotically well-behaved as $N \to \infty$ and so our conclusions hold also for networks with a very large number of nodes. The networks resulting from this algorithm have two important properties observed in real world social networks, a relatively high clustering coefficient and small-world properties − see for example Amaral et al (2000).

Figure 6 shows the fast decay of the average distance and the diameter of the network as soon as we introduce a relatively small rewiring probability μ, and how after this initial decay the parameters remain relatively stable as we increase this probability.

This huge decrease in network distance as we slightly increase μ is expected to have a huge impact on the flow of information through society. We can check this carrying a simple computational simulation. Using the Watts and Strogatz algorithm we create two networks with 300 nodes each, respectively for the cases $\mu = 0$ and $\mu = 0.1$, and assign to each of their nodes an initial belief $m_i \sim U[0,1]$.[14] Then, following Estrada (2012), we build an standard algorithm which makes each agent weight his neighbors' opinions with his own in each iteration, following a process similar to the DeGroot model, see Degroot (1974). The result is shown in Figure 7. We can appreciate how the change in the rewiring probability has enormously reduced the consensus time, the time it takes for the beliefs of all the agents in the network to converge to the same value.

Figures (6) and (7), and Table 2, make it clear that the first nodes rewiring some of their links have a crucial contribution improving the information flow. Why if this is so clear, those who are more unsatisfied with the status quo do not

[12]A subset of a graph which has every two nodes directly connected by a link.

[13]Wasserman and Faust (1994), formula 5.13, p. 192

[14]The beliefs are uniformly distributed between 0 and 1, so assuming there is a true state of the world $m = 0.5$, all agents are initially biased.

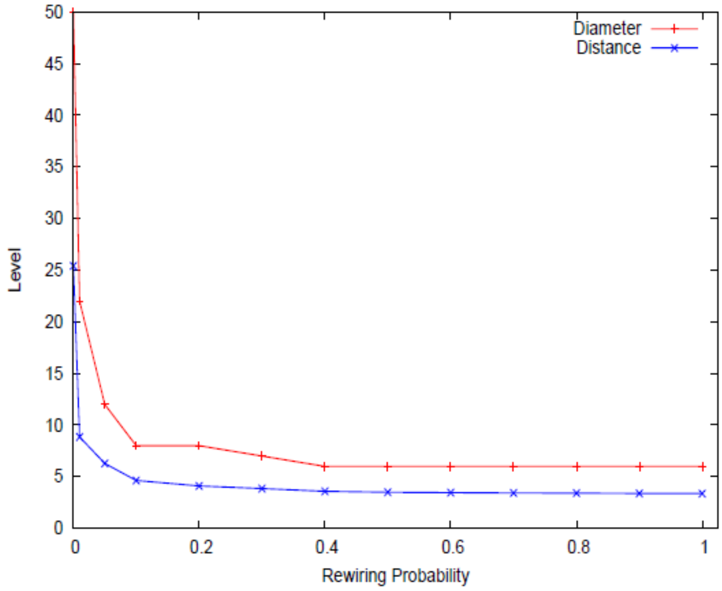

Figure 6: WS 300 nodes, k=6. Average Distance and Graph Diameter

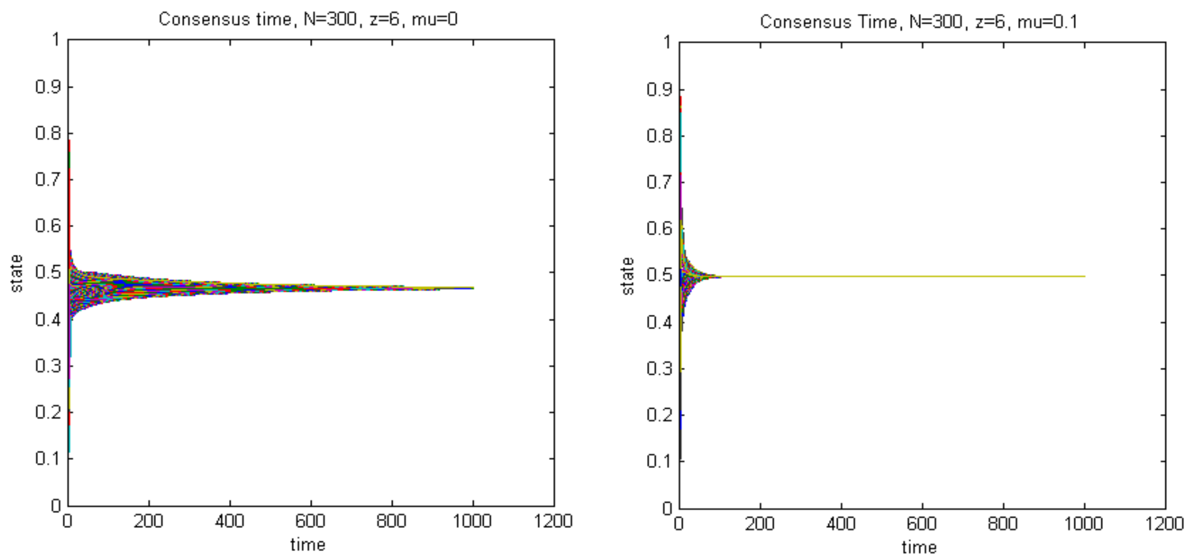

Figure 7: Consensus time for two networks with 300 nodes, degree 6 and rewiring probability μ respectively 0 and 0.1.

just begin to contact different individuals across society? The next graphic and the concept of betweenness gives us the answer.

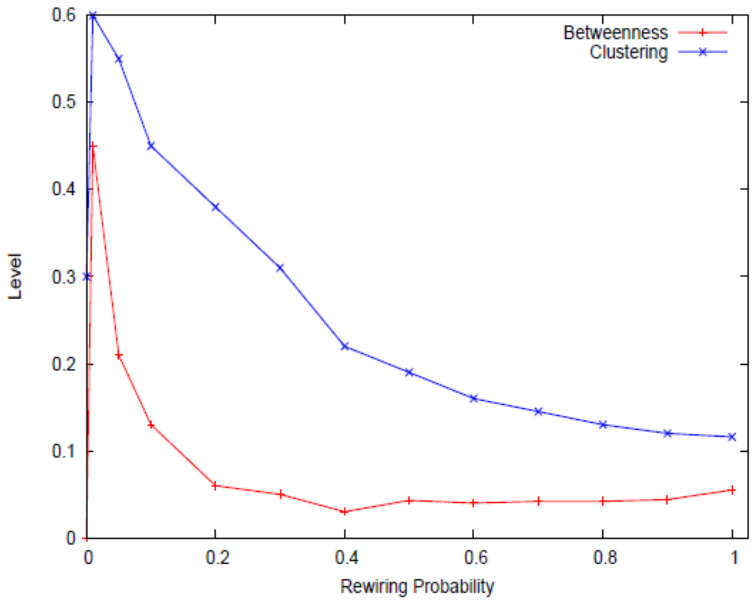

Figure 8: WS 300 nodes, k=6. Graph Betweenness and Clustering Coefficient.

Figure 8 makes clear the intuitive idea that the first nodes establishing a new connection bear a higher risk of being found and punished by authorities. A lot of information is going through them, what makes them highly visible. The fact that the Coefficient of Graph Betweenness goes from 0 to 0.45 with the change of μ from 0 to 0.01 gives us an idea of their visibility and the high level of risk this might imply under an authoritarian regime which prosecutes any kind of opposition[15]. In our model this larger risk will be taken by those who have a higher interest in the mass protest taking place, i.e. they dislike very much their current, passive situation. This gives them a stronger incentive to explore the feasibility of a mass protest happening, by trying to find out what is the state of opinion far from their neighborhood.

In the next section we study the incentives which might push those first individuals to take the risk of looking for new information, and how their degree of success will determine if the process of coordinating a significant part of the population in a mass protest will succeed or fail.

[15]According to Freedom House (2011), a blogger or other internet users were arrested for content posted online in 23 out of the 37 countries assessed.

4 The triggering of the assurance game.

Back to the N-person assurance game, assume that in a given society there is a population of N individuals $i \in \{1, 2, 3, ..., n\}$, and a government/regime whose only action is to try to repress any sign of political opposition, all living infinitely in a world with discrete time $t = 1, 2, 3, ...$ There is also a minimum requirement of a certain amount $\hat{X}$ of people for a protest to have a significant probability of succeeding, and this threshold is common knowledge among the population. If the amount of protesters were smaller than $\hat{X}$, no rational agent would find it optimal to take part in this protest.

4.1 The perfect information scenario

In the perfect information scenario, together with knowing $\hat{X}$, the population knows also p, the proportion of people willing to join a protest with at least $\hat{X}$ supporters. This known p is given by

$$p = \sum_{i=1}^{n} i \ : \left\{ \phi^i(i(A), p) > \phi^i(i(B), p) \iff p \geq \hat{X} \right\}. \tag{7}$$

There are only two possibilities,

1. $p \geq \hat{X}$, then it is optimal for at least $\hat{X}$ individuals to unite themselves for a common cause, and (mass) protests would be triggered.
2. $p < \hat{X}$, the proportion of those who would like to protest if $\hat{X}$ other individuals were protesting is less than this critical amount. It would not be optimal for anybody to mobilize against the regime/government's interests. No revolt would be triggered.

This is the solution for the standard N-person assurance game presented before. However, as it has already been argued, p is generally unknown, and certainly difficult to estimate.

4.2 Partially perfect information

Under this scenario, in the initial state $t = 1$ communication of potentially prosecuted political information across society follows the structure of a regular ring lattice with N nodes and degree z. Now, agents do not know p, but they are able to search across society to estimate it, following a process parallel to the Watts and Strogatz algorithm.

The previous analysis has given us two main conclusions:

1. By reducing the average distance in the social network, a relatively small amount of new connections between individuals who are initially far from each other implies a huge increase in the capacity the population has to estimate the general willingness to participate in a mass protest.

2. The risk for those actively trying to know the general state of opinion on this issue decreases fast as their number increases, but it is very high for those who try to do it first.

Which conditions will determine if an individual takes the risk of searching for this information in other parts of the social network?

A given individual i will only decide to try to rewire one of his links, searching for a new connection, if the utility he expects from doing this $U(\hat{w}_{it}/\hat{j}_{it})$ is higher than the utility he gets from remaining passive. We define this outside option as u_{it} and consider it to be normally distributed

$$u_{it} \sim N(\alpha_t, \sigma^2).$$

The dependence on time reflects the possibility of shocks to the average utility α_t, which can be caused by a sudden worsening of the economic situation, a success of a revolt in a neighboring country and, broadly speaking, by any of the causes explained for the assurance game in section 2.1.[16]

A new connection is formed if a potential protester contacts another agent in the network asking about his willingness to participate in a protest with at least $\hat{X}$ supporters, and the latter answers positively. We define w_t as the number of individuals who have built a new connection, an indication of the number of people trying to actively coordinate themselves into a mass protest. j_t is the number of people who have been repressed as a consequence of trying to establish new contacts. Hence, at a given period t, w_t/j_t gives a measure of the relative efficiency of potential protesters in coordinating themselves, versus that of the government/regime stopping the process. As we will see, society will estimate p as a function of this variable, $p(w_t/j_t)$.

In the initial period, $t = 1$, no links have been rewired, so the flow of information is at the least efficient possible point in the model, and this circulates very slowly. In this scenario it is assumed both w_t and j_t are public information, and as before, at the initial state before the process starts they take the value $w_1 = j_1 = 0$. Also, given the large network distance and the absence of any activity, society will initially believe that there is currently very little support for a mass protest, so we assume $w_1/j_1 = 0$.

When an individual tries to rewire one of his links he might get three outcomes.

[16] For the sake of simplicity, we assume these shocks affect the whole population in the same way. Since the effects of both a negative and a positive shock to α_t have a straightforward interpretation, our analysis focuses on the flow of information through society.

1. With probability $p(w_t/j_t)$ he will succeed, if he finds another individual who is in favor of protesting. They establish a new link, improving the flow of information across the potential protesters and increasing w_t by two.

2. With probability $1 - p(w_t/j_t)$ he will fail, if he connects with an individual who is not willing to protest. In this case he might:

   (a) Be punished, with probability $q(w_t/j_t)$ if he does not use internet to contact others, or with probability $q'(w_t/j_t)$ if he does use it, with $q(w_t/j_t) > q'(w_t/j_t)$. j_t increases by one in this case. A crucial assumption in our model is that an agent who tries to coordinate a protest using internet is less likely to be found and repressed than if he does it without using internet.

   (b) Stay free, respectively with probability $1 - q(w_t/j_t)$ and $1 - q'(w_t/j_t)$.

For large enough networks, $p(.)$ is equivalent to the proportion of N people who will respond positively to a new connection, and so it gives a measure of the number of agents willing to participate in a mass protest with at least $\hat{X}$ supporters, for a given −known− w_t/j_t.[17] $p(.)$ depends on w_t/j_t because an agent who is asked for his support to a potential protest bears also the risk of punishment. For simplicity, we will fix the probability of being repressed after being contacted by another agent to be the same as when looking for new information, $q(w_t/j_t)$ when not using internet and $q'(w_t/j_t)$ when using it.[18] As the ratio w_t/j_t increases, and a larger proportion of people is successfully establishing new connections, the proportion of these attempts which correspond to the regime forces (less flexible than society as a whole) decreases. Thereby, $q(.)$ and $q'(.)$ depend on w_t/j_t, because the probability of being repressed by authorities when trying to coordinate a mass protest decreases as the number of people who has already succeeded at doing it over those punished for trying increases,

$$\frac{dq(w_t/j_t)}{dw_t/j_t} \leq 0, \tag{8}$$

and as a direct consequence of this,

$$\frac{dp(w_t/j_t)}{dw_t/j_t} \geq 0, \tag{9}$$

the proportion of people joining the mass of potential protesters increases with w_t/j_t.

[17] For simplicity, we assume that every agent is truthful when establishing a new connection, in that if he has joined the group of potential protesters and finally a protest with at least $\hat{X}$ supporters is coordinated, he will take part in it. Assuming that only a proportion θ of those who build new connections will protest would only imply that the support for such a mobilization given w_t/j_t would not be $p(w_t/j_t)$ but $\theta p(w_t/j_t)$.

[18] We do this to simplify the model, without loss of generality. This probability could be assumed to be a different decreasing function of w_t/j_t.

Clearly, since agents are scared of declaring their real intention, given a low w_t/j_t $p(w_t/j_t)$ will in general be lower than the actual p. These two variables will coincide only after a certain threshold $(w_t/j_t)^*$.

These probabilities give the risk of being repressed when trying to rewire a connection, under perfect information,

$$R(w_t/j_t) = (1 - p(w_t/j_t))\, q(w_t/j_t), \tag{10}$$

when doing it without internet and

$$R'(w_t/j_t) = (1 - p(w_t/j_t))\, q'(w_t/j_t), \tag{11}$$

when this is done using internet.

The probabilities of being repressed satisfying $q(w_t/j_t) > q'(w_t/j_t)$ directly imply that the risk of being repressed when not using internet is larger than when using it,

$$R(w_t/j_t) > R'(w_t/j_t), \tag{12}$$

for all w, j and t. Notice that by using only one probability $p(.)$ we are implicitly assuming that the distribution of the population in terms of their attitude towards participating in a potential mass protest is the same for those who have internet access and those who do not have it.[19]

What determines whether a given agent takes the risk of looking for a new connection, or to answer positively about joining the mass of potential protesters?

We assume that agents are sophisticated enough as to know, or accurately estimate, the shape of the functions $p(.)$, $p'(.)$, $q(.)$ and $q'(.)$.[20] Given that in this scenario w_t/j_t is publicly known, the agents will know the corresponding probabilities and the implied risks $R(.)$ and $R'(.)$ in every moment.

We assume that the individual payoff when trying to rewire a link λ is flat in the current w_t/j_t. This is so because those who try to rewire a link knowing the proportion of those who have succeeded is much larger than the proportion of those who have been repressed, know there is a relatively high probability of succeeding $p(w_t/j_t)$, but at the same time, if they establish a new connection, they will get less valuable information than those who were among the first ones to do it.[21] For a matter of simplicity, we assume that these two effects compensate each other. Using the same argument, we consider the payoff for accepting to establish a new connection when asked about it δ, also flat in w_t/j_t. With these payoffs being constant, the variables which will determine the optimal action for

[19]If these are different, the validity of the model requires $p'(.)$ to be large enough as to still imply $R'(.) < R(.)$.

[20]This is equivalent to say that they are able to infer the support for a mass protest and the danger of actively trying to organize it implied by w_t/j_t.

[21]Formally, this is the case if we assume $p_1(.) > 0$ and $p_2(.) < 0$, where subscripts indicate respectively the first and second derivative.

each agent will be the costs of each choice, as measured by the risk of being repressed by the authorities, and the outside option u_{it}. Then, for an individual without internet access, it is optimal to try to rewire a link if and only if

$$U(\hat{w}_{it}/\hat{j}_{it}) = \lambda - R(\hat{w}_{it}/\hat{j}_{it}) \geq u_{it}, \tag{13}$$

and for an individual with internet access, this is optimal if and only if

$$U'(\hat{w}_{it}/\hat{j}_{it}) = \lambda - R'(\hat{w}_{it}/\hat{j}_{it}) \geq u_{it}. \tag{14}$$

Similarly, for an agent asked about joining the declared potential protesters, the equivalent conditions are respectively given by

$$V(\hat{w}_{it}/\hat{j}_{it}) = \delta - q(\hat{w}_{it}/\hat{j}_{it}) \geq u_{it}, \tag{15}$$

for an individual without internet access, and by

$$V'(\hat{w}_{it}/\hat{j}_{it}) = \delta - q'(\hat{w}_{it}/\hat{j}_{it}) \geq u_{it}, \tag{16}$$

for an individual with internet access.[22]

Thereby, for a given individual to explore the feasibility of the protest, the utility derived from doing so must be higher than the utility of remaining linked to the same agents as before, with the previous w_t/j_t. Since $R(.) > R'(.)$ and $q(.) > q'(.)$, for every w, j and t, everything else being equal, more people will tend to take the risk of working in favor of a mass protest if the access to internet is larger.

Imagine a given society is in a stationary state, with a network structure which corresponds to that of the initial period of the game in our model, a ring lattice network with N nodes and degree z. Conditions are such that given $w_1/j_1 = 0$, u_{i1}, $R(.)$, and $R'(.)$ no agent finds it optimal to look for new connections. In this same period, α_1 the mean of agents' outside option u_{i1} receives a negative shock, decreasing the utility everybody gets by remaining passive,[23] as to the point that some agents would find it optimal to try to rewire one of their links, in order to explore the general willingness to take part in a mass protest. The probability of being punished when doing this without internet is higher than when it is doing through this channel, implying, for large enough N:

1. More people are expected to take the risk of looking for new connections if there is a larger proportion of the population with internet in the country, since $R(.) > R'(.)$.

[22]In general we should assume $V(.)$ to be larger than $U(.)$ for every individual and every perceived w_t/j_t, so in every period less people are searching for connections than will answer affirmatively if they are contacted. For this to be the case, $\delta > \lambda$.

[23]Examples of such shocks could be the success of a mass demonstration in a neighboring and similar country, an increase in energy, food or transport prices, etc.

2. Among those who try, a lower proportion of them are expected to be punished for it, since $q(.) > q'(.)$. This also implies that more of those who are asked about their potential support will respond affirmatively.

In the hypothetical case that a country could face this shock in two identical scenarios, apart from a different proportion of the population having internet access, for large enough N these two effects together guarantee $w_2'/j_2' > w_2/j_2$.[24] This directly implies $w_3'/j_{i3}' > w_3/j_3$, and agents in $t = 4$ will again be less likely to be found and punished by authorities, since this probability $q(.)$ decreases as w_t/j_t increases. If many of those agents using internet succeed, the perceived risk will decrease for everybody, also for those agents without internet access, and some of them will also join the process, pushed by a relatively large w_t/j_t.

Hence, once the process starts, it might follow two different paths. The possible results are:

1. Authorities are sufficiently efficient repressing those trying to get new information so that w_t/j_t does not grow enough as for $p(w_t/j_t)$ to approach p, and it remains lower than $\hat{X}$. In this case, society, and potential protesters in particular, are not able to estimate their number, and independently of whether p is larger or smaller than the critical amount $\hat{X}$, no protests would be triggered.

2. Authorities are not efficient enough repressing agents who communicate with each other. After some point, $q(.)$ and $q'(.)$ are negligible, and w_t/j_t will grow to the point of making $p(w_t/j_t)$ close enough to the true value of potential protesters p. At this point, the situation is equivalent of that under perfect information, and so two different scenarios are plausible:
   (a) $p < \hat{X}$, so no protests would be triggered.
   (b) $p \geq \hat{X}$, and a protest with at least $\hat{X}$ supporters would be triggered.

Since the utility of looking for new connections for a given individual i is highly supermodular on the proportion of previous individuals succeeding, increasing access to internet by a small proportion of those individuals for whom $U'(.) \geq u_{it} \geq U(.)$ and of those for whom $V'(.) \geq u_{it} \geq V(.)$ can make a big difference in the resulting equilibrium. A few exemplary early punishments could prevent the process from starting, but if this begins −online− with a relatively low proportion of those trying to coordinate the protest being repressed for it, it might well become a mass phenomenon, including a significant part of the population.

This gives the first big result of this study, that a relatively small increase in the level of internet access in a given country in which the discontent with

[24](') refers, again, to the scenario with internet access, or equivalently to a scenario in which a larger proportion of the population has internet access.

the current situation is large enough, might mean the difference between nothing happening (as in point 1 above) and a mass protest being triggered (as in 2b).

Most authoritarianisms and autocraticies are not interested in potential protesters getting to know their real social support, and so assuming w_t/j_t to be public knowledge is not very realistic. In the next section, we study the situation when this variable is private information.

4.3 Support for the protest as private information

The previous results depend on the agents having perfect information about w_t/j_t, which in general cannot be known or accurately estimated. Denoting each agent's perceived w_t and j_t as $\hat{w}_{it}$ and $\hat{j}_{it}$, in the very first periods they will find the value of $\hat{w}_{it}/\hat{j}_{it}$ to be extremely low, what makes their perceived strength of potential protesters $p(\hat{w}_{it}/\hat{j}_{it})$ also low. How do agents estimate the value of these variables?

We define an updating rule similar to the DeGroot model for the study of consensus in networks,[25] assuming that in the initial state, communication of pursued political information across society follows the structure of a regular ring lattice. Given the initial condition $w_1/j_1 = 0$, the real number of connections will again increase if some agents try to search for new contacts and they succeed in their attempt. When a new connection is created, the increase in w_t will be first known by those two agents who connect with each other, and each period this information will spread through the network, beginning by their first neighbors. In the same way, when an agent is repressed when trying to get a new connection, he will know it in that same period, these news will reach his direct neighbors in the next period, and so on.[26] Formally, in each period t agents update their beliefs by taking weighted averages of their neighbors' beliefs.

The trust agent i places on agent h's opinion is given by T_{tih}, the corresponding entry of the square, row stochastic matrix T_t.[27] We assume that agents get direct information only from those who are directly linked to them in the network, and each agent places the same trust on each of his connections,[28] so the trust matrix is just equal to the adjacency matrix[29] of the network A_t with its A_{tih} entry

[25] See DeGroot (1974) or more recent papers based on this model, such as Golub and Jackson (2010).

[26] We are implicitly assuming that a repressed agent might still communicate with the rest of the network and participate in the mass protest in case this happens at all. Another possibility would be to assume he is arrested, becoming inactive, and these news directly reach his first neighbors.

[27] The sum of the elements in each row is 1.

[28] Any other trust distribution would be as arbitrary, and this significantly simplifies the model.

[29] A square, symmetric matrix with 1 in the entry A_{ih} if there is a direct link between i and h, and 0 otherwise.

divided by z_{it}, the node degree for agent i in period t. After some of the links have being rewired some nodes will be connected to more neighbors than others, so T_t will not be symmetric in general.

Agent i's belief about w_t in period t is given by $\hat{w}_{it}$, and we denote the $N \times 1$ vector containing all the beliefs in each period by $\hat{W}_t$. Then, for $t = \{2, 3, ...\}$ the updating rules governing the collective dynamics of the system are simply given by

$$\hat{W}_t = T_t \hat{W}_{t-1} + C_t \tag{17}$$

for beliefs about w_t and

$$\hat{J}_t = T_t \hat{J}_{t-1} + G_t \tag{18}$$

for beliefs about j_t. C_t is an $N \times 1$ vector with each entry c_{it} giving the number of successfully rewired connections by agent i since $t = 1$, and G_t is another $N \times 1$ vector with each entry g_{it} equal to 1 if agent i has ever been repressed by authorities and 0 otherwise. The reason to introduce these vectors is that each agent knows whether he has been punished or if he has got new connections, and so he does not discount this information with others' beliefs. Then $\hat{W}_{it}$, the ith entry of the vector $\hat{W}_t$, gives $\hat{w}_{it}$, each individual's estimation of the total number of completed successful connections by period t. Similarly, $\hat{j}_{it}$ is given by $\hat{J}_{it}$, the ith entry of the vector $\hat{J}_t$. Therefore, the main variable of interest, $\hat{w}_{it}/\hat{j}_{it}$ is directly obtained as the ratio of these two variables for each agent.

Information about each new connection and punishment flowing first to those agents closer to it in the social network implies that peer effects[30] are at work in the model. In the short run every agent i's estimation of $\hat{w}_{it}/\hat{j}_{it}$ will be more influenced by the beliefs of those who are closer to him, but at the same time, through these agents he will be receiving information coming from many other agents who are farther away in the network, all of them as $t \to \infty$.

An important issue is whether these updating rules make the group tend to, and *de facto* do reach, consensus, being this understood as every member in the group sharing the same opinion about a certain state of the world. If this is the case, it is also interesting to find out if this consensus happens to the true value of the variable of interest, if social learning converges to wisdom. Denoting x_{it} as the state or belief about a given variable x for individual i in period t, we argue a group reaches consensus at $t \to \infty$ when $|x_{it} - x_{ht}| \to 0$ for every pair of its members, see for example Estrada and Vargas-Estrada (2013). If this consensus happens to the true value x, formally $|x_{it} - x| \to 0$ as $t \to \infty$ for every i, society converges to wisdom. Do these updating rules produce consensus? And if so, does society converge to wisdom?

For inquiring whether beliefs tend to consensus, we must assume that at a certain point in time, w_t and j_t stop increasing, as otherwise new information is continuously arriving to certain nodes, and strict consensus is not possible.

[30]See Estrada and Vargas-Estrada (2013).

The introduction of C_t and G_t gives every individual who has either found a new connection or been repressed for trying to do it infinite memory about his own state, making these two updating rules explosive. Formally,

$$\sum_i \hat{w}_{it+1} > \sum_i \hat{w}_{it}; \ \sum_i \hat{j}_{it+1} > \sum_i \hat{j}_{it}, \tag{19}$$

respectively for every $w_t, j_t > 0$. This implies that even if consensus were reached, this would not converge to the true value of these variables, society will not know either w_t or j_t accurately. Simulating the model we find that consensus is in fact reached for the vectors of beliefs $\hat{W}_t$ and $\hat{J}_t$.[31] As t increases, $\hat{w}_{it}$ and $\hat{j}_{it}$ grow larger and larger, making the individual certainty about the own state, which adds either 0 or a relatively small[32] positive number to the individual estimation, negligible.

This implies that the main variable of interest in the model, $\hat{w}_{it}/\hat{j}_{it}$ also tends to consensus. Does this consensus happen to the true value w_t/j_t? Using results from Golub and Jackson (2010), we know that given a non-directed, connected network A_t, and a stochastic trust network T_t, each agent influence on the general opinion is simply given by his degree, divided by the sum of the degrees of all the nodes in the network,

$$s_i = \frac{z_i}{\sum_i z_i}. \tag{20}$$

Given some conditions on initial beliefs, wisdom will be attained if every agent's influence vanishes as the size of the network N increases. Then, just assuming a large enough N, a one shot game of our model, letting the information about its result flow through the network for a sufficiently long t, would converge to wisdom. If for example at $t = 2$ four agents successfully establish new connections, two agents are repressed while trying to do it, and there are not any new actions for $t \geq 3$, $\hat{w}_{it}/\hat{j}_{it}$ would converge to the true value 2. However, the dynamic nature of our model modifies this result in an essential way. The tendency of $\hat{w}_{it}$ and $\hat{j}_{it}$ to increase over time, as given by equation (19) implies that early news will get a larger weight in agents' estimation of w_t/j_t. This will in general prevent society from converging to wisdom. Those results, connections and punishments, in relatively early periods will be over-weighted in the estimation, while those in later periods will be penalized. Importantly, this implies that the real $p(.)$, which depends on the perceived $q(.)$ and $q'(.)$ and not on the real ones, and so, $R(.)$ and $R'(.)$ do not coincide with those under perfect information. These variables are now determined by the whole history of w_t/j_t, denoted by H_t (not just by their real value w_t/j_t). This real variables will then be written as $p(w_t/j_t|H_t)$, $R(w_t/j_t|H_t)$ and $R'(w_t/j_t|H_t)$, and in general $(w_t/j_t|H_t) \neq (w_t/j_t)$.

[31]While it is out of the scope of this paper, proving this result analytically should not be difficult, as it follows standard results in network theory, see for example Golub and Jackson (2010), with the particularity of beliefs not converging to a well-defined limit.

[32]As compared to the explosive behavior of the estimated variables $\hat{w}_{it}$ and $\hat{j}_{it}$.

As we will see, this makes the outcome of early periods in the game crucial to determine the final equilibrium.

Agents are still sophisticated enough as to know or accurately estimate the shape of the functions $p(.)$, $p'(.)$, $q(.)$ and $q'(.)$.[33] Then, for agent i and period t, given conditions (8) and (9), as long as $\hat{w}_{it}/\hat{j}_t < w_t/j_t$, the perceived risk for a given individual will be larger than that under the scenario with perfect information about w_t/j_t. Given that now information needs to flow from agent to agent, the initial increase in w_t/j_t will in general be slower than when this ratio is public information. As w_t increases and information flows faster, agents' beliefs will be closer to consensus, however, as we have seen, this consensus value $\hat{w}_t/\hat{j}_t$ will generally not converge to wisdom. Formally, $\hat{w}_{it}/\hat{j}_{it} \neq w_t/j_t$ as $t \to \infty$.

These characteristics, convergence to consensus but not to wisdom, and the dependency of those generally biased beliefs on the timing of new connections being created and agents being punished, give the main predictions of the theoretical model.

1. As w_t increases, the beliefs about the capacity of potential protesters to coordinate in a mass protest, as given by $\hat{w}_{it}/\hat{j}_{it}$, will tend to a consensus.

2. Whether a mass protest happens or it does not occur depends on agents' perception, which is determined by the whole story of w_t/j_t, and not simply by the actual value of this variable.

3. Since early events are weighted more in agents' estimation, the outcome in the first periods will be crucial to determine the resulting equilibrium.

The steps of the process are the same as in the previous scenario, in which w_t/j_t was perfectly known. However, agents' biases in this estimation might have important implications. Perceived authorities' and organizers' efficiency will now determine the result. The number of people protesting will be given by the following condition:

$$\hat{p} = \sum_{i=1}^{n} i \ : \left\{ \phi^i(i(A), \hat{p}_{it}) > \phi^i(i(B), \hat{p}_{it}) \iff \hat{p}_{it} \geq \hat{X} \right\} \wedge \left\{ \hat{p}_{it} \gg \hat{X} \right\}. \tag{21}$$

Where $\hat{p}_{it} = p(\hat{w}_{it}/\hat{j}_{it})$ This condition reads that those who get a larger utility from protesting than from not protesting, if the protest counts with at least $\hat{X}$ supporters, and also estimate the support to be much larger than this threshold, will coordinate in a protest. The much larger symbol accounts for the uncertainty potential challengers have about the accuracy of their estimate, both with respect to the true w_t/j_t and with respect to others' perception of this variable.[34]

[33]The problem is that in general their estimate of w_t/j_t is wrong.

[34]When $\hat{w}_{it}/\hat{j}_{it}$ is very large for some agents, individuals' estimations will be pretty close to consensus, but in practice there will still remain some differences. The $\gg$ sign prevents the first agent reaching $\hat{p}_{it} \geq \hat{X}$ to begin a protest and find himself alone.

Depending on the value $\hat{p}$ takes when this condition is first satisfied, the following results are possible:

1. Again, if authorities are relatively efficient in repressing those trying to coordinate the protest, the estimated $p(.)$ will not be able to increase enough for any agent. Agents will stop searching and the protest will not be triggered, independently of the real proportion of people willing to participate in a protest which gets $\hat{X}$ supporters together, p. The main difference is that now this might also happen with w_t/j_t being relatively high, if the authorities were particularly efficient in early periods, as these will have a larger weight in agents' estimation.

2. If the organizers are relatively efficient, and a protest is triggered, this might happen under different conditions, as given biased estimates the number of protesters $\hat{p}$ will in general not be equal to p:

   (a) $\hat{X} < \hat{p} = p$, the communication of private information among agents has the same implications as it would have under perfect information. This will happen only under very concrete functional forms and will not hold in general.

   (b) $\hat{X} < \hat{p} < p$, the communication process has made some potential challengers over-fearful and the protest will count with less people than in the scenarios with perfect information.

   (c) $\hat{X} < p < \hat{p}$, the communication process has taken society to over-estimate the efficiency of those organizing the protest, and more people will mobilize than under perfect information.

   (d) $p < \hat{X} < \hat{p}$, significantly biased protestors coordinate in a mass protest which would not happen in the scenarios with perfect information.

Scenario (b) will hold when authorities are relatively more efficient in the beginning, and this is over-weighted in society's estimation. Note that the last two situations are caused by a fake enthusiasm produced by an over-optimistic perception of the early success of those who were first in trying to mobilize the population into a mass protest.[35] This might make the perceived $\hat{w}_{it}/\hat{j}_{it}$ larger than its real value for many i, who will then perceive p_{it} larger than it really is, overestimating the relative strength of the discontent people. In (c), with condition (21) first satisfied when this is the case, more people will join the mass protest than it would under a perfect information scenario. In the extreme case

[35] In another case, $\hat{p} < \hat{X} < p$, some importantly biased potential protestors estimate the number of those joining them to be much larger than it really is, and they coordinate in a protest condemned to fail. This should not happen in this environment, since it requires important divergences in $\hat{w}_{it}/\hat{j}_{it}$ across agents, and these should not hold for values of w_t large enough for the required increase in some $\hat{p}_{it}$, as large w_t takes agents estimation of w_t/j_t close to consensus.

(d), with the situation being such that $\hat{p} \geq X > p$, the protest will be triggered by over-optimistic beliefs, with the upward biased estimation of the numbers of supporters becoming a self-fulfilled prophecy. This protest would not happen under perfect information, however, challengers decide to protest because they expect more people to do it than they should, and this itself takes more agents to protest than they should if knowing the true situation.

This gives us the second big result of this study, that a relatively small increase in the level of internet access in a given country might trigger a process that not only allows potential protesters to estimate their proportion accurately, but also influences their number, making possible mass protests, some of which would arguably not happen with a lower proportion of the population having internet access.

Hence, this process points towards a high potential for internet as a tool for facilitating the triggering of mass protests, and helps to explain one of the reasons why it tends to be so controlled and feared by authoritarianisms and autocracies.

5 Some illustrative evidence: The Arab Spring

The Arab Spring origins lie in the series of events that triggered the so-called *Jasmine revolution* in the Tunisian region of Sidi Bouzid in December 2010. A young man, Mohammad Bouazizi, set himself on fire in front of the regional office as a response against a pressing precarious situation. This act fuelled further mobilizations in Tunisia (additional self-immolations, but also sit-ins and marches), which spread to other Arab countries throughout 2011, like Egypt and Libya, where rulers were also forced from power. Civil uprisings and major protests also took place in Syria, Algeria, Bahrain, Lebanon, Yemen, Morocco, Kuwait, Jordan, Sudan and Iraq, and to a lesser extent in Oman, Saudi Arabia, Mauritania, Djibouti, Western Sahara and Palestine. Most of these regimes were in 2011 −and still are− under authoritarian or autocratic rule, with only a few exceptions. According to the 'Democracy Index' (Freedom House 2013), only Kuwait, Morocco and Lebanon have recurrently been classified as partly free (highest regional scores), whereas so have been occasionally Egypt, Tunisia and Iraq. The lowest scores (i.e. least free) are for Saudi Arabia and Yemen.

Notwithstanding this inter-regime variability, our specified restrictions and model developed above work as long as one basic condition holds: the perceived risk of exchanging political information online in non-democracies is lower than doing so offline.

As we have previously argued, the utility of participating in collective actions in autocracies or authoritarianisms is a function of the expected levels of collective engagement. If having more accurate information on the perceived support for a collective action, utility of demonstrating increases sharply. The internet, as reflected during the Arab Spring revolts, can supply the infrastructure that facilitates exchanges of political information, affecting individuals expectations of potential challengers willingness to resort to the (offline) protest arena.

To illustrate this point, we use Google trends. It contains data on the diachronic evolution of searches in the Google browser using specific keyword combinations during a given time span relative to the total searches done on Google over time. Data are normalized and presented on a 0-100 scale (Google trends 2014)[36]. These data can be disaggregated per country. Thus, using the input "revolt" in Arabic,[37] we can observe how the number of searches using that keyword increased dramatically just prior to mass protest events taking place in that country (25/01/2011 in Egypt, 15/02/2011 in Libya, 20/02/2011 in Morocco, 15/03/2011 in Syria, respectively).[38]

Therefore, despite inter-country variability (e.g. intensity, success, contextual factors, role of institutions religious and civic associations, parties, trade unions, etc.), people accessed the internet to improve their knowledge about the potential mass protest and/or to contribute to coordinate it, before this happened. The only partial exception is Tunisia (figure 9). We mean *partial* because, on the one hand, the catalystic event that fostered the wave of contention took place in the first instance (17/12/2010). Moreover, we should speak about *escalation of contention* to the detriment of *sudden mass gatherings*: claim-making performances and claimants involved increased steadily, as so happened with numbers of protest events and levels of repression, while repertoires widened (sit-ins, demonstrations, strikes, etc.) −see Sergi and Vogiatzoglou (2013). On the other hand, we observe a similar trend when truly anti-government/regime *mass protests* occurred. Our criterion for qualifying a contentious event as a mass protest in the context of the Arab Spring is its size: at least 10,000 challengers should be gathered. Since we only have estimations of these figures, we could set the date either on 06/01/2011 (there was only 8,000 participants in a lawyers' strike, though), 14/01/2011 (several thousands attended to the UGTT union-called strike) or 17/01/2011 (also thousands rallied against the presence of *Democratic Constitutional Assembly* members in the new Government led by Mohamed Ghannouchi after Ben Ali had fled a few days earlier). The picture will slightly change depending on which day we exactly consider the first mass protest took place. Provided we pick either 14/01/2011 or 17/01/2011, its substantive interpretation will not change dramatically: online mobilization was prior to physical mass protest events in Tunisia. In fact, most previous protests before that across Tunisia had a local scope and were relatively small in size.

[36] The amount of searches at each point on the graph is divided by the highest point, and expressed as a percentage of this, see Google trends (2014).

[37] ثْوْرْهْ

[38] Results for other countries are either similar (not included here due to space constrains; available upon request) or non-existing (information from Google trends not provided or not available).

The finding is more robust in the other cases within the Arab context online —i.e. revolt searches preceding physical mobilization. This gives us an idea of how individual curves of individual participation in collective action (and the respective tipping points) may have been altered thanks to the internet. It points towards a relationship between the timing of online political activity and the moment in which the mass protests were being organized. As Grossman (2009) states referring to the mass protests occurred after the Iranian presidential election in 2009: "there's no question that it [Twitter] has emboldened the protesters, reinforced their conviction that they are not alone and engaged populations outside Iran in an emotional, immediate way that was never possible before". Notwithstanding this, we should be careful with the importance of social networks in general and Twitter in particular as channels for the spread of information in the Arab Spring. As Aday et al (2013) show, Twitter was an information channel for non-MENA onlookers during the Arab Spring but less so for protesters on the ground. The Arab Spring did not merely consist of social media revolts. However, as we argue throughout and our country-level browser data illustrate, internet —not only social networks— can make the difference for individual utility of participation in collective action by providing the infrastructure for exchanges of political information.

Figure 9: "Revolt" searches in Google trends, Tunisia.

Figure 10: "Revolt" searches in Google trends, Egypt.

Figure 11: "Revolt" searches in Google trends, Libya.

Figure 12: "Revolt" searches in Google trends, Morocco.

Figure 13: "Revolt" searches in Google trends, Syria.

6 Conclusion.

Our model accounts for the process which makes individual estimations of the plausible numbers of challengers involved in anti-regime or anti-government protests possible in authoritarian or autocratic settings. Previous models based on the N-person assurance game established that above a given proportion of the population getting united to engage in collective action, it would be optimal for a given individual to choose to protest. After this, they explain how different social, political and economic conditions could reduce or increase these individual critical points. When certain conditions make this critical value too low for a significant amount of people, mass protests will take place.

We argue that in societies in which any expression of a political opinion against the government or regime bears the risk of strong repression, poor transmission of (often forbidden) political information across different social groups can make communication too inefficient as for mass protests to be triggered. This can happen even when the amount of people willing to coordinate against the government or regime is very high and if organized could potentially pose a real challenge for the political status quo. We make an analogy between this situation and the slow flow of information in a regular ring lattice. After this, using the Watts and Strogatz's model for Small Worlds, we illustrate how a few individuals contacting others who are socially far from them can critically improve the transmission of information across society. The model also accounts, through the concept of betweenness, for the high risk borne by those who are first in trying to contact others.

Thereby, as the quality of the communication across society increases sharply with a few new connections, the risk of looking for new contacts decreases also very fast as soon as there is a significant number of individuals doing it. The reason is that there is a high strategic complementarity between the individual strategies of looking for new information about the prevailing political opinion in other social groups. This supermodularity means that any element which facilitates the communication of information across society, will have a multiplicative effect. As a consequence, internet access can significantly affect the political equilibrium of a country. A relatively small increase in private internet access across the population, given certain circumstances, might well be enough to trigger the beginning of the flow of information required for the population to coordinate on mass protests.

The results from this theoretical model point to the Web as having a high potential to facilitate challengers' estimations of their relative size and support. Particularly, we should expect its influence to be higher at the beginning of the process, and relatively less relevant after the first events happening, once protesters and potential challengers are aware of their magnitude. After this point, completely shutting down the network, as was done in Egypt, Syria and Libya, might not help to stop the ongoing revolts.

References

[1] Acemoglu D., A. Ozdaglar and A. ParandehGheibi (2010) 'Spread of (mis)information in social networks', *Games and Economic Behavior*, 70(2), pp. 194-227.

[2] Amaral L., A. Scala, M. Barthélémy, and H. E. Stanley (2000) 'Classes of small world networks', *PNAS*, 97(21), pp. 11149-11152.

[3] Beissinger M. R. (2007) 'Structure and Example in Modular Political Phenomena: The Diffusion of Bulldozer/Rose/Orange/Tulip Revolutions', *Perspectives on Politics*, 5(2), pp. 259-276.

[4] Bennett W. L. and A. Segerberg (2013) *The Logic of Connective Action. Digital Media and the Personalization of Contentious Politics*, Cambridge, New York: Cambridge University Press.

[5] Bimber B. (1998) 'The Internet and political transformation: Populism, community, and accelerated pluralism', *Polity*, 31(1), pp. 133-160.

[6] Büchler S. (2004) 'The strange career of strain and breakdown theories of collective action', in D. A. Snow, S. Soule and H. Kriesi (eds.) *The Blackwell Companion to Social Movements*, Oxford: Blackwell, pp. 47-66.

[7] DeGroot M. H. (1974) 'Reaching a Consensus', *Journal of the American Statistical Association*, 69(345), pp. 11821.

[8] Diani M. (2000) 'Social movement networks virtual and real', *Information, Communication & Society*, 3(3), pp. 386-401.

[9] Donker T. H. (2013) *Islamism and the Arab Spring: a Social Movements Approach*, PhD thesis defended at the European University Institute.

[10] Earl J., and K. Kimport (2011) *Digitally enabled social change. Activism in the internet age*, Cambridge: MIT Press.

[11] Estrada E. (2011) *The structure of complex networks: theory and applications*, Oxford: Oxford University Press.

[12] Estrada E. (2012) 'Path Laplacian matrices: introduction and application to the analysis of consensus in networks', *Linear Algebra and its Applications*, 436(9), pp. 3373-3391.

[13] Estrada E. and N. Hatano (2008) 'Communicability in complex networks', *Physical Review*, E 77, 036111.

[14] Estrada E. and E. Vargas-Estrada (2013) 'How Peer Pressure Shapes Consensus, Leadership, and Innovations in Social Groups', *Nature Scientific Reports* 3, art. 2905.

[15] Ferdinand P. (2000) 'The Internet, democracy and democratization', *Democratization*, 7(1), pp. 1-17.

[16] Finkel S. E. (2008) 'In Defense of the Wide Rational Choice Model of Collective Political Action', in A. Diekmann, K. Eichner, P. Schmidt, and T. Voss (eds.) *Festschrift fr Karl-Dieter Opp*, Opladen: Westdeutscher Verlag.

[17] Finkel S.E., E. N. Müller and K. D. Opp (1989) 'Personal Influence, Collective Rationality and Mass Political Action', *American Political Science Review*, 83(3), pp. 885-903.

[18] Freedom House (2011) 'Freedom on the Net 2011', online, URL:http://www.freedomhouse.org/sites/default/files/FOTN2011.pdf, updated 18/04/2011, accessed 08/09/2013.

[19] Freedom House (2013) 'Democracy Index. Middle East and North Africa', online, URL: http://www.freedomhouse.org/regions/middle-east-and-north-africa#.UuApcBar0Us, updated n.d., accessed 03/01/2014.

[20] Golub B. and M.O. Jackson (2010) 'Naive Learning in Social Networks and the Wisdom of Crowds', *American Economic Journal: Microeconomics*, 2(1), pp. 112-149.

[21] Google trends (2014) 'About Trends graphs', online, URL: https://support.google.com/trends/answer/4355164?hl=en, updated n.d., accessed 11/01/2014.

[22] Granovetter M. (1974) 'The strength of weak ties', *American Journal of Sociology*, 78(6), pp. 1360-1380.

[23] Grossman L. (2009) 'Iran protests: Twitter, the medium of the movement', *Time Magazine*, online, URL: http://content.time.com/time/world/article/0,8599,1905125,00.html, updated 17/06/2009, accessed 07/01/2014.

[24] Karklin R. and R. Petersen (1993) 'Decision Calculus of Protesters and Regimes: Eastern Europe 1989', *The Journal of Politics*, Cambridge: Cambridge University Press.

[25] Kochen M. (ed.) (1989), *The Small World*, Norwood, NJ: Ablex.

[26] Lichbach M. I. (1995), *The Rebel's Dilemma*, Ann Arbor, MI: Michigan University Press.

[27] Loader B. D. and D. Mercea (2011), 'Networking democracy?', *Information, Communication & Society*, 14(6), pp. 757-769.

[28] Marwell G., and P. Oliver (1993) *The Critical Mass in Collective Action. A Micro-Social Theory*, New York: Cambridge University Press.

[29] Mercea D. (2012) 'Digital prefigurative participation: The entwinement of online communication and offline participation in protest events', *New Media and Society*, 14(1), pp. 153-169.

[30] Micó, J. L. and A. Casero-Ripollés (2013) 'Political activism online: organization and media relations in the case of 15M in Spain', *Information, Communication & Society*, online, URL: http://www.tandfonline.com/doi/full/10.1080/1369118X.2013.830634#.Usn0lRYZJES, accessed 22/11/2013, updated 27/08/2013.

[31] Oberschall A. (1994) 'Rational Choice in Collective Protests', *Rationality and Society*, 6(1), pp. 79-100.

[32] Olfati-Saber R., J.A. Fax and R. M. Murray (2007) 'Consensus and cooperation in networked multi-agent systems', *Proc. IEEE*, 95, pp. 215233.

[33] Olson Mancur (1965) *The Logic of Collective Action*, Cambridge, MA: Harvard University Press.

[34] Opp K.-D. (1989) *The Rationality of Political Protest*, Boulder, CO: Westview Press.

[35] Opp K.-D. (2009) *Theories of Political Protest and Social Movements. A Multidisciplinary Introduction, Critique and Synthesis*, London: Routledge.

[36] Opp K.-D. (2013) 'Rational Choice Theory and Social Movements', in D. A. Snow, D. della Porta, B. Klandermans and D. McAdam (eds.) *The Wiley-Blackwell Encyclopedia of Social and Political Movements*, Oxford: Blackwell Pub.

[37] Opp K.-D., S. E. Finkel, E. N. Müller, J. Green and G. Wolfsfeld (1995) 'Ideological Commitment and Collective Political Action', in B. Klandermans and J. C. Jenkins (eds.) *The Politics of Social Protest: Comparative Perspectives on States and Social Movements*, Minneapolis, MN: University of Minnesota Press.

[38] Ostrom E. (1998) 'A behavioural approach to the Rational Choice theory of Colllective Action', APSA Presidential Address, printed in *American Political Science Review*, 92(1), pp. 1-22.

[39] Rousseau, J. J. (1754) *Discourse on the Origin and Basis of Inequality Among Men*, online, URL: http://www.constitution.org/jjr/ineq.txt, accessed 27/12/2013, updated n.d.

[40] Sen A. K. (1967) 'Isolation, Assurance and the Social Rate of Discount', *Quarterly Journal of Economics*, 81(1), pp. 112-124.

[41] Schelling T. (1985) *Micromotives and Macrobehavior*, New York: Gordon and Breach.

[42] Sergi V. and M. Vogiatzoglou (2013) 'Think globally, act locally? Symbolic memory and global repertoires in the Tunisian uprising and the Greek anti-austerity mobilizations', in C. Flesher Fominaya and L. Cox (eds.) *Understanding European Movements: New social movements, global justice struggles, anti-austerity protest*, London: Routledge.

[43] Skyrms B. (2004) *The Stag Hunt and the Evolution of Social Structure*, Cambridge: Cambridge University Press.

[44] Sueur C., J. L. Deneubourg and O. Petit (2012) 'From social network (centralized vs. decentralized) to collective decision-making (unshared vs. shared consensus)', *PLoS one*, 7(2), e32566.

[45] Topkis D.M. (1998) *Supermodularity and Complementarity*, Princeton, NJ: Princeton University Press.

[46] Van Aelst, P., and Walgrave, S. (2002) 'New media, new movements? The role of the Internet in shaping the "anti-globalization" movement', *Information, Communication & Society*, 5(4), pp. 465-493.

[47] Vega-Redondo, F. (2007) *Complex Social Networks*, Econometric Society Monographs, Cambridge: Cambridge University Press.

[48] Watts D. J. and S. H. Strogatz (1998) 'Collective dynamics of 'small-world' networks', *Nature*, 393, pp. 440-442.

[49] Wasserman S. and K. Faust (1994) *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences, Cambridge: Cambridge University Press.

[50] Yang Lu, Wing Suen and Heng Chen (2013) 'The Power of Whispers: A Theory of Rumor, Communication and Revolution', Meeting Papers 411, Society for Economic Dynamics.