Regime Change in Large Information Networks

Joan de Martí
Pau Milán

June 2018

Barcelona GSE Working Paper Series

Working Paper nº 1049

Regime Change in Large Information Networks*

Joan de Martí[†] Pau Milán[‡]

June, 2018

Abstract

We study global games of regime change within networks of truthful communication. Agents can choose between attacking and not attacking a status quo, whose strength is unknown. Players share private signals on this state of the world with their immediate neighbors. Communication with neighboring players introduces local correlations in posterior beliefs and also allows for the pooling of information. In order to isolate the latter effect, we provide, as a methodological contribution, sparseness conditions on networks that allow for asymptotic approximations that eliminate covariances from equilibrium strategies. We ask how changes in the distribution of connectivities in the population affect the types of coordination in equilibrium as well as the likelihood of successful rally. We find that without a public signal strategic incentives align, and the probability of success remains independent of the type of network. With a public signal the distribution of degrees unambiguously affects the probability of success, although the direction of change is not monotone, and depends crucially on the cost of attack.

*This paper was previously circulated under the title "Regime Change in Large Social Networks" and was presented for the first time in the 17th Coalition Theory Network Workshop in Paris in February 2012. After that, it was not circulated. The authors would like to thank Coralio Ballester, Yann Bramoullé, Vasco Carvalho, Jan Eeckhout, Rachel Kranton, Gabor Lugosi, Alessandro Pavan, Yves Zenou, and seminar participants at CTN-Paris, UPF, and CREI for helpful comments. Joan de Martí gratefully acknowledges financial support from the Spanish Ministry of Education through project SEJ2006-09993/ECON and a Juan de la Cierva fellowship (co-financed by the European Social Fund), from the Spanish Ministry of Science and Innovation through grant ECO2011-28965, from the Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa de la Generalitat de Catalunya through Beatriu de Pinós 2009 BP-A 00116 postdoctoral grant, and from Barcelona GSE and the Government of Catalonia. Pau Milàn acknowledges financial support from the Spanish Ministry of the Economy and Competitiveness/FEDER through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015- 0563) and by ERC AdG - GA324048 grant. All errors are ours.

[†]Universitat Pompeu Fabra, and Barcelona GSE

[‡]Universitat Autònoma de Barcelona, MOVE, and Barcelona GSE
(corresponding author: pau.milan@gmail.com)

1 Introduction

Important features of political and economic life are determined by the coordinated movement of large numbers of individuals. The Arab Spring or the sub-prime crises, for instance, are fascinating for the great power of many concentrated (and otherwise negligible) individual decisions. In these environments, individuals' payoffs depend on the share of other individuals that are simultaneously choosing to act a certain way. Moreover, in the case of political uprisings or currency attacks, payoffs respond discretely, and only jump once a sufficient amount of people have coordinated. These type of situations describe frameworks of regime change.

Global games of regime change describe coordination games of incomplete information in which the status quo – i.e. a currency peg, a bank's balance sheet, or a political regime, – is abandoned when a sufficient fraction of the population attacks it. So far, most previous work has treated the population as an infinite, homogeneous, mass of individuals, each with a private noisy signal of the fundamentals, abstracting from the potential patterns of communication that admittedly exist amongst individuals (See for instance: Angeletos et al. (2006), Bueno de Mesquita (2010), and Edmond (2013)). In this paper we propose a new approach to introduce networks of information transmission within this class of games. We show that the network structure imposes important considerations on the outcomes of the game.

Our model assumes a population of individuals connected according to a given (and fixed) network, each with an independent noisy signal on the underlying strength of the status quo. The state of the world describes the minimal fraction of individuals that are necessary for regime change. We assume that players observe the signal realization of their neighbors – including themselves – but do not observe the private information of others in the population. Each agent can then either choose to attack or not attack the status quo. Attacking can yield a positive payoff, if regime change is successful, or a negative payoff, if it is not. Not attacking always yields a 0 payoff. Payoffs are discontinuous both in the state and other players' actions so that the payoff structure does not exhibit strategic complementarities in the strict sense.

This model describes a game of incomplete information where the level of informational asymmetry corresponds to the underlying network structure, which is fixed. At the extreme, in which everyone is connected, information is symmetric and the structure of the game is common knowledge. However, the information structure also implies that players' posterior beliefs are locally correlated: if player $i's$ signal moves, so too do the beliefs of all neighbors of i, even if they are not directly connected. This means that certain signals are more "valuable" than others in coordinating individuals. At the extreme, if one player is connected to everyone, then her signal may be used as a coordinating device. In this sense, players' strategies cannot simply take a simple average of all incoming signals and attack if the resulting value is below some threshold. Instead, strategies may trade-off between two signals in a non-linear fashion, thus complicating the analysis. As in Bueno de Mesquita (2010)'s analysis of political vanguards, this paper also analyzes models of regime change with non-uniform populations, where the asymmtery is not in the payoff structure, but rather in individuals' network position.

To avoid dealing with intractable correlation effects that plague the system for a generic class of networks, we focus instead on networks where each agent's neighborhood is sufficiently small relative to the entire population, such that these correlation effects are negligible. In the end, we are left with an infinite population split into various partitions of varying connectivity (or degree) and therefore of varying precision in their private information. We identify the degree distribution of the network as a crucial determinant of individual strategies and aggregate behavior, and compare the equilibrium outcomes for various degree distributions ranked according to different measures of Stochastic Dominance. Moreover, we present some results on uniqueness of equilibrium that relate meaningfully to previous models with a homogeneous population.

In the case of a diffuse common prior (i.e. no public information) we prove that the probability of successful regime change does not vary with the degree distribution: As remarked by Vives (2005), Ashworth and Bueno de Mesquita (2006) and others, maximal strategic uncertainty with respect to others' behavior induces "flatter" best response so that individuals are less concerned with the aggregate composition of the population. It turns out that each individual responds to common beliefs about aggregate behavior by selecting a threshold strategy commensurate with the probability that their private signal zis extreme. This implies that players with smaller tails in their private signals compensate by selecting larger thresholds at exactly the proportions that offset any differences across individuals. As a result, everyone in the population (regardless of their degree) is equally propense to attack the regime, so that altering the proportion of highly connected individuals does not affect the aggregate share of belligerents. This does not mean, however, that all equilibria are identical. While the success probability remains unchanged, the size and composition of attacks conditional on success/failure indeed respond to the varying connectivity of the population. Populations with a larger (smaller) share of highly connected individuals will exhibit larger (smaller) successful attacks and smaller (larger) failed attacks.

Introducing public information has important implications for equilibrium outcomes: the degree distribution becomes an instrumental determinant of the success probability, albeit with surprising considerations. It turns out that the cost endured by attackers in the event of failure defines the direction of the comparative statics and sheds some light into the strategic interactions of a population with varying informativeness of the relevant state of the world. When costs are sufficiently low, increasing the connectivity of the population (by way of First Order Stochastic Dominance) actually lowers the success probability. The opposite is true for sufficiently high costs. This surprising result reflects the fact that less informed individuals place less weight on their private signal and respond more to aggregate behavior. As costs fall less connected individuals increase their propensity to attack far more than more connected agents. Moreover this difference increases and becomes arbitrarily large as the costs of revolt approach zero.

The model attains uniqueness for a larger set of parameter values than under a homogeneous population (i.e. without incorporating a communication network). This is explained by noticing that the presence of varying degrees in this model essentially translates into a convex combination of weights placed on the public signal. With a larger proportion of well-connected individuals (who pay less attention to the public signal) we obtain uniqueness for smaller prior variances than previous models allowed.

Finally the paper provides a methodological contribution to the literature by identifying sufficient conditions on network sparseness that allow for an approximation of large networks by an infinite population partitioned by their connectivity. Not only do we gain tractability and insight but we are able to isolate the effect of connectivity on informativeness by disregarding the local correlations induced by the network. We show these conditions are quite general and applicable to a wide range of network global games and we hope these can be useful for future research in the area.

2 Literature

Our paper mainly contributes to two general strains of literature: that on global games pioneered by Carlsson and Van Damme (1993) and Morris and Shin (2001), and on network models of information transmission. The paper essentially extends the static version of the global game of regime change put forth by Angeletos et al. (2007) in considering the role played by the exchange of private information within a network. Indeed the presence of local communication lays the ground for a number of additional questions on the role of connectivity in coordination not adressed in the basic model. Others, such as Edmond (2013) and Bueno de Mesquita (2010), have similarly dealt with discrete action global games in large populations, tackling diverse aspects such as the possibility of strategic action by the status quo, or by political vanguards.

In a particularly relevant study, Bueno de Mesquita (2010) presents an information model of revolutions with information transmission between motivated extremists and moderates. He finds that the model obtains multiple equilibria, and that structural factors may therefore influence the likelihood of regime change. We focus instead on settings that guarantee a unique equilibrium in threshold strategies, and we extend some of the insight of these models by considering the impact of connectivity in shaping equilibrium behavior. We find that the probability of regime change responds to the underlying structure of connections. Our paper shares with Hellwig (2002) and Angeletos and Werning (2006) the study of the interaction of private and public information in determining uniqueness of coordinating equilibria.

More precise efforts to model regime change with heterogeneous agents, such as Chwe (2000) and Guimaraes and Morris (2007), distinguish agents' possibly different action spaces contingent on types or their network position in sequential action games, but similarly make no effort to model varying connectivity and its role in the sharing of private information. Moreover, the latter focus on continuous action spaces, which disregards some of the inherent complications of correlated signals in threshold equilibria with finite players. We show that these considerations are not innocent, and that the strategic impact of connectivity on equilibrium outcomes is far from obvious. Most recently, an attempt to adress the role of networks by Dahleh et al. (2015) has provided a partial characterization for finite populations. Their results are silent to non-regular network structures and their focus on multiplicity is strangely at odds with the solution concept employed. Finally, Hassanpour (2010) provides an applied study that underscores the empirical importance of these types of models in recent experiences with large scale coordinated attacks on regimes. His theoretical model, however, allows for continuous belief updating à la deGroot, which fundamentally undermines the impact of limited local communication in coordinated attacks (DeGroot, 1974).

The network literature includes a rich tradition of modelling communication. Bloch and Dutta (2009) propose a model of network formation where agents can choose to invest in links of communication with varying degrees. Their work establishes stable and efficient architectures, rather than exploring the impact of exchange on games of coordination. Galeotti and Goyal (2010) consider a model where individuals are partially informed about the structure of the social network and provide results characterizing how the network structure shape individual behaviour and payoffs. Finally, Hagenbach and Koessler (2010) consider strategic communication in networks by modelling a cheap talk communication stage within networks.

We assume in this paper that communication is truthful and limited to direct network neighbors. This is a modelling assumption shared with Calvó-Armengol and De Martí (2009) that deals directly with the role of communication networks in a class of global games with continuous quadratic payoffs. They provide a knowledge index that essentially compounds higher-order expectations in order to map beliefs into actions. Regime change models, however, are discrete action games that require a consideration of the entire posterior distributions. As such, a new approach that resolves the underlying correlations is warranted.

The focus of this paper on information structures resembles other efforts to model information acquisition in large scale coordination models, such as Szkup and Trevino (2015), Myatt and Wallace (2012), and Hellwig and Veldkamp (2009). However, these papers model uniform populations with symmetric signal accuracy. We complement this literature by relating existing network structures to informational distributions and therefore to the probability of regime change. In this paper we explore signal heterogeneity, as in Sakovics and Steiner (2012), but from a very different perspective. We take a network view of varying signal strength, and we don't consider vanishing noise, obtaining a unique equilibrium that can be related precisely to the payoff structure.

Closer to this paper, recent work by Iachan and Nenov (2015) explore the relationship between the quality of information and fundamentals. They show that if payoffs (conditional on regime change) are sensitive to fundamental parameters (such as the strength of the regime) then less precise signals can induce greater probability of regime change. In this model we provide a network-based micro-foundation for varying signal qualities across players, and we show that the effect of signal precision on regime change exists even with conditionally fixed payoffs.

Finally, the recent paper by Barberà and Jackson (2016) characterizes the set of monotone threshold equilibria for a discrete version of a similar collective action game, with an infinite number of players but with a finite number of possible signals. Their model assumes that all individuals receive the same number of signals and, in terms of the network story of this paper, this could represent the case of a regular network where all players have exactly the same number of connections. We allow for more general information transmission structures.

3 Model

This section develops a model that builds on the original model of regime change by introducing a stage of communication in networks loosely inspired by the work of Calvó-Armengol and De Martí (2007). By approximating large networks with an infinite population the model essentially introduces heterogeneous variances to the original analysis by Angeletos et al. (2007) which allows for a different set of comparative statics exercises and, more importantly, introduces different informational roles across society's members in the spirit of Bueno de Mesquita (2010).

3.1 Actions, Payoffs and Network

There is a population N of individuals connected according to some network G to be specified below. Each agent takes an action $a_i \in \{0,1\}$ where $a_i = 1$ will represent an attack on the status quo. The payoffs are as follows:

	Regime Change ($A \geq \theta$)	Failure ($A < \theta$)
$a_i = 1$	$1 - c$	$-c$
$a_i = 0$	0	0

where θ is some exogenous parameter, $A = \frac{1}{N}\sum\limits_{n} a_i$ is the proportion of the population that chooses to attack the status quo and $c \in (0,1)$ represents the cost of attack.

There is a network G that captures the communication process. We assume $g_{ij} = g_{ji}$ (undirected channels) and $g_{ij} \in \{0,1\}$, with $g_{ij} = 1$ meaning that agents i and j communicate with each other. For computational simplicity me let $g_{ii} = 1$. We define the neighborhood of i as $N_i = \{j \in N | g_{ij} = 1\}$ and we denote its cardinality as the degree (that is, $d_i = |N_i|$). Let N_d represent the number of individuals in the network with degree d. Finally, let $\mathcal{D} = (d_i)_{i \in N}$ represent the set of degrees of all nodes in the network, and $D = max(\mathcal{D})$ be the maximum degree.[1] Notice that, contrary to other models of regime-change, conditional payoffs are fixed, and not a function of fundamentals, such as in Iachan and Nenov (2015). This distinction matters because we show here that the effects of different information quality distributions across the population need not depend on the sensitivity of payoffs to fundamentals, as is required in their setting.

[1]We will also use the typical notation in game theory where $D_{-i} = \{d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n\}$

3.2 Information, Communication, and Belief Formation

Agents have a common prior belief θ_0 with a corresponding variance σ_0^2, which can be diffuse or not. Each agent then receives an i.i.d. signal $x_i = \theta + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$. There is one round of truthful communication in which each agent transmits their signal to his/her neighbors.[2] Alternatively you could think of the network as describing a technology whereby being connected to someone implies that their private information is readily available. In any case, after the dispersal of information, each agent i contains a vector composed of d_i independent signals with which to update beliefs about the strength of the status quo, θ, using Bayes' rule. Clearly, in the measure that agents' signal vectors overlap, their posterior beliefs about fundamentals will correlate, meaning that agents that share most of their neighbors will also have very similar information; in the limit, a complete network corresponds to a situation of common knowledge.

Formally, agent i forms the following posterior distribution of θ conditioned on the entire vector of private signals,

$$\theta \mid \mathbf{x}_i \sim N\left(\frac{\sigma^2}{\sigma^2 + d_i\sigma_0^2}\theta_0 + \frac{\sigma_0^2}{\sigma^2 + d_i\sigma_0^2}\left\langle \mathbf{x_i}, \mathbf{1_{d_i}} \right\rangle, \frac{\sigma^2\sigma_0^2}{\sigma^2 + d_i\sigma_0^2}\right) \tag{1}$$

where $\langle \cdot \rangle$ represents the dot product of two vectors and $\mathbf{1_{d_i}}$ is a vector of ones of dimension d_i. This updating process is instrumental to equilibrium since it refines agents' beliefs about the probability of success (holding everyone's equilibrium behavior fixed) and therefore allows players to obtain their optimal best response.

3.3 Strategies

In this context a strategy is defined as a function, $a_i(\mathbf{x}_i) \in \{0, 1\}$, that maps a player's vector of signals into one of the two available actions. A natural equilibrium notion for our setup is Bayesian Nash equilibrium. At equilibrium each player is best-replying to others' strategies:

[2]We focus on only one round of communication for ease of exposition. There is nothing special about one round, and we can easily accomodate more rounds of communication by extending our currrent communication protocol. If there are, say, t rounds of communication, and we assume that any given player communicates all signals she knows to her neighbors, the ensuing information structure is equivalent to our current setup with a new (and denser) network where there is a link between any two nodes that were at most t links apart in the original network. As an example, with 2 rounds of communication we would define $G^{(2)}$ as the new primitive network structure, where $G^{(2)} = 1$ if and only if there exists a path of length at most 2 between i and j.

player i will choose to attack the status quo for all signal vectors in the set,

$$B_i(A) = \{\mathbf{x_i} \mid Pr\left(A \geq \theta \mid \mathbf{x_i}\right) \geq c\} \tag{2}$$

where the equilibrium size of attack, A, in terms of equilibrium strategies, is defined as,

$$A = \frac{1}{n}\sum_{i}^{n} a_i = \frac{1}{n}\sum_{i}^{n} \mathbf{1}_{\{\mathbf{x_i} \in B_i\}} \tag{3}$$

with $\mathbf{1}_S$ representing the indicator function over the set S.

We can now formally define a Bayesian Nash Equilibrium of this model as a situation where, given others' strategies B_{-i}, and given the vector of incoming signals, $\mathbf{x_i}$, player i forms beliefs about the size of attack, A, and chooses an optimal strategy, B_i.

Definition 1. *An equilibrium corresponds to a profile $B = (B_1 \ldots, B_n)$ such that equations (2) and (3) hold simultaneously.*

Notice that the definition of the equilibrium is circular: equilibrium strategies (characterized by the set of signal realizations, for each individual, that induces her to attack) depend on the aggregate A, which, in turn, depends on individuals' equilibrium decision to attack. To find the equilibrium we need to resolve such circularity, and this requires a number of steps.

First, the set $B_i(A)$ contains all vectors of signal realizations that induce agent i to attack the status quo. Recall, however, that agent i knows that some of these realizations are observed by other neighboring players. This means that if she observes, say, a very high realization for the signal of an influential neighbor, she can infer that many other players have observed it as well. Then, she might expect that the share of attackers, A, will be low and will therefore need to observe a much lower realizations in her remaining signals to convince her that a low size A is sufficient for regime change. In other words, the upper boundary of the set B_i can be highly non-linear, following the network structure and ensuing correlations. The following section works through an example of this sort, but in general these considerations complicate the analysis considerably, and it is precisely what we try to avoid with an infinite population approach.

Furthermore, the size of attack, A, is essentially a binomial random variable (or a sum of bernoulli random variables), but the correlations implicit in the network structure guarantees that these are not independent Bernoulli draws. This means that a player's position in the network admittedly affect her expectation about A. To see this notice that players are correlated amongst individuals up to two links apart (i.e. "I am correlated with my friend's friend since both of us received my friend's signal"). This guarantees that my belief about the possible states in which, say, $A = \frac{1}{3}$ is not the same as someone else with a different set of neighbors (and thus a different set of correlations). This seems to imply that any two individuals (even of the same degree) can arrive at radically different beliefs about the possible value of A. However, common knowledge of the network structure would guarantee that every player knows everyone's correlation structure when calculating their thresholds, so that in equilibrium every player would know each others' threshold strategies perfectly and in fact would end up calculating the exact same distribution for A. We are going to find sufficient conditions on the network geometry that make sure we don't have to worry about correlation effects. As a consequence, all necessary information to analyze regime change in large networks is going to be captured in the degree distribution.[3]

Finally, notice that nothing in the definition of the equilibrium hinges on the assumption of normality. Indeed, all we need is that players can form proper beliefs about θ, conditional on their private information, so that equation (2) is well defined. However, by assuming that signals are normally distributed we can use equation (1) to obtain convenient analytic expressions for posterior beliefs that allow us to establish the existence and characterization of a unique equilibrium. The advantage of the normality assumption is that we obtain simple linear estimates of the posterior mean, and we retain a full description of the posterior distribution with which to properly define $Pr(A \geq \theta \mid \mathbf{x}_i)$. Most of the classic global games literature has focused on cases where players receive only one signal. In this context, a unique Bayes equilibrium is typically defined in threshold strategies, in which players choose one action if their signal is below a certain threshold and play the other action if the signal is greater than this threshold (Morris and Shin, 2001; Vives, 2005) . Our problem here is multidimensional because each individual receives different signals from multiple sources (i.e. their direct neighbors in the network) and it is therefore more difficult to characterize threshold strategies – we could end up with a nonlinear frontier, for each individual, that characterizes the mapping between vectors of signals and actions. The linear estimates of the posterior mean in equation (1), together with the sufficient condition on correlation effects from 5, reduce the dimensionality of the problem, and provide a natural way to generalize threshold strategies to our setup: a player attacks if and only if the

[3]This model deals with very large networks and the idea that the entire geometry is somehow known by everyone is untenable. Fortunately it is not necessary. It suffices that players all know the degree distribution of the network and that they have a common prior belief about the likelihood of each particular architecture that is possible given this degree distribution.

average of signals received is below a given threshold. In the discussion section at the end of the paper we briefly mention how other distributional assumptions within the class of conjugate prior distributions could deliver similar results.

4 A Finite Network Example

In this section we seek to underline the correlation effects that guide equilibrium behavior in finite networks by working through an amenable example with three players. Given the difficulties in solving the model for a general finite network of size n, the next section will show that these correlation effects disappear for large (and sufficiently sparse) networks; this will allow us to solve the model asymptotically.

In order to solve the model with finite agents we must consider the possibility that players a priori will not weight all signals equally when defining equilibrium strategies, i.e. the upper boundary of set B_i is not linear. In the infinite population scenario agents simply take an average of all signals in anticipation of the negligible impact of correlations. This in turn means that the position in the network turns out to be irrelevant (all that matters is the degree of each player) and all incoming signals are equally useful when calculating posteriors. But with finite agents the position in the network is crucial in determining equilibrium strategies. As an example, a player might choose to weight one of his neighbor's signals more if this neighbor happens to be in a privileged position- i.e. a "hub"'s signal gets read by a large share of the total population. This implies that depending on others' equilibrium strategies, best response functions may take on different shapes corresponding to different weights placed on each signal. Formally, best responses here are not a linear mapping of all incoming signals (as is the case for infinite players) but instead will be shaped by the relative position of each neighbors who transmitted each signal.

As an example, consider the game described above played by three agents (call them 1, 2, and 3) connected in a star-like network as shown in Figure 7. It should be clear from the communication process that after the signals have dispersed through the network all agents' beliefs are correlated, and in particular agents 2 and 3 are correlated vis-a-vis 1's signal.[4] An equilibrium here corresponds to a vector of equilibrium strategies $B^\star = (B_1^\star, B_2^\star, B_3^\star)$ defined

[4]Because communication occurs for one round only, correlations emerge across players with at most 2 degrees of separation. If we added a fourth player 4 connected to 3, then 2 and 3 would lie 3 links apart and would not be correlated in their posterior estimates of θ

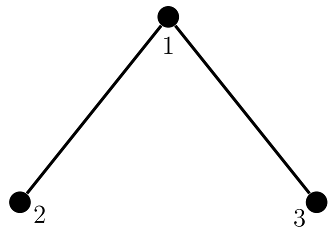

Figure 1: A Simple Three-Player Network

as in equation (2) and a share of agents that attack in equilibrium, A, given by the relation $A(\theta) = \frac{1}{3}\sum_i \mathbf{1}_{\left\{\mathbf{x_i} \in B_i^\star\right\}}$.

As the hub, 1's private information will take on a larger share of others' best response correspondences. To develop some intuition consider how player 2 finds her optimal strategy. Given the strategies of other players fixed, imagine 2 observes signals (x_1, x_2) such that x_1 is very large. Although player 2 has no direct contact with 3, she knows 3 has also observed this signal and can reason that it will drive 3's posterior beliefs about θ upward. Then, given 3's beliefs about A fixed, player 2 can argue that 3 will need to observe a very low realization of x_3 (the only remaining signal observed by 3) in order for $Pr(A \geq \theta \mid \mathbf{x}_3) \geq c$, which we know induces 3 to attack. In sum, a high realization of x_1 gives player 2 some confidence that 3 is very unlikely to attack, even as 2 and 3 are not directly connected. It should be clear that 2 can perform similar reasoning with respect to 1's equilibrium behavior ex-post. Now consider what happens when 2 observes a very low realization of x_1 instead. In this scenario 2 will argue that 3 will attack more often than before because, keeping others' strategies fixed, 2 can reason that this low realization of x_1 will induce 3's beliefs about θ downward and thus $P(A \geq \theta \mid \mathbf{x}_3) \geq c$ holds for larger realizations of x_3 than before. Together, these arguments suggest that low values of x_1 raise 2's belief about 3's (and 1's) propensity to attack, while high values of x_1 lower these beliefs.

This reasoning will affect 2's equilibrium strategy because it affects her belief about the aggregate size of attack A. Since networks here are small, forming expectations about other's equilibrium action will in fact affect the belief about A and therefore will impact best responses. To see this notice that our indifference condition $Pr(A > \theta \mid \mathbf{x_i^\star}) = c$, which defines the boundary of our set B_i, can be reformulated as

$$\int_{-\infty}^{\infty} Pr(A(\theta; B) \geq \theta)\, Pr(\theta \mid \mathbf{x}_i^\star)\, d\theta = c$$

where the dependence of A on players' equilibrium strategies $B = (B_1, B_2, B_3)$ is made explicit. After rearranging and noticing that necessarily $A \in [0,1]$ gives the following meaningful expression

$$\int_0^1 Pr\left(A\left(\theta; B\right) \geq \theta\right) Pr\left(\theta \mid \mathbf{x_i^\star}\right) d\theta = c - Pr\left(\theta \leq 0 \mid \mathbf{x}_i^\star\right) \tag{4}$$

Notice this equation essentially defines a best response correspondence $B_i\left(B_{-i}\right)$ for each i that is non-linear following the intuition above: when determining the set of vectors $\mathbf{x_i^\star}$ that satisfy the indifference condition, players must take into account that signal realizations will not only affect their inference of θ as determined by $Pr\left(\theta \mid \mathbf{x}_i^\star\right)$ but also how they will affect their beliefs about aggregate equilibrium behavior as defined by $Pr\left(A\left(\theta; B\right) \geq \theta\right)$. While all signals will contribute equally to statistical inference on the state of the world (the first effect), the previous discussion makes clear that signals will nonetheless have a varying impact on the belief about aggregate behavior in equilibrium (the second effect).

In Section 5 we show that, under certain conditions, this entire reasoning may be disregarded: the fact that 3 is more or less likely to attack (given an observation of a common signal x_1) will not affect 2's belief about $A(\theta)$ when the population tends to infinity.[5] As a result, there is no strategic reason for players to take into account the correlation structures generated by the network; Indeed, 1's signal is no more valuable than 2's signal if the information it provides about 3's equilibrium behavior does not move 2's beliefs about $A(\theta)$. More formally, since beliefs about $A(\theta)$ remain constant the above equation can be simplified by identifying regions $\left(-\infty, \hat{\theta}\right)$ where success occurs with probability 1 and regions $\left(\hat{\theta}, \infty\right)$ where success never occurs, so that the above equation can be rewritten as

$$\int_{-\infty}^{\hat{\theta}} Pr\left(\theta \mid \mathbf{x}_i^\star\right) = Pr\left(\theta \leq \hat{\theta} \mid \mathbf{x}_i^\star\right) = c \tag{5}$$

In short, when signals cannot inform players about aggregate behavior they merely serve to update beliefs about the underlying state θ, and they are all equally valuable in this regard.

Returning to 2's equilibrium strategy, consider how her reasoning above affects her best response correspondence. We have seen that a high realization of x_1 will not only force 2's posterior beliefs about θ upward, but it will also allow 2 to conclude that both 1 and 3 will now attack less

[5]In truth, it is not sufficient that the population tend to infinity. We must also ensure that no player remains too central as population grows or else her signal would indeed move beliefs on aggregate behavior. This sparseness condition will be specified in the next section.

often in equilibrium, so that aggregate behavior A goes down. This will make 2 more reluctant to engage in attacking the status quo, unless his remaining signal, x_2, is extremely favorable for success. What is important here is that although we could make the same argument for a high realization of x_2, this signal will only allow 2 to form beliefs about 1's equilibrium behavior, while x_1 allows 2 to form beliefs about the equilibrium behavior of both 1 and 3. In other words, 1's signal is more "informative" than 2's. The unequal weight of signals that emerges from correlations of beliefs in the network generates nonlinear strategies in finite networks. In the next section we present an asymptotic strategy to resolve this issue.

5 A Large Network Approximation

The reader should think that the network in this model generates two main effects: imbuing the system with correlation and allowing for the pooling of information. This paper focuses on the latter by assuming that we are in sufficiently large networks with only local correlations, such that they become strategically irrelevant. As a result, the only effect of the network is that agents with larger degree have more precise signals. Of course not all network architectures exhibit a sufficiently local correlation structures – consider as a counterexample the complete network where correlation is maximal across all players. We are therefore looking for a condition on network sparseness that guarantees sufficiently local correlation effects, which will induce a weak form of the Law of Large Numbers.

Recall that the equilibrium size of attack is given by the proportion of the population that takes action 1 in equilibirum, and can be written formally using our definition of the set B_i as,

$$A = \frac{1}{n} \sum_{j}^{N} 1_{\{\mathbf{x_j} \in B_j\}}$$

Notice that each for each $j \in N$ the random variable $1_{\{\mathbf{x_j} \in B_j\}}$ is a weakly dependent Bernoulli with probability parameter $Pr(\mathbf{x_j} \in B_j)$. The trick is to arrive at a convenient expression for the limiting size of attack that is symmetric on the degree of individuals: this expression will allow us to solve for the equilibrium as a function of the distribution of connectivities in the society, which vastly reduces the dimensionality of the network into a manageable object.

First, imagine we can apply a Law of Large Numbers (LLN) to the expression above to obtain the following convergence result,

$$A = \frac{1}{n}\sum_j 1_{\{\mathbf{x}_j \le B_j^\star\}} \underset{a.s.}{\longrightarrow} \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_j 1_{\{\mathbf{x}_j \le B_j^\star\}}\right] \tag{6}$$

This implies that, if indeed we can guarantee that a LLN applies, then the limiting size of attack converges to the limit of the expected proportion of attackers in the population. Notice that, as a limit, this expected proportion does not respond to individual changes in players' best responses. In particular, if player, say, k decides to change her equilibrium strategy from $B_k^\star$ to $\tilde{B}_k$, the value of A will remain fixed. Moreover, we show below that equation (6) implies all players' signals are observed by a vanishing fraction of the population. This effect implies that the identity of each incoming signal is irrelevant, so that in equilibrium players' will weight all signals equally in a simple average, and the upper boundary of B_i is therefore linear with a slope of -1. In other words, to maintain indifference between attacking or not, a low realization of one signal can be offset by a higher realization of another by exactly the same amount – something that did not occur in the finite case above. Formally, this means that strategies can now be formulated in terms of the average of all incoming signals,

$$B_j = \left\{\mathbf{x}_j \mid \frac{1}{d_j}\left\langle \mathbf{x}_j, \mathbf{1}_{\mathbf{d_j}}\right\rangle \le x_j^\star\right\}$$

where we have that $x_j^\star$ satisfies our indifference condition $P\left(A \ge \theta \mid x_j^\star\right) = c$.[6] With this, we can now rewrite the above expression as

$$A \underset{a.s.}{\longrightarrow} \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_j 1_{\{\bar{x}_j \le x_j^\star\}}\right]$$

where $\bar{x}_j$ refers to the average of all signals received by individual j, and where $x_j^\star$ corresponds to the equilibrium scalar value that defines j's attack threshold: j attacks if and only if the average of her signals is below this value.

We can further simplify the above expression. As argued above, the position in the network becomes irrelevant in the limit. Incoming signals can therefore only resolve fundamental uncertainty about θ, but are not able to resolve strategic uncertainty about the equilibrium size of attack, A.[7] Moreover, all signals are equally valuable in terms of forming posterior beliefs about

[6]Notice the probability now is conditioned on the average realization, rather than on the entire vector. A cursory look at equation (1) should reveal that these are equivalent formulations

[7]This distinction between *fundamental uncertainty* on the one hand and *strategic uncertainty* on the other is

θ, so the only relevant form of heterogeneity across players boils down to the number of signals they receive, which corresponds to their degree. This symmetry across players of the same degree arises from the fact we deal with asymptotically large populations where local correlations become strategically negligible. Formally, let N_d be the subset of players with degree d. Then we can extend the previous definition of A by imposing symmetric cutoff strategies for all players of a given degree, and with some manipulations we can write down the following set of relations:

$$A \underset{a.s.}{\longrightarrow} \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{d=1}^{D}\sum_{j\in N_d} 1_{\{\bar{x}_j \le \bar{x}_d^\star\}}\right] = \lim_{n\to\infty} \mathbb{E}\left[\frac{n_d}{n}\sum_{d=1}^{D}\frac{1}{n_d}\sum_{j\in N_d} 1_{\{\bar{x}_j \le \bar{x}_d^\star\}}\right] = \sum_{d=1}^{\infty} P_d\, Pr(\bar{x}_d \le \bar{x}_d^\star)$$

The first equality above comes from partitioning the total population n into the different possible degrees, while the second equality comes from applying the expectation and taking limits. The last sum goes to infinity because in an infinite population we cannot rule out infinite degrees.[8] The previous arguments allow us to conclude that, if indeed we can establish a LLN to equation (6) above, we would obtain the following limiting result:

$$A \underset{a.s.}{\longrightarrow} \sum_{d=1}^{\infty} P_d Pr(\bar{x}_d \le x_d^*) \tag{7}$$

where P_d is the share of the population with degree d.

The above arguments require that we may apply a LLN on equation (6). We now proceed to provide precise conditions on the network structure such that this is feasible. Notice that the elements summed in (6) are not independent random variables so a standard LLN does not apply. Instead we rely on a LLN for weakly dependent random variables. The weak correlation structures that allow for a LLN also require a minimum level of sparseness in the network. This is formalized in the following proposition that places conditions on the growth rate of the maximal degrees in order to assure that these limit properties hold.

Proposition 2. *Let $D_1 = max\,(\mathcal{D})$ and $D_2 = max\,(\mathcal{D}_{-i})$, where $i \in N$ is such that $d_i = D_1$. Then, if $D_1 \cdot D_2 \in o\,(n)$, the convergence result in equation (7) holds.*

Proof. See Appendix A for a proof of this result and a more general result on convergence for a

first made by Myatt et al. (2002), who demonstrate how "even as the fundamental uncertainty becomes smaller and smaller, strategic uncertainty can remain as large as ever" (page 410).

[8] Indeed, Proposition 2 places bounds on the growth rate of the maximal degrees such that this convergence result holds, but it does not rule out that the maximal degree tends to infinity. In fact, section 7.1 deals with scale free networks with degree distributions that follow a power-law, in which case the support is the natural numbers.

wider class of correlation structures. ■

By placing an upper bound on the growth rate of maximal degrees, Proposition 2 essentially forces us to concentrate on networks that are sufficiently sparse, in the sense that no one individual contains too many links relative the size of the population. Intuitively, if we are trying to coordinate with a large set of individuals, any local correlations that exist within a small subset of the population become strategically insignificant, and we can therefore disregard them when forming our optimal strategies. As a result, all incoming signals are equally valuable, and best-reply functions are linear. This is the key step in the model, and what allows us to proceed with an analytic solution that focuses on the network effect on precision of beliefs, and ignores correlations.

To get a better idea of the type of restriction imposed by Proposition 2, consider a family of networks that satisfies the sparseness requirements. Imagine that, due to some cognitive or behavioral limitations, individuals cannot realistically sustain more than some predefined number, $\bar{d}$, friendships at any given point in time. This implies that no one individual can remain influential as the network grows in size, and, therefore, correlations vanish asymptotically. We write this down as an immediate corollary to Proposition 2.

Corollary 3. *If $d_i \leq \bar{d}$ for some $\bar{d} \in \mathbb{N}^+$, and for all $i \in N$, then Proposition 2 holds.*

Another family of networks that satisfies the sparseness requirements of Proposition 2 is introduced in Section 7.1. This is the class of scale-free networks whose degree distribution follows a Power Law (i.e. $Pr(k) \propto k^{-\gamma}$). The value of γ defines the size of the right tail of the degree distribution, and we show that for sufficiently large γ, which implies the tail is thin enough, we satisfy Proposition 2 almost surely.

Let us also consider a case in which Proposition 2 fails: the star network of Figure 7. In the star, one individual called the "hub" is connected to everyone (and therefore has degree equal to n) while everyone else, called "spokes", are only connected to the "hub" (and therefore have degree equal to 2). Trivially, the maximal degree in this network grows linearly with n. Correlations here remain a crucial determinant of equilibrium for all population sizes: all "spokes" know that everyone else has observed the hub's signal, so the "hub"'s realization will move everyone's beliefs about $A(\theta)$ much more than their own individual realization. In other words, the frontier of B_i is not linear because correlation effects imply that not all signals are equally valuable.

If we are willing to focus on the networks prescribed by Proposition 2, we may proceed with the analysis, summarizing any network simply by its degree distribution. In other words, we are left with an infinite population that is partitioned into classes with proportions P_d for $d \in \mathbb{N}$. In this sense, this model essentially extends the standard global game of regime change to a population with heterogeneous variances.

6 Equilibrium With Diffuse Prior

This section proceeds with the equilibrium analysis, and we focus first on the case where priors are completely uninformative. In the next section we analyze the case where individuals hold particular prior beliefs about the state of the world θ. We analyze both sections separately because the equilibrium implications are very different.

We follow classic arguments on global games and super-modular games (Morris and Shin, 2001; Athey, 2002; Vives, 2005) that, under certain conditions on the relation between actions and payoffs, and for particular signal structures, the unique Bayesian Nash equilibrium is defined in terms of threshold strategies.[9] This implies that if player i receives a vector of signal realization below some frontier she takes action 1, and conversely takes action 0 if she observes a private signal above. Of course, precisely at the threshold, player i must be indifferent between the two available actions, which, given our payoff structure, can be written as, $P\left(A \geq \theta \mid \mathbf{x_i}\right) = c$. As such, player i will choose to attack the status quo for all signal vectors in the set,

Agents have a diffuse prior over the state of the world (i.e. θ is distributed uniformly over the real line). Next, suppose there is a degree-specific threshold strategy $x_d^\star \in \Re$ such that each agent, i, with degree d attacks if and only if $\bar{x}_i \leq x_d^\star$. Then, the right-hand side of expression (7) transforms into,

$$A(\theta) = \sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}\left(x_d^\star - \theta\right)\right) \tag{8}$$

where Φ is the CDF of the standard normal. Because $A(\theta)$ is decreasing and continuous in θ, there exists a unique value $\hat{\theta}$ that is a fixed point of $A(\theta)$. Formally, $\hat{\theta}$ solves,

$$A(\hat{\theta}) = \sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}\left(x_d^\star - \hat{\theta}\right)\right) = \hat{\theta} \tag{9}$$

[9] For a complete list of these conditions see Morris and Shin (2001). See, for example, Morris and Shin (1999) and Sakovics and Steiner (2012) for the analysis of setups that are closest to our own modeling assumptions.

Finally, notice that there is regime change whenever $\theta \leq \hat{\theta}$. Standard Bayesian updating implies that the posterior expectation for an agent i of degree d that receives a signal realization x_i is $\theta \mid x_i \sim N(x_i, \frac{\sigma^2}{d})$. Therefore, to this particular agent, the probability of regime change is given by, $Pr(\theta \leq \hat{\theta} \mid x_i) = \Phi\left(\frac{\sqrt{d}}{\sigma}\left(\hat{\theta} - x_i\right)\right)$. The agent will find it optimal to attack whenever the posterior probability of regime change is greater than the marginal cost from attacking c, or whenever, $x_i \leq x_d^\star$ where $x_d^\star$ solves,

$$\Phi\left(\frac{\sqrt{d}}{\sigma}\left(\hat{\theta} - x_d^\star\right)\right) = c \qquad for \;\; d \in \mathbb{N} \tag{10}$$

A Bayesian Equilibrium is a sequence, $\left(\hat{\theta}; x_1^\star, x_2^\star, \dots\right)$ that solves equations (9) and (10). Notice players are strategic, and respond to everyone else's strategy through the parameter $\hat{\theta}$. We are essentially describing a coordination game of incomplete information with heterogeneous precisions.

Strategic response to others' strategies does not mean, however, that we face pathological, corner equilibria where everyone chooses never (always) to attack. Consider the response of individual i to a population where every other player chooses never to attack (i.e. everyone chooses $x_j^\star = -\infty$). In that case $A(\theta) = 0$ for all values of θ so that $\hat{\theta} = 0$. But consider introducing this to equation (10); Because $c > 0$ player i will choose an equilibrium threshold $x_i^\star$ bounded away from $-\infty$, so that never attacking is not an equilibrium strategy. In fact this will constitute an equilibrium if and only if $c = 1$. Intuitively, even though everyone else is choosing not to attack, there is still a positive probability that $\theta \leq 0$ (recall that player i's posterior belief about theta is $\theta \mid x_i \sim N(x_i, \sigma^2)$) so player i will choose to attack for some positive probability as a result.[10] This is true as well when considering the equilibrium where everyone always attacks (i.e. this will only be an equilibrium for $c = 0$). Together this implies we have a unique equilibrium. Our first result establishes the existence of this unique equilibrium by explicitly solving for $\hat{\theta}(c, \mathbf{p})$ in equation (9), and then establishes its lack of response to the degree distribution.

Proposition 4. *There exists a unique equilibrium to this static one-shot game in monotone strategies where* $\hat{\theta} = 1 - c$. *The probability of regime change,* $Pr(\theta < \hat{\theta})$, *does not change with*

[10] With a bounded distribution in private noise corner equilibria would be retrieved. To see this imagine instead that $\epsilon_i \sim U(\underline{\epsilon}, .\bar{\epsilon})$. Players' posterior beliefs about θ would be $\theta \mid x_i \sim U(x_i - \underline{\epsilon}, x_i + \bar{\epsilon})$. If $\underline{\epsilon}$ and $\bar{\epsilon}$ are sufficiently close, we cannot rule out a situation where all players believe the probability that $\theta = 0$ is 0 (alternatively 1), so that each player reacts to $\hat{\theta} = 0$ ($\hat{\theta} = 1$) by choosing an equilibrium strategy consistent with never (always) attacking the regime, which means choosing $x_i^\star = -\infty$ ($x_i^\star = +\infty$).

degree distribution $\boldsymbol{p}$ *and decreases with* c.

Proof. First rewrite equation (10) to get the best reply function for a player i of degree d,

$$x_d^\star = \hat{\theta} - \frac{\sigma}{\sqrt{d}}\Phi^{-1}(c) \tag{11}$$

Then substituting this into equation (9) gives

$$A(\hat{\theta}) = \sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}\left(\hat{\theta} - \frac{\sigma}{\sqrt{d}}\Phi^{-1}(c) - \hat{\theta}\right)\right) = \sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}\left(-\frac{\sigma}{\sqrt{d}}\Phi^{-1}(c)\right)\right)$$

$$= \sum_{d=1}^{\infty} P_d \cdot \Phi\left(-\Phi^{-1}(c)\right) = \sum_{d=1}^{\infty} P_d\left(1 - \Phi\left(\Phi^{-1}(c)\right)\right) = \sum_{d=1}^{\infty} P_d(1-c) = 1 - c = \hat{\theta}$$

Then $\hat{\theta} = 1 - c$ is independent of the degree distribution $(p_1, p_2, \ldots)$ and clearly decreases with c. This implies that the probability of regime change $Pr\left(\theta \leq \hat{\theta}\right)$ is also invariant to the degree distribution. Finally, we can calculate threshold strategies $(x_1^\star, x_2^\star, \ldots)$ that define the equilibrium by substituting in for $\hat{\theta}$ to obtain $x_d^\star = 1 - c - \frac{\sigma}{\sqrt{d}}\Phi^{-1}(c)$. ■

This result at first glance is not at all intuitive. In fact, we would expect the degree distribution to have an impact on the probability of regime change. It turns out that the range of θ where attacks are successful does not change as we alter the average connectivity of the society. Why? With a diffuse prior, players pay no attention to prior information; only private signals feed the inference of posterior beliefs. As a result, the strategic uncertainty of each agent is maximal with respect to the behavior of others. Formally, each player's higher-order beliefs on θ characterized by equation (10) have the same shape as the commonly observed distributions of private signals that sum in equation (9). Players have no way of improving on these beliefs. So even though different degrees select different threshold strategies, it turns out that precisely at the value of θ where the size of attack is the smallest successful attack, the propensity to take action is identical across the entire population. This is important because it implies that shuffling the distribution of degrees will not modify where the smallest successful attack is defined, so it will not modify the range of θ where successful attacks begin. As a result, the probability of observing a success will also remain fixed.

It is important to stress that threshold strategies are not identical. In fact threshold strategies depend crucially on d. However, the behavior of the population in equilibrium cannot be glimpsed

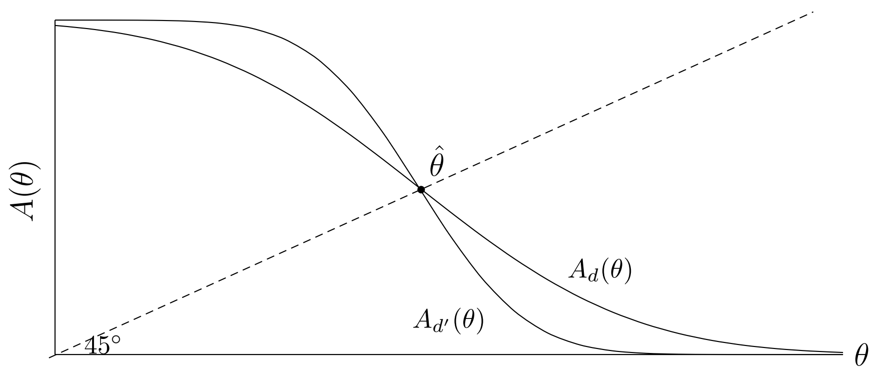

Figure 2: Size of Attack as a function of θ for $d' > d$

directly from these values. Instead, the share of individuals of degree d that decide to attack the status quo in equilibrium is defined by $Pr\left(x_d \leq x_d^\star \mid \theta\right)$, and this need not obey the ordering of threshold strategies $x_d^\star$. To see this consider the case of $c > \frac{1}{2}$. Players with a larger degree have a tighter posterior of θ and will choose a strictly higher threshold defined by $x_d^\star = \hat{\theta} - \frac{\sigma}{\sqrt{d}}\Phi^{-1}\left(c\right)$. However, notice that for the same threshold $x^\star$, these high degree individuals will observe $x \leq x^\star$ less often than low degree individuals. It turns out that at $\hat{\theta}$ high degree players will choose a threshold that is larger by the amount that exactly compensates the lower probability of observing a signal to the left of said threshold. As a result, even though it is true that $x_1^\star < x_2^\star < \cdots$, in equilibrium we have that $Pr(x_1 \leq x_1^\star \mid \hat{\theta}) = Pr(x_2 \leq x_2^\star \mid \hat{\theta}) = \cdots$. This striking result comes from the fact that the shape of the posterior belief about θ (which chooses the threshold $x_d^\star$) and the shape of the distribution of signals (which determines $Pr\left(x_d \leq x_d^\star\right)$) are equal and they therefore cancel out. Once we introduce a prior with finite variance players make us of it as a public coordinating device and higher order beliefs about θ will depart from higher order beliefs about $x_d \sim N\left(\theta, \frac{\sigma^2}{d}\right)$. As we will see, this affects the above result. However, too little strategic uncertainty in the form of too low prior variance, leads to multiplicity by increasing strategic complementarities of the model.

Proposition 4 should not convey the idea that the propensity to attack is independent of the degree distribution. This is only true at the point $\hat{\theta} = A\left(\hat{\theta}\right)$, which as mentioned above implies that the likelihood of observing a successful attack remains fixed for all $\mathbf{p}$. However, the equilibrium is defined for all values of $\theta \neq \hat{\theta}$. In these cases, the size of the attack responds directly to the weights chosen for each degree. In other words, conditional on the attack being too small to be successful, its size will vary with the degree distribution. The same applies for attacks large enough to be successful. In order to gain some intuition, Figure 2 plots the size of attack $A\left(\theta\right)$ for two different populations: one with $p_d = 1$ and the other with $p_{d'} = 1$ (where

$d' > d$). You can think that a population with positive weight on both these degrees will have an $A(\theta)$ line somewhere in between. What is important to note first is that Proposition 4 can be thought of as stating that these two lines (and in fact all other lines for all degrees $d \in \mathbb{N}$) intersect at the 45° line. It should be clear that any convex combination of these two functions (i.e. for any degree distribution) will always cross the 45° at the same point of intersection $\hat{\theta} = 1 - c$. For all values different form $\hat{\theta}$, however, the two curves take on quite different values and it is here where the degree distribution will determine the size of attacks. This is the content of our next result.

Proposition 5. *Define $A_{\mathbf{p}}(\theta)$ as the equilibrium size of attack under degree distribution $\mathbf{p}$ and state of the world θ. Let $\mathbf{p}'$ FOSD $\mathbf{p}$ then:*

- *for $\theta > \hat{\theta}$ (failure) $A_{\mathbf{p}}(\theta) > A_{\mathbf{p}'}(\theta)$ and $A_{\mathbf{p}}(\theta) - A_{\mathbf{p}'}(\theta)$ increases with θ*
- *for $\theta < \hat{\theta}$ (success) $A_{\mathbf{p}}(\theta) < A_{\mathbf{p}'}(\theta)$ and $A_{\mathbf{p}}(\theta) - A_{\mathbf{p}'}(\theta)$ increases with θ*

Moreover, this effect is largest whenever $c = \frac{1}{2}$ and decreases monotonically as $c \to \{0, 1\}$

Proof. Consider the equilibrium size of attack by plugging in equilibrium threshold strategies found in Proposition 4 into the definition of $A(\theta)$

$$A(\theta) = \sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}\left(1 - c - \frac{\sigma}{\sqrt{d}}\Phi^{-1}(c) - \theta\right)\right) =$$

$$\sum_{d=1}^{\infty} P_d \cdot \Phi\left(\frac{\sqrt{d}}{\sigma}(1 - c - \theta) - \Phi^{-1}(c)\right)$$

Notice that whenever $\theta < \hat{\theta}$ ($\theta > \hat{\theta}$) then the factor multiplying $\frac{\sqrt{d}}{\sigma}$ above is positive (negative) so that the argument of Φ is larger (smaller) for $d' > d$. Then, since Φ is an increasing function, by shifting weight to larger degrees (FOSD) we increase the weights on those terms in the summation that are larger (smaller) leading to a total value of $A(\theta)$ that is larger (smaller). Moreover, as we increase the value θ, this effect becomes larger if $\theta > \hat{\theta}$ (since the factor multiplying $\frac{\sqrt{d}}{\sigma}$ becomes larger in absolute value) and this effect becomes smaller if $\theta < \hat{\theta}$ (since the factor multiplying $\frac{\sqrt{d}}{\sigma}$ converges to zero). For the second part of the proof notice that for $c = \frac{1}{2}$ we have $\Phi^{-1}(c) = 0$ and that as c tends to the extremes, $\Phi^{-1}(c)$ tends to $\pm\infty$. ∎

Proposition 5 essentially states that in a population with equal weights, unsuccessful attacks are composed by a majority of less connected individuals, while successful attacks are composed by a majority of more connected individuals. This confirms our intuition that more connected individuals, because they are more informed, miscoordinate less often. After all, they obtain a more precise estimate of the true parameter, so it only makes sense that once θ is too large to guarantee success they retreat from attacking in larger shares. Graphically, you can see that for $\theta > \hat{\theta}$ the slope of $A(\theta)$ is steeper for the more connected individual.

What seems harder to reconcile, however, is that the difference in performance across degrees diminishes monotonically as the costs of revolt become more extreme. The intuition here is that the informational advantage of more informed individuals is greatest when the costs of attack are the least extreme. In other words, for costs neither too high nor too low more informed individuals shirk from attacking much more quickly as θ rises, creating a large advantage vis-a-vis the less informed. But for extreme costs these individuals respond less to the value of θ and choose to attack more or less the same for all state of the world (after all, either the costs are so low that attacking is almost always a better option, or too high to attack regardless of the state of the world).

7 Equilibrium With Non-Diffuse Prior

Next we turn to the case where player's hold a finite-variance prior about the state of the world, and must therefore incorporate it into their posterior beliefs. It turns out that in this scenario the previous results change dramatically. In particular, with the presence of a public signal the degree distribution completely determines the probability of success. Most interestingly, this comparative static is not monotone- so that more average connectivity means a greater probability of success- and instead depends on the cost of attack, c. When the costs are high more connectivity translates to a larger share of success, but the opposite is true for sufficiently low costs.

To begin the analysis, notice that if the prior is not diffuse, and instead follows a normal distribution,

$$\theta \sim \mathcal{N}(\theta_0, \sigma_0^2)$$

then the posterior distribution of the state of the world, θ, looks like,

$$\theta|x_i \sim \mathcal{N}(\frac{d_i\sigma_0^2 x_i + \sigma^2\theta_0}{d_i\sigma_0^2 + \sigma^2}, \frac{\sigma_0^2\sigma^2}{d_i\sigma_0^2 + \sigma^2})$$

The computation of $A(\theta)$ remains unaffected. However, the latter part of the analysis changes because posterior beliefs about θ now must trade off public and private signals according to their relative qualities. This leads to a modified expression for the cutoff decision of an individual of connectivity equal to d, which is now given by,

$$x_d^* = (1 + R_d)\hat{\theta} - \frac{\sigma}{\sqrt{d}}\sqrt{1 + R_d}\Phi^{-1}(c) - R_d\theta_0. \tag{12}$$

where we define,

$$R_d = \frac{\sigma^2}{d\sigma_0^2}$$

Notice that this expression resembles the cutoff expression for diffuse priors in section 6, except for two main differences. The first difference is the term $1 + R_d$ that replaces the term 1 in equation (11). Notice that this term depends on the relative precision of public and private signals, and tends to zero if the public signal is noisy relative to the private one. The second difference is that now cutoff strategies depend on the value of the common signal (or prior), and for larger values of θ_0, individuals will attack less often all else equal.

Plugging expression (12) in the equality $A(\hat{\theta}) = \hat{\theta}$ of equation (9) leads to the following fixed point representation of $\hat{\theta}$,

$$\sum_d P_d \Phi\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_d}\Phi^{-1}(c)\right) = \hat{\theta} \tag{13}$$

An Equilibrium is a $D + 1 - tuple$ $\left(\hat{\theta}, (x_d^\star)_{d=1}^D\right)$ that simultaneously solves equations (12) and (13).

Equilibrium Analysis

One first thing to notice is that, in equation (12), as $\sigma_o^2 \to \infty$, the equilibrium strategy for all degrees $x_d^\star$ tends to $\hat{\theta} - \frac{\sigma}{\sqrt{j}}\Phi^{-1}(c)$ (the solution to section 6) and that similarly $x_d^\star$ tends to $\hat{\theta}$ for sufficiently large degrees. This seems to suggest that as the public signal becomes more noisy and agents switch to the private signal, then the same effect as in diffuse prior starts kicking in. In other words the marginal effect of degree on the weights placed on each signal disappears. In a sense, we can think of the degree distribution as determining how strongly the population will value their private signal against the public signal.

Before we proceed to establish the effect of the degree distribution on the size and outcome of coordinated attacks, we shall establish the existence of a unique equilibrium. Unlike the previous section, the presence of a public signal means we can only guarantee uniqueness above a minimum public variance. The reason rests on arguments from Morris and Shin (2001). As the public signal variance diminishes, all players will shift their posterior beliefs towards said signal, in effect increasing the level of correlated beliefs and, in a sense, continuously increasing the level of common knowledge. At a certain point the level of strategic uncertainty is sufficiently low to generate multiplicity. The following result establishes the lower bound on public variance to guarantee uniqueness.

Proposition 6. *A unique equilibrium exists if* $\sigma_0^2 > \frac{\sigma}{\sqrt{2\pi}} \sum\limits_d \frac{p_d}{\sqrt{d}}$

Proof. We want to show that there exists a $\hat{\theta}$ that solves equation (13). Rewrite equation (13) as $F\left(\theta;\theta_0,\sigma^2,\sigma_0^2\right)=0$ where

$$F\left(\theta;\theta_0,\sigma^2,\sigma_0^2\right) = \sum_d P_d \Phi\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\left(\hat{\theta}-\theta_0\right) - \sqrt{1+R_d}\Phi^{-1}\left(c\right)\right) - \theta$$

Note that $F\left(\theta;\cdot\right)$ is continuous and differentiable in $\theta \in (0,1)$ and that

$$F(0;\cdot) = \sum_d p_d \Phi\left(\frac{-\theta_0\sigma}{\sigma_0^2\sqrt{d}} - \sqrt{1+R_d}\Phi^{-1}\left(c\right)\right) > 0$$

and

$$F(1;\cdot) = \sum_d p_d \Phi\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\left(1-\theta_0\right) - \sqrt{1+R_d}\Phi^{-1}\left(c\right)\right) - 1 < 0$$

Then, we need to show F is monotonically decreasing in θ. Notice that

$$\frac{\partial F}{\partial \theta} = \sum_d p_d \phi\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\left(\theta-\theta_0\right) - \sqrt{1+R_d}\Phi^{-1}\left(c\right)\right)\cdot\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\right) - 1$$

and given that $\underset{\hat{\theta}}{max}\phi\left(\cdot\right) = \frac{1}{\sqrt{2\pi}}$ the condition $\frac{1}{\sqrt{2\pi}}\frac{\sigma}{\sigma_0^2}\sum\frac{p_d}{\sqrt{d}} < 1 \implies \sigma_0^2 > \frac{\sigma}{\sqrt{2\pi}}\sum\frac{p_d}{\sqrt{d}}$ is both necessary and sufficient for F to be monotonic in θ. ■

This result suggests that introducing heterogeneity in the model retrieves uniqueness for a greater range of parameter values. This is to be expected since the presence of a degree distribution essentially introduces a mass of individuals with a lower private variance with respect to the homogeneous population scenario. The presence of more informed individuals strengthens the role of private information, making coordination more difficult. In other words, the presence of varying degrees in this model essentially translates into a convex combination of weights placed on the public signal. As more weight is placed on high degree individuals (who in turn pay less attention to the public signal) then we retain uniqueness for smaller public-signal variances than the previous models allowed.

There is no explicit analytical solution to equation (13) of the form $\hat{\theta}\left(c, \mathbf{p}\right)$ that would describe the impact of $\mathbf{p}$ on the probability of regime change. However, we can use the implicit function theorem to say something about the comparative statics across various types of degree distributions. We will focus on the set of parameter values that guarantees uniqueness (i.e. $\sigma_0^2 > \sigma$). The following result establishes a surprising non-monotone comparative statics for a population composed of only two arbitrary values of degrees, and a completely general class of degree distributions.

Proposition 7. *Let $\sigma_0^2 > \sigma$ and $D = 2$. Define two general degree distributions $\mathbf{p} = (p_1, (1 - p_1))$ and $\mathbf{p}' = (p_1', (1 - p_1'))$ such that $\mathbf{p}$ FOSD $\mathbf{p}'$ (i.e. such that $p_1 < p_1'$). Then, there exists a threshold $\hat{c} \in (0, 1)$ such that*

- *for all $c \in (0, \hat{c})$ the probability of regime change is lower under $\mathbf{p}$ than under $\mathbf{p}'$*
- *for $c \in (\hat{c}, 1)$ the probability of regime change is larger under $\mathbf{p}$ than under $\mathbf{p}'$.*

Moreover, $\hat{c}\left(\sigma, \sigma_0, \theta_0\right)$ decreases with θ_0 and increases with $\frac{\sigma_0^2}{\sigma^2}$.

Proof. Rewrite equation (13) for $D = 2$

$$p_1 \Phi\left(\frac{\sigma}{\sigma_0^2}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_1}\Phi^{-1}\left(c\right)\right) +$$

$$(1 - p_1)\Phi\left(\frac{\sigma}{\sigma_0^2\sqrt{2}}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_2}\Phi^{-1}\left(c\right)\right) = \hat{\theta}$$

define an implicit function $F\left(\hat{\theta}\left(p_1\right), p_1\right) = 0$. Applying the implicit function theorem, as before, we obtain,

$$\frac{\partial \hat{\theta}}{\partial p_1} = -\frac{\frac{\partial F}{\partial p_1}}{\frac{\partial F}{\partial \hat{\theta}}} =$$

$$-\frac{\Phi\left(\frac{\sigma}{\sigma_0^2}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_1}\Phi^{-1}\left(c\right)\right) - \Phi\left(\frac{\sigma}{\sigma_0^2\sqrt{2}}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_2}\Phi^{-1}\left(c\right)\right)}{\sum\limits_d p_d \phi\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\left(\hat{\theta} - \theta_0\right) - \sqrt{1 + R_d}\Phi^{-1}\left(c\right)\right)\left(\frac{\sigma}{\sigma_0^2\sqrt{d}}\right) - 1}$$

Whenever $\sigma_0^2 > \sigma$ the denominator is negative (see previous proof). As a result, the sign of the comparative static is determined entirely by the sign of the numerator. Notice that

$$\frac{\partial\hat{\theta}}{\partial p_1} = 0 \implies \frac{\sigma}{\sigma_0^2}\left(\hat{\theta} - \theta_0\right) - \sqrt{1+R_1}\Phi^{-1}(c) = \frac{\sigma}{\sigma_0^2\sqrt{2}}\left(\hat{\theta} - \theta_0\right) - \sqrt{1+R_2}\Phi^{-1}(c)$$

rearranging we get:

$$\Phi\left(\frac{\left(\hat{\theta} - \theta_0\right)\left(\frac{\sigma}{\sigma_0^2} - \frac{\sigma}{\sigma_0^2\sqrt{2}}\right)}{\sqrt{1+R_1} - \sqrt{1+R_2}}\right) = c$$

where $\hat{\theta}$ is endogenously determined in equilibrium and decreases with c (check equation (13)). Since $\Phi(\cdot)$ is a continuous, monotone function defined over the interval $[0,1]$ and moves positively with $\hat{\theta}$, we can be sure there exists a unique $\hat{c}$ that solves,

$$\Phi\left(\frac{\left(\hat{\theta}(\hat{c}) - \theta_0\right)\left(\frac{\sigma}{\sigma_0^2} - \frac{\sigma}{\sigma_0^2\sqrt{2}}\right)}{\sqrt{1+R_1} - \sqrt{1+R_2}}\right) = \hat{c}$$

Finally, notice that for all $c < \hat{c}$ the left hand side of this equation is larger than the right hand side so that $\frac{\partial\hat{\theta}}{\partial p_1} > 0$ and for all $c > \hat{c}$ the left hand side is smaller than the right so that $\frac{\partial\hat{\theta}}{\partial p_1} < 0$, thus proving the result. ■

The non-monotonicity implied by Proposition 7 is surprising. Indeed, increasing the average connectivity of the population *does not* increase the likelihood of success unambiguously. Instead, a low cost of failure increases the marginal propensity to attack of low connected individuals by far more than the corresponding increase experienced by highly connected individuals. As a result, low connected players choose to attack more often for a greater range of θ values, including the value $\hat{\theta}$ that determines the likelihood of success. It is still true (as in Proposition 5) that all failed attacks will register greater participation by less connected (less informed) individuals. But when costs are low some successful attacks also will contain greater shares of low connected players, so the overall likelihood of success increases.

Intuitively, the low risk involved in attacking the status quo gives less informed individuals an advantage by being more "reckless" (i.e counting more on rare tail events). Anticipating this behavior, the minorities of more informed individuals respond by attacking more than their signals would normally prescribe: equilibrium threshold strategies respond to $\hat{\theta}$, which captures the strategic response of players to aggregate behavior. Of course, as more informed individuals

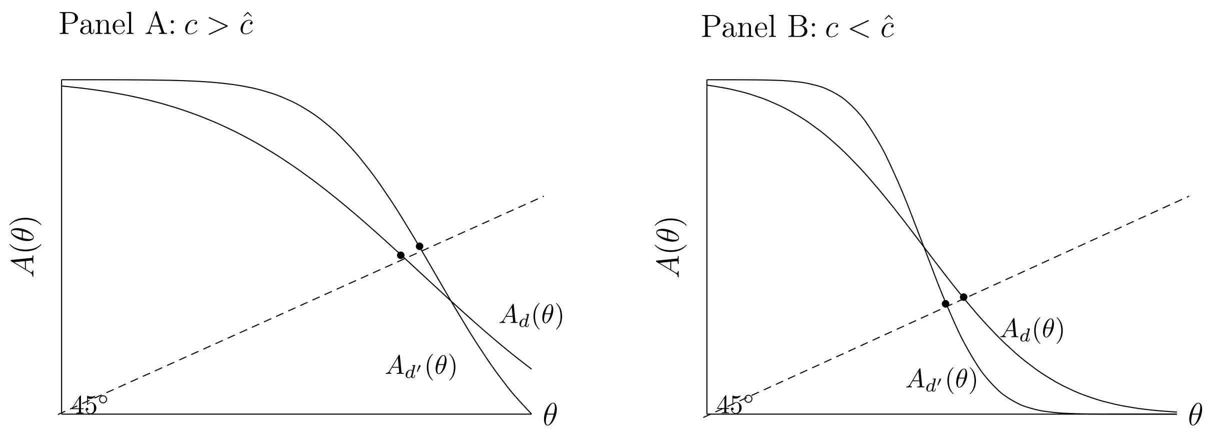

Figure 3: Size of Attack as a function of θ for $d' > d$.

cease to be a minority, the opposite will occur and less informed individuals will strategically respond to expected aggregate behavior: they will choose to be more cautious about attacking the regime than their low-quality information would prescribe.

To provide a better intuition, Figure 3 plots the size of attacks as a function of θ for both low high costs. You can see in the right-hand panel that for costs sufficiently low, the fraction of attackers for the low degree case d is in fact superior to the fraction for the high degree case d' for all values of $\theta \geq \hat{\theta}$ corresponding to failed attacks (as was the case in the previous section). However, now this fraction of attackers is also greater for some range of values of $\theta \leq \hat{\theta}$, corresponding to cases of successful attacks. In other words, less connected players surpass more connected players in their shares of attack at a lower value of θ than was the case with no prior information. Therefore, if the shares of low connected individuals increase the value of $\hat{\theta}$ will increase, and so will the overall probability of regime change.

Notice, however, that increasing the probability of regime change is not without a cost, and the welfare implications are far from obvious. Indeed, although shifting the degree distribution alters the probability of regime change, it also affects the relative fractions of participants reaping success or enduring failure. For instance, for the high cost case on the left panel of Figure 3, notice that while increasing connectivity raises the probability of regime change, it also raises the fraction of the population that attacks when regime change fails (i.e. for those values of θ to the right of the 45° line and to the left of the point where the two curves intersect). This larger share of attackers must incur a negative payoff of $-c$ for this range of θ, so the final effect on total welfare of moving from d to d' is therefore not obvious (a similar and converse argument can be made for the low cost case on the right panel).

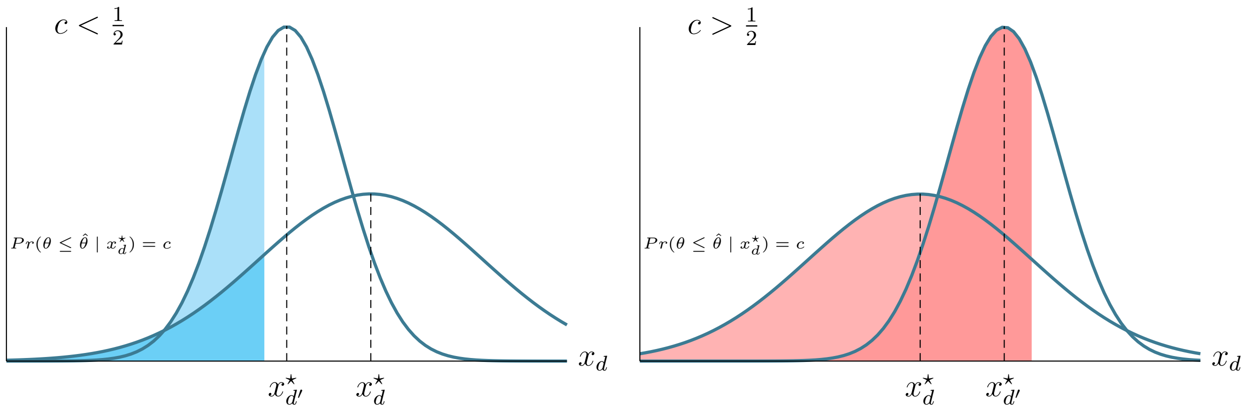

Figure 4: Less Informed Players choose larger $x_i^\star$ for c sufficiently low

Why are players with less precise posterior beliefs more willing to act when the costs are low? A look at figure 4 reveals that this result comes from the symmetry of distributions. Less connected individuals attain a more dispersed posterior belief about the state of the world, θ, and therefore must center their distributions further away from the critical cutoff $\hat{\theta}$ in order to attain a probability of success equal to c. When $c > \frac{1}{2}$ (Right-hand panel) the cutoff will certainly lie to the right of every player's expected belief, and less connected players will therefore choose a lower equilibrium threshold (With a lower threshold the propensity to attack is lower). But when $c < \frac{1}{2}$ (Left-hand panel) the cutoff will lie to the left of the distribution's center. In that case, more dispersed distributions will locate further to the right than less dispersed ones. As a result, low degree individuals will choose an equilibrium threshold that is larger (i.e. higher propensity to attack). Moreover this effect increases with the value of c and the distance between the threshold strategies for high and low-degree players can be made arbitrarily big. If costs are sufficiently low, so that this distance is very large, less connected individuals can be much more prone to attack and value of $\hat{\theta}$ is larger with less connected populations. (see figure 3).

Intuitively, the high expected gains and low expected losses that come with low values of c mean that players need not be very sure of the probability of success (remember at equilibrium players choose threshold such that probability of success equals c). In that case being less informed has a strategic advantage in so far as more weight is given to rare events, and these rare events are now enough to trigger action. Then as a low connected individual you will attack unsuccessfully more often than others, but you will also sometimes attack successfully more often than others. As a result, the status quo will need to be stronger to survive a population with your equilibrium behavior.

7.1 Power Law Degree Distribution

The previous result showed that the model exhibits non-monotone comparative statics with respect to the degree distribution for a general class of distributions and $D = 2$. Of course we would like to say something for a wider support of the degree distribution. The difficulty emerges in expressing a FOSD shift in one single parameter. We need this since our proofs rest on totally differentiating an implicit function in θ. Indeed for a great many number of distributions this cannot be done.

Following a vast number of studies that have documented the prevalence of scale-free characteristics across most types of large networks, we focus on degree distributions where the probability that a vertex is connected to k other vertices decays as a power law following $Pr(k) \propto k^{-\gamma}$, for $k \geq 1$.[11] This type of scale-free distributions have been shown to accurately describe the behavior of online social networks, such as twitter, that proved instrumental in transmitting information prior to large scale mobilization and regime change in the Arab Spring.[12]

We must first guarantee that this type of scale-free networks satisfy the sparseness requirements in Proposition 2. Following Barabási (2016), it is well known that, in a scale-free network, the expected maximum degree (also known as the *natural cutoff*), d_{max}, satisfies the following relation

$$d_{max} = d_{min}\, n^{\frac{1}{\gamma-1}}$$

where d_{min} defines the minimum degree. It therefore follows that the product of D_1 and D_2 in Proposition 2 can be bounded above by an increasing, but concave function of the network size n:

$$D_1 \cdot D_2 \leq d_{min}^2\, n^{\frac{2}{\gamma-1}}$$

Therefore, in terms of the sparseness condition in Poposition 2, we can conclude that

$$\frac{D_1 \cdot D_2}{n} \leq d_{min}^2\, n^{\frac{2}{\gamma-1}-1} \underset{n\to\infty}{\longrightarrow} 0$$

whenever $\gamma > 3$. In other words, if the decay parameter is sufficiently strong, then a scale free network's expected maximal degree grows sufficiently slow to apply Proposition 2. Moreover, γ above 3 is reasonable in many contexts, and in fact describes many real-life communication networks, such as mobile phone calls ($\gamma = 4.69$), email ($\gamma = 3.43$), scientific collaboration

[11] See, for instance, Barabási and Albert (1999)

[12] For evidence on the Power Law properties of online social media, see for instance Goel et al. (2015), Bakshy et al. (2011), and Kwak et al. (2010).

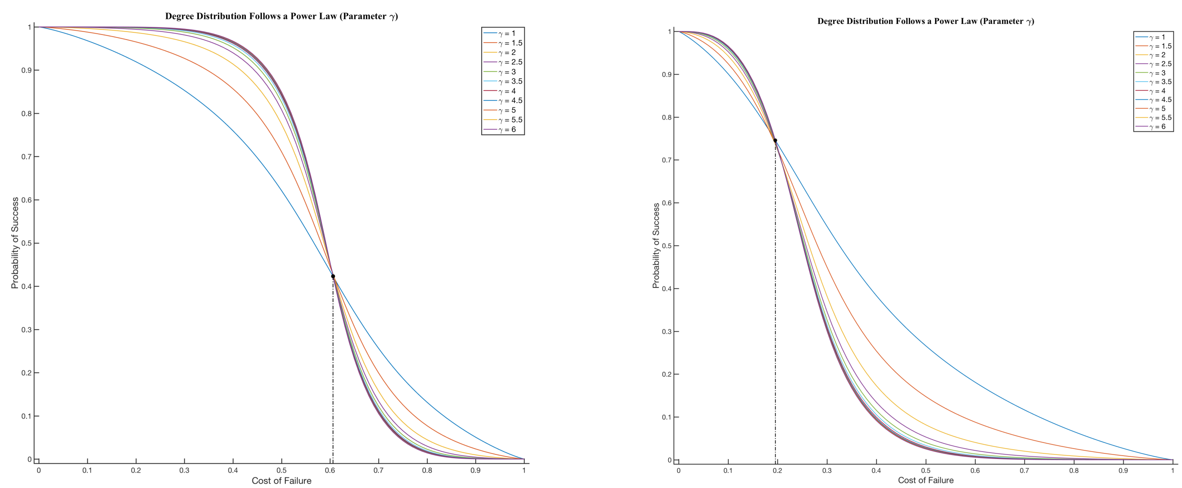

Figure 5: Probability of Success against Cost of Failure for model simulations under a powerlaw distributions and with parameters $D = 200$, $\sigma_0^2 = 4$, $\sigma^2 = 16$. Panel A: $\theta_0 = 0$. Panel B: $\theta_0 = 2$.

($\gamma = 3.35$), and the router internet network ($\gamma = 3.42$).[13]

We will consider shifts in γ as FOSD movements in the degree distributions. To see that this is equivalent, notice that lower values of γ imply that the probably of degrees decays more slowly, so that at least for large values of d, a power-law distribution with $\gamma' > \gamma$ is First Order Stochastically *dominated* with respect to a power-law distribution with parameter γ.

We solve the model numerically by first simulating power law degree distributions for $D = 200$ and a wide array of different $\gamma's$. We then find the $\hat{\theta}$ that solves the equilibrium condition shown in equation (13). Recall that the value of $\hat{\theta}$ specifies the likelihood of successful regime change in equilibrium. From this exercise we therefore obtain the success probability as a function of c, for each different power law distribution (parametrized by γ). We plot the results in Figures 5 and 6 for different values of θ_0, σ_0, and σ. The two panels of Figure 5 capture the effect of changing θ_0, the mean of the prior beliefs about the strength of the status quo. The left panel corresponds to $\theta_0 = 0$ and the right panel corresponds to $\theta_0 = 2$. As found in Proposition 7 for the case of $D = 2$, raising the prior beliefs about θ lowers the value $\hat{c}$ at which the comparative statics are reversed. In any case, notice that there exists one unique value $\hat{c}(\sigma, \sigma_0, \theta_0) \in (0, 1)$ such that, to the left of $\hat{c}$ increasing γ increases the probability of regime change, and to the right of $\hat{c}$ the probability of success decreases with γ. The two panels of Figure 6 capture the effect of an increase in $\frac{\sigma_0^2}{\sigma^2}$. Again, it is clear that there exists one unique value $\hat{c}(\sigma, \sigma_0, \theta_0) \in (0, 1)$ where the ordering of the curves is reversed. However, in this case, raising $\frac{\sigma_0^2}{\sigma^2}$ in fact increases

[13]See, for instance, Barabási (2016)

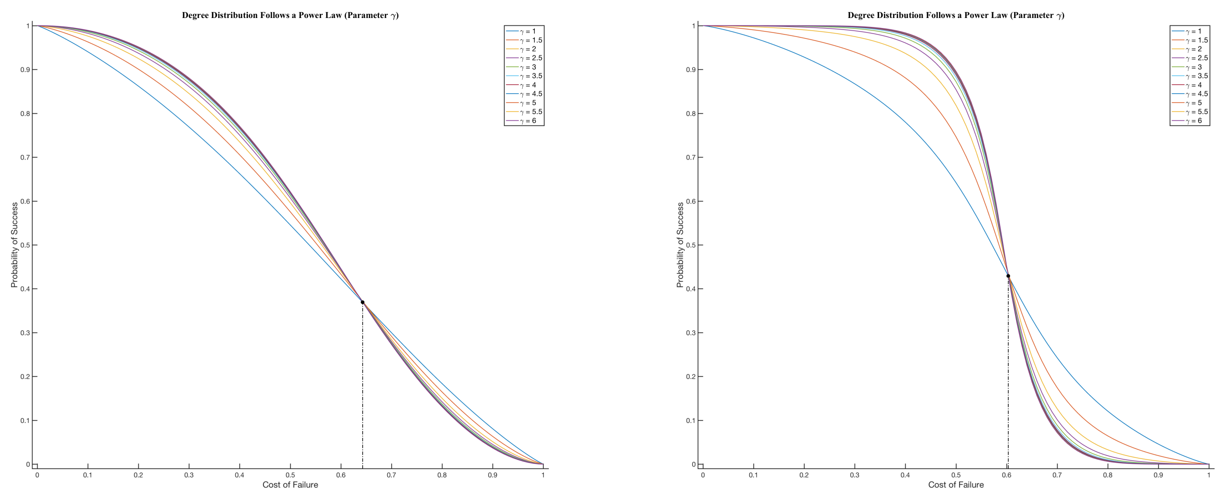

Figure 6: Probability of Success against Cost of Failure for model simulations under a powerlaw distributions and with parameters $D = 200$, $\theta_0 = 0$. Panel A: $\sigma_0^2 = \sigma^2 = 4$. Panel B: $\sigma_0^2 = 4$, $\sigma^2 = 32$.

the value of $\hat{c}$. This was also true in Prop 7 for the case of $D = 2$.

In general, while an analytic solution for power law distributions ($Pr(k) \propto k^{-\gamma}$) is hard to come by, the numerical results presented in Figures 5 and 6 have been carried out for an extensive range of parameter values.[14] We can therefore state the following useful observation.

Simulation Results (Power Law Distribution): *Let $\sigma_0^2 > \sigma$. For the class of degree distributions following a Power Law (i.e. $Pr(d) \propto d^{-\gamma}$ for some $\gamma > 0$), there exists a unique threshold $\hat{c}(\sigma, \sigma_0, \theta_0) \in (0, 1)$ such that for all $0 < c < \hat{c}$ the probability of regime change increases with γ, and for all $1 > c > \hat{c}$ the probability of regime change decreases with γ. Moreover, $\hat{c}(\sigma, \sigma_0, \theta_0)$ decreases with θ_0, increases with $\frac{\sigma_0^2}{\sigma^2}$.*

The results above suggest that, as with the case where $D = 2$, the success probability does not respond monotonically to a FOSD shift in the degree distribution, and instead identifies a threshold cost where the direction of comparative statics is reversed. The intuition corresponds to the arguments presented above and can be seen clearly in Figures 5 and 6.

[14] Only a few examples are shown here (code available upon request). Notice that, although strictly speaking Proposition 2 requires that $\gamma > 3$, we also plot curves for γ below 3. We do this to make the relationship between the curves stand out as much as possible – curves for $\gamma > 3$ are much closer together and it is therefore harder to visualize the existence of a unique threshold that reverses their order. However, the results are qualitatively identical for all values of γ.

8 Discussion

Models of large-scale coordination with incomplete information have usually neglected the role of communication, and in particular the role of connectivity in pooling information. In this paper, we propose a model of large-scale coordination within a network of information transmission. We assume agents observe the private information held by their neighbors within a given social network. This generates a situation of locally public signals that correlate posterior beliefs according to the structure of social interactions. Moreover, a connectivity effect guarantees that the strength of posterior beliefs In this environment, we describe the equilibrium for two-action, two-outcome global games with large networks.

The main technical contribution of this paper provides an upper bound on network density as a function of size, such that, for any communication protocol, the correlation of posterior beliefs is sufficiently mild relative to the connectivity effect. This implies that, for large networks, the connectivity effect dominates and, in the limit, the problem reduces to independent posterior beliefs with strength proportional to connectivity. This allows us to solve for an equilibrium, simply as a function of a network's degree distribution.

After characterizing the equilibrium, we perform comparative statics on the network by shifting the degree distribution. We show that these considerations are not innocent, and that the strategic impact of connectivity on equilibrium outcomes is far from obvious. Indeed, largely connected individuals, while they care little (a priori) for publicly observed information, must strategically respond to the behavior of less connected individuals, and therefore to the public signal indirectly. We show that, when prior beliefs are diffuse (i.e. publicly-held information is completely uninformative of the state of the world), then the probability of successful coordination does not depend at all on the degree distribution. However, the size of successful and unsuccessful attacks does vary with the degree distribution – more informed populations will correspond with smaller, unsuccessful attacks and larger successful attacks.

On the other hand, when public information provides information, shifting the degree distribution affects the likelihood of successful coordination, and we show that the effect is non-monotone and depends crucially on the cost paid when mis-coordination happens. In particular, if the cost of failure is sufficiently small, then the probability of success increases as networks are, on average, less connected. The opposite is true for large failure costs. Intuitively, if payoffs don't fall by much when coordination fails then the probability of success is maximized by having less informed individuals that are less selective about when to attack. Indeed in this scenario failure is also more ubiquitous, but so are successful attacks, increasing the total probability.

Although we frame most of the paper in the context of attacking a regime to overthrow the status quo, the current model can apply in a great deal of additional setting in which agents must choose between two possible choices and there is incomplete information about others' beliefs. One particularly exciting context is the adoption of new technologies. Switching costs and network externalities (i.e. that the value of the product to a consumer depends on the level of consumption by others) implies that individuals will only choose to adopt a new technology if they can be sure that a minimum mass of other individuals will also choose to adopt it. Typical examples are social media platforms, such as Twitter, Facebook, and Whatsapp. In this respect, the model can speak to the type of pre-existing communication network that will maximize the probability of adoption, or can guide firms and developers towards markets or communities that are more likely to switch to the new product. This is just one example, and Myatt et al. (2002) allude to this and a number other settings in which these type of models can be useful – for instance, speculative currency attacks, fire sales of assets, and, of course, popular revolts.

We have considered here equilibrium and comparative statics results in the world where information's noise does not vanish. Part of the literature of global games has focused attention in what happens when noise tends to zero. In that sense, the work of Sakovics and Steiner (2012) is complementary to our approach: in their paper, different groups can receive information with different probability distributions, and they provide a closed form expression of the common threshold for all groups in the limiting case where noise vanishes.

Although we have assumed normally distributed signals, the current model could be extended to other information structures. Indeed, it is true and well-known that, if we restrict ourselves to linear estimators, the minimum variance estimator (and indeed the minimum mean squared error estimator) is given precisely by a weighted average of signals, where the weights correspond to the relative precision of each signal, such as in the expression of the posterior mean of θ in equation (1) (Scharf and Demeure, 1991). However, restricting ourselves to linear estimators is often not the best strategy. In the statistics literature, Diaconis and Ylvisaker (1979) show that, within the class of exponentially distributed signals, only conjugate proper priors satisfy the property that the Bayes estimate of the mean corresponds to a linear function of the signals.

Moreover, even if we could obtain a linear mean as in (1), the equilibrium of this model requires that we know a lot more about the posterior distribution. Notice that our equilibrium is defined by equating the probability of success to the cost of failure, given formally by $Pr(\theta \leq \hat{\theta} \mid x_d^\star) = c$. This requires integrating the posterior distribution of θ, which requires knowing all moments of the distribution. The elegant thing about the normal setting is that, as a conjugate prior, the posterior distribution is known and well-defined, so we can readily obtain our equilibrium condition. This is the main reason why normality is assumed in most papers on global games of regime change in the literature (i.e. Angeletos and Werning (2006), Angeletos et al. (2007) etc.). To summarize, we argue that our model could probably extend to other settings with conjugate symmetric priors, such that the posterior distributions is known and symmetric, and such that the mean and variance are given as in equation (1). This corresponds, for instance, to the gamma prior for the Poisson distribution, or the beta prior for the negative binomial. Notice, however, that this would require justifying a very specific prior and signal relationship.

A point must be made about the communication protocol that is assumed in this paper. Indeed, there are ex-post mechanisms that could be applied in this setting to extract information from others' signals. For instance, Cremer and McLean (1988) show that truthful messages can be induced by certain mechanisms with side payments, as long as player's types are correlated. In the hard overlapping information structure that we assume in this model, these type of mechanisms could be implemented locally by soliciting two separate reports on a single person's signal realization. This would lead to full information revelation at a local level. However, as long as information revelation is only local, one can think that this leads to an augmented network where signals are observed at a greater distance. Since our model applies to any network, this environment can be readily incorporated in this framework.

A more thorough investigation of these type of communication protocols on coordination games is also warranted. Particularly, the solution for finite populations introduces complicated correlation effects. It would be worthwhile to provide more nuanced predictions on the impact of social structure on equilibrium actions. As mentioned above, if posterior beliefs are correlated across nearby players, particularly popular signals will provide information on the state of the world (as always), but at the same time they will also provide information on others' equilibrium actions. Equilibrium considerations therefore will depend on a more detailed description of the network structure than that provided solely by the degree distribution.

References

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006): "Signaling in a global game: Coordination and policy traps," *Journal of Political economy*, 114, 452–484.

——— (2007): "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks," *Econometrica*, 75, 711–756.

Angeletos, G.-M. and I. Werning (2006): "Crises and prices: Information aggregation, multiplicity, and volatility," *The American economic review*, 96, 1720–1736.

Ashworth, S. and E. Bueno de Mesquita (2006): "Monotone comparative statics for models of politics," *American Journal of Political Science*, 50, 214–231.

Athey, S. (2002): "Monotone comparative statics under uncertainty," *The Quarterly Journal of Economics*, 117, 187–223.

Bakshy, E., J. M. Hofman, W. A. Mason, and D. J. Watts (2011): "Everyone's an influencer: quantifying influence on twitter," in *Proceedings of the fourth ACM international conference on Web search and data mining*, ACM, 65–74.

Barabási, A.-L. (2016): *Network science*, Cambridge university press.

Barabási, A.-L. and R. Albert (1999): "Emergence of scaling in random networks," *science*, 286, 509–512.

Barberà, S. and M. O. Jackson (2016): "A Model of Protests, Revolution, and Information," Tech. rep.

Bloch, F. and B. Dutta (2009): "Communication networks with endogenous link strength," *Games and Economic Behavior*, 66, 39–56.

Bueno de Mesquita, E. (2010): "Regime change and revolutionary entrepreneurs," *American Political Science Review*, 446–466.

Calvó-Armengol, A. and J. De Martí (2007): "Communication networks: Knowledge and decisions," *The American Economic Review, p&p*, 97, 86–91.

Calvó-Armengol, A. and J. De Martí (2009): "Information gathering in organizations: Equilibrium, welfare, and optimal network structure," *Journal of the European Economic Association*, 7, 116–161.

Carlsson, H. and E. Van Damme (1993): "Global games and equilibrium selection," *Econometrica: Journal of the Econometric Society*, 989–1018.

Chwe, M. S.-Y. (2000): "Communication and coordination in social networks," *The Review of Economic Studies*, 67, 1–16.

Cremer, J. and R. P. McLean (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica: Journal of the Econometric Society*, 1247–1257.

Dahleh, M. A., A. Tahbaz-Salehi, J. N. Tsitsiklis, and S. I. Zoumpoulis (2015): "Coordination with Local Information," *Operations Research*, 64, 622–637.

DeGroot, M. H. (1974): "Reaching a consensus," *Journal of the American Statistical Association*, 69, 118–121.

Diaconis, P. and D. Ylvisaker (1979): "Conjugate priors for exponential families," *The Annals of statistics*, 269–281.

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

Galeotti, A. and S. Goyal (2010): "The law of the few," *The American Economic Review*, 100, 1468–1492.

Goel, S., A. Anderson, J. Hofman, and D. J. Watts (2015): "The structural virality of online diffusion," *Management Science*, 62, 180–196.

Guimaraes, B. and S. Morris (2007): "Risk and wealth in a model of self-fulfilling currency attacks," *Journal of Monetary Economics*, 54, 2205–2230.

Hagenbach, J. and F. Koessler (2010): "Strategic communication networks," *The Review of Economic Studies*, 77, 1072–1099.

Hassanpour, N. (2010): "Dynamic models of mobilization in political networks," .

Hellwig, C. (2002): "Public information, private information, and the multiplicity of equilibria in coordination games," *Journal of Economic Theory*, 107, 191–222.

Hellwig, C. and L. Veldkamp (2009): "Knowing what others know: Coordination motives in information acquisition," *The Review of Economic Studies*, 76, 223–251.

Iachan, F. S. and P. T. Nenov (2015): "Information quality and crises in regime-change games," *Journal of Economic Theory*, 158, 739–768.

Kwak, H., C. Lee, H. Park, and S. Moon (2010): "What is Twitter, a social network or a news media?" in *Proceedings of the 19th international conference on World wide web*, ACM, 591–600.

Li, B. and K. Zhang (2010): "Strong law of large numbers for*-mixing sequence," *World Academy of Science, Engineering and Technology, International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, 4, 1098–1100.

Morris, S. and H. S. Shin (1999): "A theory of the onset of currency attacks," *The Asian Financial Crisis: Causes, Contagion and Consequences*, 2, 230.

——— (2001): "Global games: theory and applications," .

Myatt, D. P., H. S. Shin, and C. Wallace (2002): "The assessment: games and coordination," *Oxford Review of Economic Policy*, 18, 397–417.

Myatt, D. P. and C. Wallace (2012): "Endogenous information acquisition in coordination games," *The Review of Economic Studies*, 79, 340–374.

Sakovics, J. and J. Steiner (2012): "Who matters in coordination problems?" *The American Economic Review*, 102, 3439–3461.

Scharf, L. L. and C. Demeure (1991): *Statistical signal processing: detection, estimation, and time series analysis*, vol. 63, Addison-Wesley Reading, MA.

Szkup, M. and I. Trevino (2015): "Information acquisition in global games of regime change," *Journal of Economic Theory*, 160, 387–428.

Vives, X. (2005): "Complementarities and games: New developments," *Journal of Economic Literature*, 43, 437–479.

Supplementary Appendix

A Sparseness Condition for Approximating Correlated Networks

A.1 Background

The following discussion proves that we can impose conditions on the network architecture that guarantee correlations across players are sufficiently local and can be disregarded when dealing with large populations. Recall that players optimally choose a threshold strategy following an updated belief about the share of the population that will choose to attack the status quo. In particular, players must form beliefs about the value of A defined as the share of the population that choose $a_i = 1$

$$A = \frac{1}{n} \sum_{j}^{N} 1_{\{\mathbf{x_j} \in B_j\}}$$

With a finite population, A is a binomial random variable and calculating the equilibrium translates into a grueling combinatorics exercise that requires calculating, for each player, the level of correlations with every other individual in order to form high-level posterior beliefs. For networks exceeding 5 players the calculations become highly intractable and provide little insight. Instead, we propose approximating a large network with an infinite population in order to rid the model of correlation effects and focus instead on the relation between connectivity and informativeness. Formally, we provide conditions on the network architecture that guarantees that,

$$A = \frac{1}{n} \sum_{j} 1_{\{\mathbf{x}_j \leq B_j^\star\}} \underset{a.s.}{\longrightarrow} \lim_{n \to \infty} \mathbb{E}\left[\frac{1}{n} \sum_{j} 1_{\{\mathbf{x}_j \leq B_j^\star\}}\right]$$

Then, following arguments in the text, we can conclude that

$$A \longrightarrow \sum_{d=1}^{D} p_d \cdot Pr\left(x_d \leq x_d^\star\right)$$

where p_d is the share of the population with degree d. Not only do we gain in tractability, but we are now able to express the equilibrium as a fixed prediction of success/failure and size/composition of attacks for each value of θ- the alternative would provide for each state of the world θ a probability of success and failure and a distribution of possible sizes and compositions of attacks. In short, the network here introduces two effects: local correlation in private information, and precision stemming from the connectivity of each individual. The following is a methodological contribution for ridding the model of the former effect in order to exploit the impact of the latter effect on the equilibrium.

A.2 Finite-Range Dependence and Strong Mixing Sequences

First we define a class of sequences with bounded correlations called finite-range dependent sequences. We then relate them to another class known as $\star$-mixing sequences, for which a large and well-known class of laws of large numbers (LLN) exist. The reason we proceed this way is that in the following section we develop an algorithm for naming nodes on a network that guarantees, under some intuitive conditions on the growth rates of the largest degrees in the network, that the sequence of Bernoulli random variables in the definition of A in equation (6) indeed satisfy finite-range dependence. Once we show that finite-range dependence implies $\star$-mixing, we can refer to the classic convergence results to establish the desired LLN.

Let (Ω, F, P) be a probability space and let $\{X_n, n = 1, 2, \ldots\}$ be a sequence of real-valued random variables defined on (Ω, F, P). For each positive integer n, let F_n be the smallest σ-algebra in which X_n is measurable, and for $n \leq m$ let F_n^m be the smallest σ-algebra in which $X_n, \ldots, X_m$ is jointly measurable.

Definition 8 (Finite-Range Dependence)**.** *A sequence $\{x_i\}_{i=1}^{\infty}$ of random variables defined on (Ω, F, P) exhibits finite-range dependence if and only if there exists an I such that if $\mid i - i' \mid \geq I$ then x_i and $x_{i'}$ are independent:*

$$P(A \cap B) = P(A)P(B)$$

for all $A \in F_i$ and $B \in F_{i'}$.

Finite range dependence implies that if we take two elements far enough apart in the sequence, we are guaranteed that these elements must be independent. Similarly, $\star$-mixing sequences place bounds on the strength of the correlation as a function of the distance between elements in the sequence, and imposes that as the distance grows the strength of the correlation vanishes.

Definition 9 ($\star$-Mixing Sequences from Li and Zhang (2010))**.** *Let $\{X_n, n \geq 1\}$ be a sequence of random variables. X_n is called a $\star$-**mixing sequence** if there exists a positive integer I, and* *a function f such that $f \downarrow 0$ and for all $n \geq I$, $m \geq 1$, $A \in F_1^m$, and $B \in F_{n+m}^{\infty}$,*

$$| \, P(A \cap B) - P(A)P(B) \, | \leq f(n)P(A)P(B).$$

Next we need to show that if a random sequence satisfies finite-range dependence, then it must necessarily satisfy the $\star$-mixing (or uniformly strong mixing) property.

Lemma 1. *Any sequence that exhibits Finite Range Dependence is a $\star$-Mixing Sequence.*

Proof. Our definition of finite range dependence in Definition 8 can be expressed in terms of σ-algebras as saying that for all $j \in \mathbb{Z}$, there exists an I such that for every $s > I$ we have $\hat{\alpha}\left(s, j\right) = 0$, where,

$$\hat{\alpha}\left(s, j\right) = \sup\left\{\left|P\left(A \cap B\right) - P\left(A\right)P\left(B\right)\right| : A \in F_j, \;\; B \in F_{j+s}, \right\}$$

But notice that, by finite range dependence, if x_j is independent of x_{j+s} then it is also independent of x_k for all $k > j+s$. Similarly all x_m with $m < j$ are also independent of x_{j+s} (and consequently also independent of all x_k with $k > j+s$). As a result we can establish that,

$$\hat{\alpha}\left(s, j\right) = \sup\left\{\left|P\left(A \cap B\right) - P\left(A\right)P\left(B\right)\right| : A \in F_1^j, \;\; B \in F_{j+s}^{\infty}, \right\}$$

Now, since $\hat{\alpha}\left(s, j\right) = 0$ for all $s > I$, definition 9 is satisfied, for instance, for $f(n) = 1/n$ since we have that

$$| \, P(A \cap B) - P(A)P(B) \, |= 0 \leq 1/n \, P(A)P(B).$$

for all $s \geq I$, $j \geq 1$, $A \in F_1^j$, and $B \in F_{j+s}^{\infty}$ ■

Finally, by Theorem 2.1 of Li and Zhang (2010) that establishes SLLN for uniformly strong mixing sequences, we can establish the following result:

Lemma 2. *If a sequence $\{x_i\}_{i=1}^{\infty}$ of non-negative random variables defined on (Ω, F, P) exhibits finite-range dependence, such that $\mathbb{E}x_i = \mu_i < \infty$ for all i, and $\sum_{i=1}^{\infty} \mathbb{E}X_i^2/i^2 < \infty$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_i) = 0 \quad a.s.$$

so a Strong Law of Large Numbers applies.

A.3 A Naming Algorithm

Once the above Lemma is shown to hold, we can construct an algorithm that assigns indexes to the players in the network such that the resulting sequence exhibits finite-range dependence. By the above lemma, the LLN applies to sub-sequences corresponding to players of the same degree, and thus Proposition 2 holds. The following algorithm is essentially a *Breadth-First-Search* algorithm on a graph.

Algorithm. Start from any node in the network, call it 1. Assign consecutive indexes to each of $1's$ neighbors. Next, starting from the neighbor with the lowest index, assign consecutive indexes to the neighbors of $1's$ neighbors. If a node is already named, do not rename it. Continue in this way.

Recall that D represents the maximal degree and that $\mathcal{D}$ represents the set containing the degree of each player in the network. Now define $D_1 = max\left(\mathcal{D}\right)$ and $D_2 = max\left(\mathcal{D}_{-i}\right)$ for some $i \in N$ with $d_i = D_1$. Given the naming algorithm, we can find a value of $I < n$ such that the sequence of all players has finite-range dependence. Specifically, we have that when

$$I = D_1\left(1 + D_2\right)$$

there is finite-range dependence for the entire sequence of nodes. To see this notice that any two players i and j with $|i - j| > D_1\left(1 + D_2\right)$ must necessarily lie more than two links away from each other. Given our information aggregation procedure, this guarantees that they are not correlated. Of course this is not true for the complete network, but in that case $I = (n-1)\,n$ which is greater than n for $n > 2$. Clearly the network must be sufficiently sparse such that $I < n$[15]. Although the algorithm gives some freedom as to the precise labeling of the nodes it guarantees that any two nodes with labels I units away will necessarily lie more than two links away from each other. Figure **??** illustrates the algorithm and shows how the value $I = D_1\left(1 + D_2\right)$ guarantees two degrees of separation for a tree network of 21 players. The reason for using the tree is that, because no neighbor of 1's neighbors is also 1's neighbor, it constitutes the starkest example imaginable.

The argument above has assumed that I corresponds to a fixed integer. It is easy to see that fixing the maximum degree while increasing the total population increases network sparseness.

[15] In fact, the next section imposes additional conditions on the behavior of I as a function of the total population n. That is, $I < n$ is a necessary, but not sufficient condition for convergence.

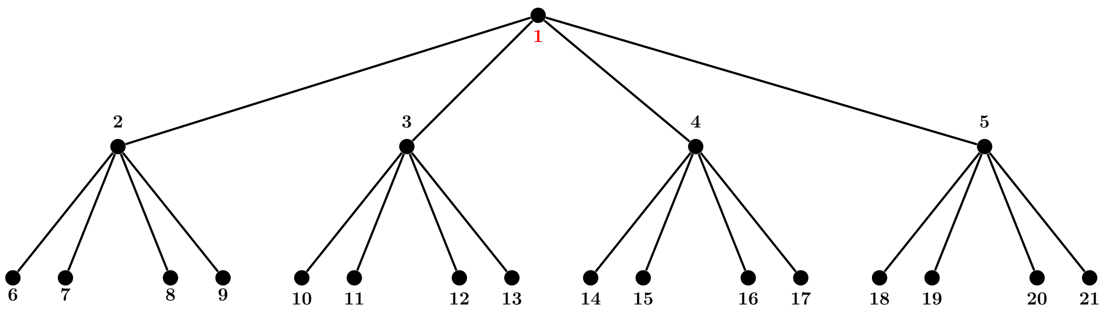

Figure 7: The Breadth-First-Search Algorithm for a Tree network with $n = 21$ and $D_1 = 4$, $D_2 = 4$ (i.e. $I = 20$). Notice that for all i and j with $\mid i - j \mid > 20$ will necessarily lie more than two links away.

It turns out, however, that weaker conditions exist. Specifically we can establish the same result for values of I that grow with n, provided the growth rate is sufficiently slow. The following section formalizes this result.

A.4 Conditions on the Growth Rate of Degrees

In this section we specify sufficient conditions on the behavior of the largest degrees in the network, $D_1 \geq D_2 \geq \ldots$ in order to guarantee that we can construct sequences of nodes with finite-range dependence, and hence the law of large numbers applies. As mentioned above, we need LLN to hold so that we can approximate the network with an infinite population and gain tractability. The algorithm above guarantees that we can construct $I = D_1 (1 + D_2)$ sequences of independent random variables, each with $\frac{n}{I}$ elements. Of course, we need that as n tends to infinity, the value of I does not grow too fast so that we can be sure that $\frac{n}{I}$ also tends to infinity. Otherwise the sequences would only contain a finite number of terms (which is not possible). We could just impose that $D_1, D_2, \ldots$ are fixed to some constant, but we are interested in finding weaker conditions. So we need that

$$\frac{n}{I(n)} \underset{n \to \infty}{\longrightarrow} \infty$$

this implies that $I'(n) \longrightarrow 0$.

In general for $I = D_1 + D_1 D_2$ we have that $I'(n) = D_1'(n) + D_1'(n) D_2(n) + D_1(n) D_2'(n)$. We know that in general for any $i \in E$, $D_i'(n) \geq 0$ (otherwise nothing to prove) and so the condition that $I'(n) \to 0$ implies that all three terms go to zero. The first condition simply implies concavity- i.e. $D_1''(n) < 0$. This makes sense, it says that as n grows, the maximum degree should grow at a slower rate. The next two conditions however impose additional conditions on the concavity of these functions. So far we only have results for the case where degrees follow a power function of the entire population.

A General Result

In general we can think of a number of communication protocols that generate all sorts of local correlations. We have assumed in this model that correlations are present up to 2 links of separation. But there is no reason why this should be the case. In general, if correlations emerged at k degrees of separation, then we would need to redefine our I. In this case

$$I = D_1 + D_1 D_2 + D_1 D_2 D_3 + \cdots + D_1 D_2 D_3 \ldots D_k = \sum_{j=1}^{k} \prod_{l=1}^{j} D_l \tag{14}$$

It is clear that as the aggregation procedure generates correlations that stretch farther across the network, then the restrictions on the growth rate of the degrees becomes stronger in the sense that it imposes structure on the shape of smaller degrees. Finally, we provide a general characterization that guarantees $\frac{I(n)}{n} \to 0$. Notice that for any general k the convergence rate of $I(n)$ is determined by the last term in the sum in equation (14), so that the necessary condition becomes $\frac{\prod_{j=1}^{k} D_j}{n} \to 0$ or, what is the same, that

$$\prod_{j=1}^{k} D_j \in o(n)$$

which is a general form of the expression in Proposition 2.