*Article*

RATIONALITY AND SOCIETY

Rationality and Society
25(1) 3–40

DOI: 10.1177/1043463112473734
rss.sagepub.com
SAGE

# Homophily, networks, and critical mass: Solving the start-up problem in large group collective action

**Damon M. Centola**
M.I.T. Sloan School, USA

## Abstract

Formal theories of collective action face the problem that in large groups a single actor makes such a small impact on the collective good that cooperation is irrational. Critical mass theorists argue that this 'large group problem' can be solved by an initial critical mass of contributors, whose efforts can produce a 'bandwagon' effect, making cooperation rational for the remaining members of the population. However, critical mass theory requires an explanation of how a critical mass can form in the first place. I present a model of collective action that solves this problem by showing how aspects of social structure – including network topology, homophily, and local coalition formation – can allow rational actors to endogenously form a critical mass. The findings indicate that as the mobilization effort becomes more 'complex', clustered networks and homophily become increasingly important for critical mass collective action.

## Keywords

Collective action, critical mass, large group problem, social networks, homophily

**Corresponding author:**
Damon M. Centola, M.I.T. Sloan School, 100 Main Street, E62-462, Cambridge, MA 02142, USA.
Email: dcentola@mit.edu

The large group problem in collective action has a very simple logic: the greater the number of people that is required to produce a collective good, the less the value of any single individual's contribution. Mancur Olson (1965) famously argued that this logic implies that, barring additional incentives, no individual will be rational to contribute, and voluntary collective action in large groups is doomed to failure. Despite Olson's 'large group argument', large group collective action occurs all the time: political protests, social and environmental movements, and electoral campaigning are all examples of 'mass action', in which large numbers of people mobilize for a collective good (Finkel et al., 1989; Heckathorn, 1993, 1996; Kim and Bearman, 1997; Marwell and Oliver, 1993). Which poses the puzzle: if large group collective action is formally impossible, how come we empirically observe so many instances of mass action?

The most significant response to Olson's large group argument comes from Marwell and Oliver's (1993) study of critical mass dynamics (see also Marwell et al., 1988; Oliver et al.,1985; Oliver and Marwell, 1988; Prahl et al., 1991). While Olson argues that the only way to initiate sizable collective action is to add 'selective incentives', such as punishments for defection or rewards for cooperation, Marwell and Oliver show that contributions to collective action can create 'positive externalities', whereby initial contributions create more incentives for subsequent actors. This generates a 'bandwagon' effect that can mobilize large populations to support a common cause. According to Marwell and Oliver, the solution to Olson's puzzle of large groups is not that people are irrational, but that Olson fundamentally misunderstood the self-reinforcing logic of cooperation. In the last decade, this critical mass argument has played a major role in the development of collective action theory (Heckathorn, 1993, 1996; Kim and Bearman, 1997; Macy, 1990; Marwell and Oliver, 1993; Oliver, 1993; Rasler, 1996; Yin, 1998), and provided an essential foundation for incorporating rational choice theory into the mainstay of sociological approaches to collective action (Oliver and Marwell, 2001).

The issue for critical mass theory is explaining where the critical mass itself comes from.[1] Until the bandwagon gets started, how does anyone have an incentive to participate at all? Marwell and Oliver (1993) propose that significant heterogeneity in the distribution of resources will create a group of highly motivated, high-resource individuals whose contributions are sufficient to generate large positive externalities for others (Kim and Bearman, 1997; Marwell and Oliver, 1993). While this solution is certainly applicable to some forms of collective action, such as philanthropic donations to build a hospital, a library, or a church (Marwell et al., 1988; Oliver and Marwell, 1988), it begs the question of how these individuals coordinate their efforts to organize their critical mass contribution. Even high-resource individuals must solve the 'coordination dilemma' of figuring out who is interested, and how best to collaborate to initiate collective action (Marwell and Oliver, 1993).[2] This problem is exacerbated in large populations with less extreme resource distributions, in which actors who want to form a critical mass will have fewer resources, and must self-organize to get their campaign off the ground.

There are many examples of 'mass action', such as large strikes (Klandermans, 1988), political protests (Opp and Gern, 1993), social revolution (McAdam, 1986, 1988), and violent revolution (Wood, 2003), in which people with moderate resources coordinate to initiate a bandwagon of growing participation. In many examples of grassroots mobilization, organizers may have non-monetary resources to offer, such as time and enthusiasm (Marwell and Oliver, 1993).[3] However, the primary measure of success for these kinds of collective action is typically not the amount of support given by any individual, but the number of participants that can be mobilized (Granovetter, 1978; Schelling, 1978). This is typically true for strikes, protests, and revolutions, where a high-resource individual simply cannot do the work for everyone (see, e.g., Heckathorn, 1993, 1996; Kim and Bearman, 1997; Marwell and Oliver, 1993; Polletta, 1998). Large numbers of people need to mobilize in order for the action to succeed. The question that needs to be answered is how these people can coordinate to get their action off the ground.

Previous attempts to solve this problem have focused on the role of structural properties, such as social networks (Chwe, 1999; Gould, 1993; Kim and Bearman, 1997; Macy, 1990) and homophily (Chiang, 2007; Chwe, 1999), in the emergence of bandwagon dynamics.[4] These studies have found that weak ties in the social network (à la Granovetter's (1973) 'strength of weak ties' argument), and moderate levels of homophily, can be very effective in helping to promote the spread of social cooperation.[5] Significantly, studies also found that too much homophily can actually impede the success of collective action by limiting the diversity of people's social networks (Boyd and Richerson, 2002; Chiang, 2007). These approaches to collective action demonstrate how structural factors can explain the emergence of bandwagon dynamics; however they face the problem that they eliminate population size from the calculus of cooperation, thereby failing to engage the basic question of how cooperation emerges in large groups.

I present a model of critical mass formation that builds on Marwell and Oliver's (1993) theory of critical mass and Heckathorn's (1996) theory of social dilemmas, to show how the formation of local coalitions can allow rational actors to initiate large group collective action. I show that if actors are allowed to interact locally, creating coalitions of cooperators, they can endogenously mobilize the critical mass necessary to initiate population-wide cooperation. Contrary to earlier studies (Centola et al., 2005; Chiang, 2007), I find that the effect of homophily on initiating collective action is not curvilinear; rather, I find that more homophily is better. I also find that increasing the number of weak ties can prevent critical mass formation, which suggests that clustered networks (with no weak ties) can sometimes be better for promoting bandwagon dynamics. These results show how homophily and social networks can dramatically, and unexpectedly, help to explain the emergence of collective action in large groups.

## *N*-person games: From prisoner's dilemma to critical mass

The classic approach to modeling collective action begins with a population of *N* actors who are embedded in an '*N*-person game', in which an individual's payoffs for cooperation and defection are contingent upon the decisions of the rest of the population taken as a whole. The traditional formalization of the *N*-person prisoner's dilemma (Bonacich et al., 1976; Hamburger, 1973) states that regardless of how many people are cooperating, from 0 to *N*, it is always rational to defect. Or, more technically, defection 'dominates' cooperation at every level of public goods production.

Despite this dominance of defection over cooperation at each point along the production curve,[6] the formal definition of the *N*-person prisoner's dilemma also states that the payoff for cooperating when there is universal cooperation is always better than the payoff for defecting when there is universal defection. This is a familiar rendition of the social paradox embodied by collective action (Hardin, 1982; Olson, 1965), in which although everyone is better off with universal cooperation than with universal defection, cooperation is forestalled by the fact that at every point along the way it is individually beneficial to defect. Consequently, everyone defects, and no collective goods are produced.

In the critical mass model developed by Marwell and Oliver (1993), a small group of interested individuals can provide a substantial amount of the collective good, creating incentives for others to also participate. A cascade of collective action can thus be triggered by the initial contributions of a few motivated individuals. However, it is clear that in the *N*-person prisoner's dilemma there cannot be a critical mass. No matter how large the group of initial contributors is, defection dominates cooperation. To generate critical mass dynamics, we must model collective action not as an *N*-person prisoner's dilemma, but as an *N*-person assurance game (Chong, 1991; Runge, 1984). Unlike the prisoner's dilemma, in which actors have competing incentives, in an assurance game the basic problem is one of coordination. The formal definition of an *N*-person assurance game states that each person has a threshold for cooperating,[7] which is defined as the number of other people that need to cooperate before a person will be willing to also contribute. Below each person's threshold level, they will not participate, but above it they will gladly help out. Formally, this translates into the dominance of defection over cooperation below the threshold, and the reverse above it.

Threshold models like those studied by Schelling (1978) and Granovetter (1978) show how populations with distributed threshold values are susceptible to mobilization by a critical mass of initiators, activists, or innovators whose actions trigger a chain reaction of participation. Each contributor increases the level of cooperation, which further increases the likelihood that the remaining actors will also have their thresholds triggered.

In an *N*-person assurance game, beginning from universal defection, no actor has an incentive to be the initial contributor, and thus the population is faced with the 'start-up' problem of critical mass. I follow Granovetter (1978), Schelling (1978), Marwell and Oliver (1993), Rogers (1995) and others (Dodds and Watts, 2004, 2005) in defining the 'critical mass' as the gap between zero cooperation and the level of cooperation at which the growth of participation becomes self-sustaining. By definition, below the critical mass level, individuals will not cooperate.

The challenge of bridging this gap is the key problem for large group collective action. It is the situation that organizers find themselves in when they want to organize a strike, but they know that no one will join until they see enough others join to suggest that the strike will proceed (Klandermans, 1988). Activists are faced with a similar problem when they want to organize a political demonstration, but know that oppressive state forces will keep citizens away until they see others going out to demonstrate (Finkel et al., 1989; Opp and Gern, 1993). The classic start-up problem for critical mass theory is how to get anyone to be the first to participate.

Opp and Gern's (1993) study of the 1989 Berlin protests sheds significant light onto how populations self-organize to solve this problem. Before the protests began, "[c]itizens were faced with a dilemma: They had strong desires to engage in action against the government, but the costs of doing so were high… A citizen considering participation in such spontaneous gatherings is faced with a *coordination problem*" (Opp and Gern, 1993: 662). Opp and Gern argue that the solution to the start-up problem came from social groups and personal networks. "Networks of friends, colleagues, or neighbors constitute micro-contexts for mobilizing citizens… even in authoritarian regimes, politically homogeneous networks whose members trust each other and communicate in a relatively uninhibited way may be established" (Opp and Gern 1993: 662).

This solution to the start-up problem has not escaped the attention of formal theorists. Kim and Bearman (1997) attempt to provide a formal solution to the problem of critical mass in large groups by using actors embedded in social networks. However, even by their own standards Kim and Bearman's model "falls short of a large group solution" (1997: 71). This is because the actors in their model assume that the public good will only be divided among their immediate neighbors in the network. They do not evaluate the costs and benefits of a public good shared over a large population (see Kim and Bearman, 1997: 77−79). Similarly, Chwe (1999) studies the effects of social network structure on the initiation of collective action. However, he also uses a model in which actors do not evaluate their payoffs using a 'global' production function, but instead use their local networks. Consequently, the size of the population is irrelevant to the actors' decision-making. Other formal models of social network mobilization in collective action (Centola and Macy, 2007; Chiang, 2007; Kitts, 2000, 2006, 2008; Kitts et al., 1999; Macy, 1991), while making interesting contributions in their own right, have similarly failed to address the large group problem head on. None of these studies shows how a voluntary critical mass can be initiated when the collective good is divided over a large population (Chwe, 1999; Kim and Bearman, 1997).

My approach to collective action in large groups combines Marwell and Oliver's (1993) theory of critical mass with Heckathorn's (1996) model of strategic interactions in local coalitions to address this problem. Heckathorn's (1996) study shows how actors' strategic incentives change over the course of collective action. I develop a simple extension of this model, which shows that changing the size of actors' local coalitions can alter their strategic incentives – turning a small number of defectors into a critical mass of cooperators. My model provides a general solution to the problem of critical mass in large groups by showing how to solve for the minimum coalition size required to initiate social cooperation. I then use this model to investigate the effects of homophily and network structure on the dynamics of critical mass formation. I find that increasing homophily helps critical mass dynamics, while increasing weak ties can sometimes impede these dynamics. These results show that the effects of homophily and network structure on the success of large group collective action can depend decisively on the size of the critical mass coalition.

## A coalition-based model of collective action

Following Heckathorn (1996), in the present model the decision whether to contribute is not modeled as a solitary action in which the actor must evaluate her rewards in a 'one-versus-all', *N*-person framework (Bonacich et al., 1976; Granovetter, 1978; Hamburger, 1973; Marwell and Oliver, 1993; Schelling, 1978).[8] Instead, actors form local coalitions of $n$ individuals who collectively try to initiate social change. My approach relaxes Heckathorn's (1996) assumption that all coalitions are of size $n = 2$, which allows me to investigate the effects of coalition size on the dynamics of initiating critical mass. An actor's decision whether to cooperate or defect is determined by the incentives for mutual cooperation, mutual defection, unilateral defection, and unilateral cooperation among the members of the coalition. In a local, or $n$-person, coalition, agents evaluate the outcomes of each of these alternatives in terms of the increase in the public good, $L$, that each strategy produces.

The level of the public good, $L$, is determined by a production function that aggregates individual contributions into collective goods (Hamburger, 1973). The production function commonly used both in economic analysis and in studies of collective action is an S-shaped curve.[9] Following Macy (1990) and others (Heckathorn, 1996; Kim and Bearman, 1997), I model this production curve with the cumulative logistic function given in equation (1).

$$L(\pi) = \frac{1}{1 + e^{(.5 - \frac{\pi+\delta}{N^{1-J}})\beta}} \tag{1}$$

In equation (1), $N$ is the total population size, $0 \leq \pi \leq n$ is the number of people in the coalition that are cooperating, $0 \leq \delta \leq N-n$ is the level of participation by members outside the coalition, and $0 \leq J \leq 1$ is the jointness of supply − the degree to which increases in the size of the population decrease each person's enjoyment of the good (discussed in more detail below).[10] The quantity $\frac{\pi+\delta}{N^{1-J}}$ represents the relative impact of the coalition's cooperation on the provision of public goods. For studying the initiation of collective action, I assume that the initial level of cooperation by members of the population is $\delta = 0$. Figure 1 shows a graphical representation of $L$ for levels of participation ranging from zero cooperation to full cooperation. The key feature of this function for critical mass dynamics is that increasing

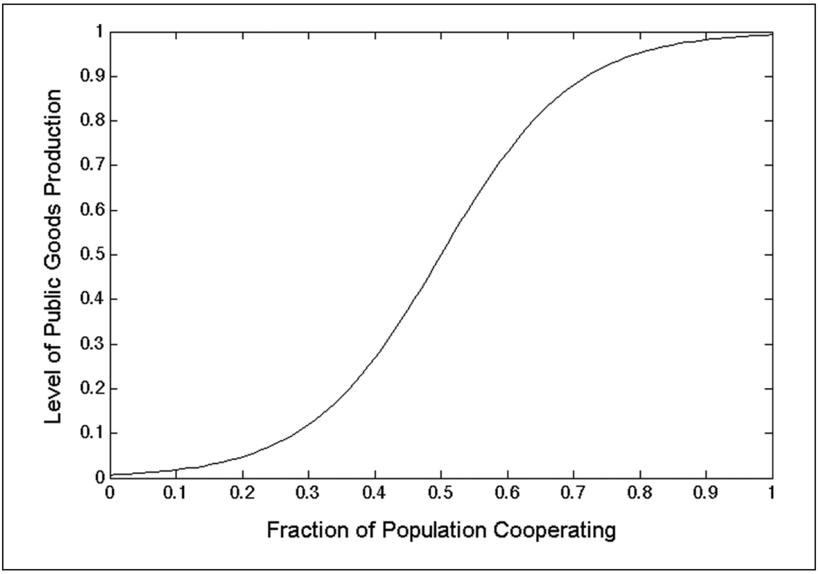


**Figure 1.** S-shaped production function.

levels of participation in the early stages increase the rewards for further participation.[11]

Equation (2) shows the payoff, $U_i$, that a player $i$ receives during one iteration of the game, where $V_i$ is the value that an individual actor assigns to the collective good (randomly distributed across the population), and $C_i = 1$ if actor $i$ cooperates, otherwise $C_i = 0$.

$$U_i = V_i L - C_i K \tag{2}$$

When $C_i = 1$, an individual's rewards for contributing, $V_i L$, are reduced by the cost of cooperation $K = 1$, which is the same for every player.[12] Heterogeneity in $V$ indicates that some members of the population place a higher value on the collective good, and thus are more likely to join the collective action early on. Actors with lower $V_i$ find participating more costly and thus will wait until others have generated more incentives before they participate.[13] Critical mass models of collective action (Granovetter, 1978; Heckathorn, 1990, 1993; Kim and Bearman, 1997; Macy, 1990, 1991; Marwell and Oliver, 1993; Young, 2009) have traditionally used the normal distribution of interests as the foundation for their analyses. Following this tradition, I will also focus my analysis on normal distributions of "interests in participation," paying particular attention to the robustness of the model's dynamics for different values of both the mean ($\bar{V}$) and standard deviation ($\sigma$) of the distribution. All actors are assumed to have homogeneous resources, equal to one unit of contribution per person.

An individual's action is based on the payoffs for cooperation or defection in a strategic game with the other members of the actor's coalition. To set up the strategic interaction between an actor and the rest of her coalition members, I represent the fellow coalition members as a single actor. This allows the decision problem within a coalition to be represented as a simple 2×2 game. One half of the 2×2 game is an actor who is randomly selected from the population and asked to decide whether to cooperate or defect. The actor bases her calculation on her own interest in the public good plus the willingness of her fellow coalition members to cooperate. The other half of the 2×2 game is a collective 'player' representing the aggregated interests of the other coalition members.

If every member of an actor's coalition is willing to cooperate, the 'player' that collectively represents them is also willing to cooperate. However, if any of the members of a players' coalition would prefer defection, then their collective representation is defection.[14] This accurately reflects the fact that mutual cooperation within a coalition only occurs when everyone in the coalition agrees to cooperate. Thus, I make the conservative assumption that it is only under unanimous cooperation that a coalition can create a critical mass and mobilize collective action.

The game among the coalition members has four possible outcomes. First, both the active actor and the player representing the other coalition members can cooperate. This is denoted as 'R', in deference to the prisoner's dilemma, since it corresponds to the 'reward' payoff where both players win. Second, the active player can cooperate, but the other player can defect – which would indicate that at least one of the other coalition members was unwilling to contribute to the collective action. This is called 'S', again with reference to the prisoner's dilemma, since it is the 'sucker' payoff, where you contribute but no one else does. Third, the active player can defect while the other player cooperates. This is denoted as 'T' since it is the 'temptation' in the prisoner's dilemma, where the active player tries to get the benefits of the other actor's cooperation without herself contributing. Finally, both players can defect. This is called 'P', referring to the mutual punishment payoff in the prisoner's dilemma, and results in no one contributing anything. This outcome corresponds to Olson's expectations for free-riding in collective action.

**Table 1. Core game's payoff matrix.** The payoffs show a 2×2 game with row versus column. The term in each cell is the row's payoff

| | C (Rest of coalition cooperates) | D (Rest of coalition defects) |
|---|---|---|
| C (*i* Cooperates) | $R_i = V_iL(n) - K$ (3)<br>($R$ = the 'reward' payoff) | $S_i = V_iL(1) - K$ (4)<br>($S$ = the 'sucker' payoff) |
| D (*i* Defects) | $T_i = V_iL(n-1)$ (5)<br>($T$ = the 'temptation' payoff) | $P = 0$ (6)<br>($P$ = the 'punishment' payoff) |

Table 1 shows the payoffs that an actor *i* receives for all four outcomes ($R_i$, $S_i$, $T_i$, and $P$, equations (3)−(6)), which are represented in terms of the interaction between an individual actor ('row') and the other members of the coalition treated as a collective actor ('column'). In the upper left cell, where both the row and column cooperate, the coalition produces the maximum increase in the collective good. All players receive the full value of local goods production, but each player's payoff is also reduced by the cost of cooperation, *K*. In the upper right cell in Table 1, the active actor (row) contributes while the remainder of the coalition (column) defects, so the row gets the 'sucker' payoff, *S*. This reflects the lower level of collective goods production resulting from only a single actor cooperating, and the row's absorption of the contribution cost, *K*. In the lower left cell, the active actor (row) defects while the other coalition members (column) contribute. The temptation payoff, *T*, also reflects a lower level of collective goods production than mutual cooperation, but the actor does not incur any contribution costs. Finally, in the lower right cell, everyone defects, so no change is made to the level of the collective good and no marginal utility is gained from the interaction. Since no contribution costs are incurred, the net payoff from the 'punishment' outcome, *P*, is zero.

Based on Heckathorn (1996), the payoffs for cooperation and defection within a coalition correspond to unique values of *T*, *R*, *P*, and *S*. Each ordering of these values, in turn, corresponds to a distinct strategic situation. This is essential for understanding collective action since it means that neither the prisoner's dilemma nor the assurance game have sole province over the strategic problem. There are five qualitatively different orderings of the payoffs *T*, *R*, *P*, and *S*, each of which corresponds to a different strategic situation, or game type, that can be encountered by actors in a collective action problem. These games are: the prisoner's dilemma, the altruist's dilemma, the privileged game, the assurance game, and the chicken game.

The first of the five games, the prisoner's dilemma ($T > R > P > S$), corresponds to a tension between individual and collective interests, resulting in a dominant strategy of defection. This is the classic framing of the problem of collective action. The altruist's dilemma ($P > T > S > R$, which falls outside the present discussion) corresponds to both an individual and collective interest in defection – which makes defection both dominant and unproblematic. The flip side of this is the third game, the privileged game ($R > T > S > P$), in which individual and collective interests coincide on cooperation, making cooperation the dominant strategy. In the fourth game, the assurance game ($R > T > P > S$), interests are conditional – actors will cooperate only if every other member of the coalition also cooperates. Finally, in the chicken game ($T > R > S > P$), interests are also conditional, but the interest structure is the opposite of the assurance game. Actors cooperate in the chicken game only when the rest of the coalition defects (see Heckathorn, 1996, for a more detailed description of these games and their empirical applications).

Figure 2 represents these four strategic situations (the altruist's dilemma is omitted) in terms of a dynamic strategy space for individuals engaged in mobilizing collective action. Each individual has a unique location in the strategy space based on her values of $T$, $S$, $R$, and $P$. An individual's location is determined by three variables. The first is her utility from contributing to the collective action ($U_i$), which is represented by the $z$-axis. Lower values of $U_i$ correspond to the prisoner's dilemma. Intuitively, this means that actors who get less utility from the public good, will be less inclined to cooperate. Similarly, very high values of $U_i$ correspond to the privileged game. This means that people who derive great utility from the good are likely to cooperate regardless of what others do.

As stated in equation (2), the value of $U_i$ is determined both by $V_i$, which is endogenous to each actor, and by the impact that each coalition makes on the collective good, which changes based on the current level of participation in the population. Figure 1 shows that as participation increases, the rate at which new contributions matter for the public good (i.e., the steepness of the production curve) also increases. This is the cumulative effect of a bandwagon dynamic – as more people cooperate it makes more people want to cooperate. Contributions have a greater impact on the public good further along the curve than at the very beginning (Finkel et al., 1989; Heckthorn, 1996).[15] Consequently, actors' vertical positions in the space can change as a function of other actors' behaviors. A few actors' decisions to cooperate can alter the strategic situations of hundreds of others, bringing them into the collective action.

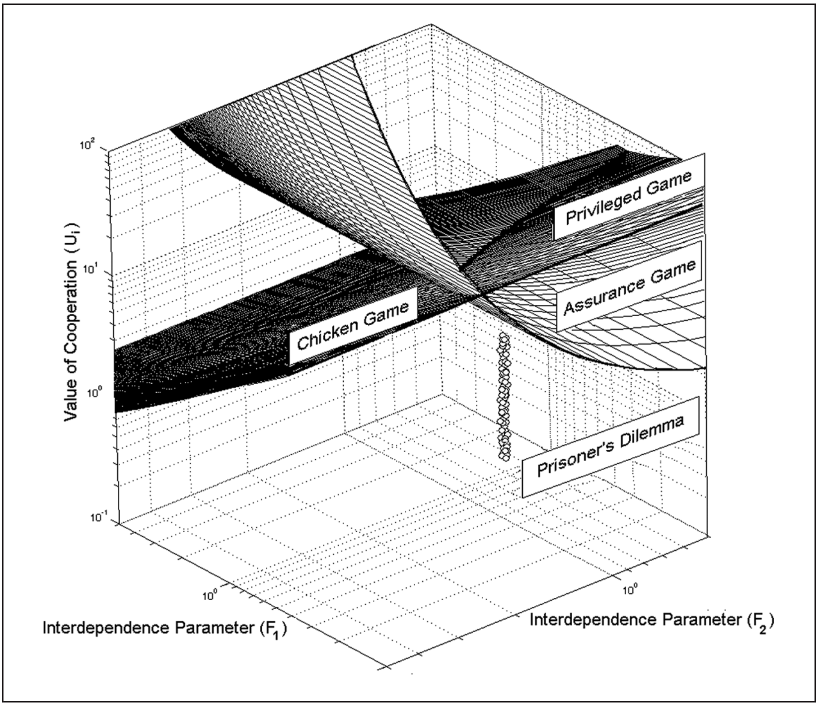


**Figure 2.** Prisoner's dilemma incentives in the *N*-person assurance game with no coalitions ($N$ = 1000, $J$ = 0.3, $\bar{V}=600$, $\sigma=200$). The strategic interests of actors playing an *N*-person assurance game without coalitions.

The second and third variables that determine an individual's position in the strategy space (the $x$ and $y$ axes) correspond to the strategic interdependence of the actors *within* a coalition. The $x$ axis, $F_1$, is the influence that the active actor's decision has over the decisions made by the rest of the coalition members, while the $y$ axis represents $F_2$, the influence of the rest of the coalition on the active actor's decision. These reciprocal effects of coalition members on each other's interests in cooperating evolve over the course of a collective action, altering the relative values of $F_1$ and $F_2$, and moving actors horizontally within the strategy space. Formal definitions of $F_1$ and $F_2$, along with complete analytical details of the model, are presented in Appendix 1.

The strategy space shown in Figure 2 (plus the altruist's dilemma, which is located *below* the prisoner's dilemma) exhaustively captures the possible

**Table 2.** First and second mover payoffs in the sequential game-playing model

| Strategy choice by move | Second mover | | | | |
|---|---|---|---|---|---|
| First Mover | | **PD** | **AG** | **PG** | **CG** |
| | **PD** | D, D | D, D | D, C | D, C |
| | **AG** | D, D | C, C | C, C | D, C |
| | **PG** | C, D | C, C | C, C | C, D |
| | **CG** | C, D | C, C | D, C | D, C |

set of strategic interactions within a coalition.[16] Over the course of collective action, actors move through this space, which alters their strategic interests in cooperation, ultimately causing them to either cooperate or defect. At any given time, the best strategy for a player $i$ depends upon both the relative values of $P$, $R_i$, $S_i$, and $T_i$, and the collective interests of the $n-1$ other actors in the coalition.[17]

Because of heterogeneity in the population, players within a coalition can be located in different regions of the strategy space. Consequently, the set of strategic interactions in a coalition includes not only the four symmetric games corresponding to each region of strategy space, but also a set of asymmetric games, occurring when actors play across regions. To simplify the model, strategy selection within a coalition is based on backward induction. This reduces the possible set of choice outcomes to a series of either dominant strategies (such as defection in the prisoner's dilemma) or conditional strategies (such as defection in the assurance game), depending on an actor's strategic situation. The basic rules for the model are as follows.

Players with a dominant strategy always act accordingly. Thus players located in the prisoner's dilemma region always defect, and players located in the privileged game region always cooperate. Players in a region without a dominant strategy play their best response to their alter. Thus, a first-mover facing a player with a dominant strategy chooses her own best response. A player in the assurance game region will cooperate on the condition that all of the other members of the coalition are also willing to cooperate. Sequential best response can also produce a first-mover advantage; for example, when a player in the chicken game region meets fellow chicken game players, the first player defects, thereby eliciting cooperation from the alter. Table 2 summarizes outcomes by move status and preference type.

In summary, the setup of the dynamic model is based on two ideas: (1) because of group heterogeneity, all individuals within the system do not occupy the same point in strategy space, with some occupying regions that are more conducive to cooperation than others; and (2) through the formation of local coalitions actors can alter the incentives for one another, making cooperation a rational course of action. The agent-based model developed in this study resembles the sequential decision models used in other formal theories of collective action (Heckathorn, 1993, 1996; Kim and Bearman, 1997; Macy, 1990; Marwell and Oliver, 1993). In each iteration, a single player is randomly chosen from the population and activated. The active player $i$ then randomly selects $n-1$ neighbors and constructs a local coalition. Each player in the coalition evaluates the payoffs associated with each of the four possible outcomes of her $N$-person game and chooses the action that yields the greatest payoff. Thus, coalitions only cooperate when it is individually rational for every member of the coalition to cooperate. The players' collective choices result in one of the four outcomes. After the coalition acts, a new player is randomly selected, and the process repeats itself. This procedure is run asynchronously, allowing each coalition's decision to affect the state of the world before a new player is activated. Technical details of the simulation model are provided in Appendix 1.

## *The emergence of critical mass*

The model is initialized in a state of universal defection ($\delta = 0$) with a population size, $N$, of 1000 actors, and a distribution of interests ($V$) such that the population is in an $N$-person assurance game. Each actor has a threshold past which cooperation would be individually rational. However, until the threshold is reached, each actor prefers defection over cooperation. Thus, a critical mass is necessary in order to initiate collective action (Granovetter, 1978; Heckathorn, 1993; Kim and Bearman, 1997; Marwell and Oliver, 1993; Schelling, 1978).

Figure 2 shows the strategic makeup of the population when actors are not allowed to form coalitions. With zero cooperators, the incentives for a single person to initiate cooperation are insufficient to overcome the costs of contributing (i.e., for all actors $i$, $V_iL(1) < K$). Thus, each individual in the population is located in the prisoner's dilemma region of the strategy space: they would like others to cooperate, but are unwilling to cooperate themselves.

To initiate critical mass dynamics and solve the collective dilemma, actors are allowed to self-organize into local coalitions. The coalition size required to initiate successful collective action can be analytically derived from equations (1)–(6), as shown in Appendix 1. Briefly stated, the size of

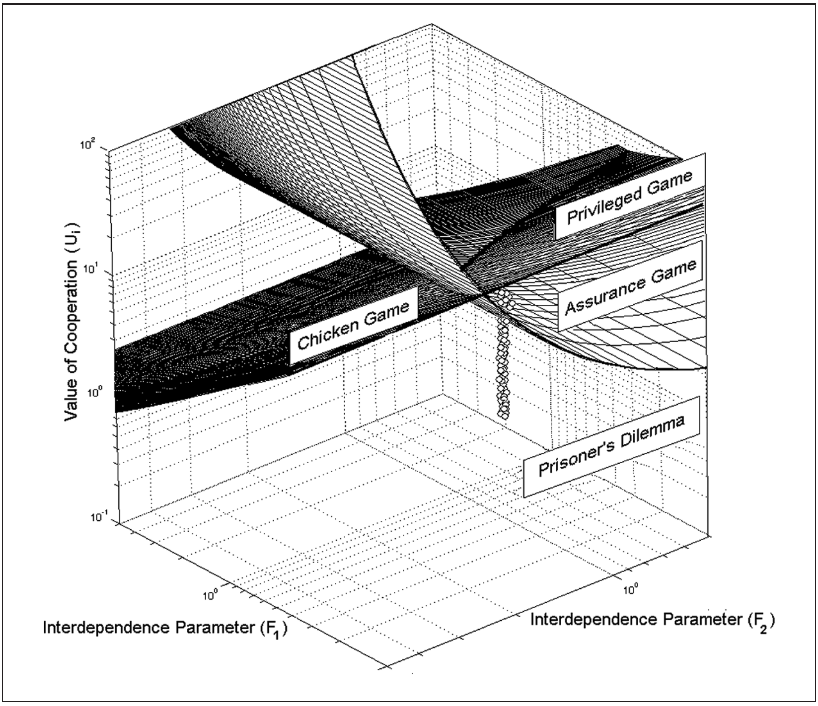


**Figure 3.** Critical mass formation with local coalitions ($N$ = 1000, $J$ = 0.3, $\bar{V} = 600$, $\sigma = 200$, $n$ = 8). The strategic interests of the actors from Figure 1, after they are allowed to form into coalitions of size $n$ = 8. The small group in the assurance game can form a critical mass coalition and initiate collective action.

the critical mass is given by the minimum coalition $n$, such that if actors organize into coalitions of size $n$, at least $n$ people will prefer mutual cooperation to unilateral defection, as given by expression 7.

$$\min(n) \; s.t. \left( \sum_{i=1}^{N} H(R_i - T_i) \right) \geq n \tag{7}$$

For each combination of parameters $N$, $J$, $\bar{V}$, and $\sigma$ we can determine the minimum coalition size, $n$, that is sufficient to initiate collective action. For the settings shown in Figures 2 and 3, the coalition size necessary to provide a critical mass is eight players. Figure 3 illustrates how coalition formation changes the strategic situation for the actors stuck in the prisoner's dilemma in Figure 2 by allowing coalitions of size $n$ = 8 to form. In Figure 3, ten individuals have moved from the prisoner's dilemma region of the strategy space up into the assurance game region. Each of these ten individuals is willing to cooperate if they can find seven other coalition members who will also cooperate.

The population shown in Figures 2 and 3 is located in a fully connected network, which means that everyone is connected to everyone else. Each person can sample the entire population to see if there are others 'like them' who would be willing to cooperate.[18] As long as actors can form coalitions of the required size, they keep searching until they find seven others with whom they can coordinate to initiate collective action – at which point, the start-up problem is solved. As shown in Appendix 1, this model is quite general. For smooth, single-peaked distributions of thresholds for cooperation (see Granovetter, 1978), and production functions with critical mass properties (see Dodds and Watts, 2005; Heckathorn, 1993; Marwell and Oliver, 1993), we can derive the size of the critical mass required to initiate collective action. If actors in the population can form coalitions of the required size, then mobilization will succeed.

Robustness analysis shows that the size of the critical mass (i.e., the required coalition size, $n$) varies only slightly across changes in the mean and standard deviation of $V$.[19] However, changes in the population size ($N$) can have a dramatic effect on the size of the critical mass needed to initiate collective action. Figure 4 shows the change in the size of the critical mass for populations ranging from $N$ = 100 to $N$ = 20,000. Results are shown for four different distributions of $V$.

While increases in the population size ($N$) produce a corresponding increase in the critical mass ($n$), the amount of increase depends upon the jointness of supply of the public good. That is, it depends upon the degree to which an increase in the size of the population that benefits from the public good results in a decrease to each person's utility from contributing to it. For instance, a good with zero jointness is a raffle in which each person who buys a ticket directly reduces everyone else's chances of winning the prize. As the size of the population buying tickets becomes very large, the expected value of buying an additional ticket goes to zero. This presents a bleak picture for large group collective action, since a sizable number of people (10,000 or 20,000) would seemingly make any given individual's contribution not worth the cost. However, *zero* jointness is an extreme that is empirically quite rare. Almost all goods have some level of jointness (Marwell and Oliver, 1993). And, for some goods, like winning an election, or passing a bill, jointness is quite high since, for example, the value of new

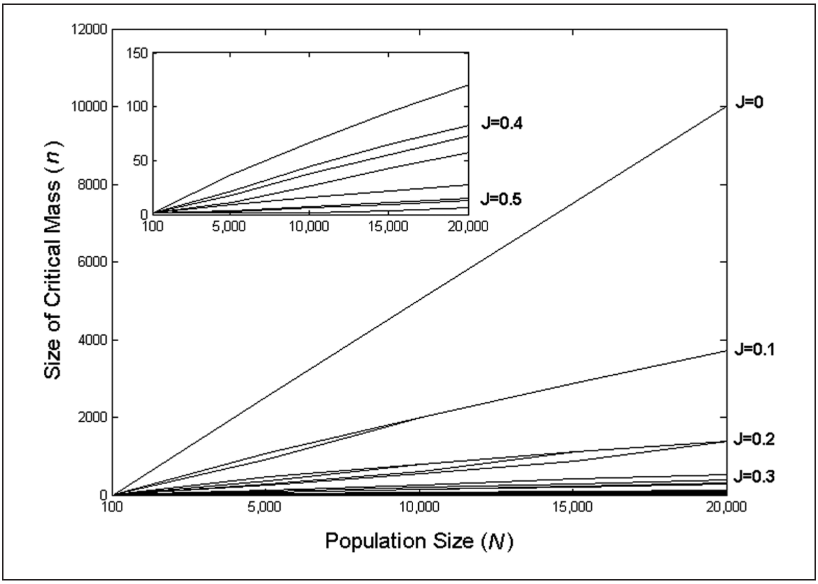


**Figure 4.** Change in critical mass with increasing population size. For four different distributions of the valuation of the public good ($\overline{V} = 200, \sigma = 50$), ($\overline{V} = 200, \sigma = 200$), ($\overline{V} = 800, \sigma = 50$), ($\overline{V} = 800, \sigma = 200$), the size of the critical mass ($n$) increases with increasing population size $N$, shown ranging from $100 \leq N \leq 20,000$. The rate of increase in the critical mass changes dramatically depending on the jointness of supply $J$, shown ranging from $0 \leq J \leq 0.5$.

legislation does not typically depend upon the size of the constituency. Public goods such as creating parks, building bridges, cleaning up the streets, and cleaning up the environment also typically have very high levels of jointness. However, these forms of public good can also suffer from ‘crowding effects’, which may reduce the jointness of supply to less than unity (Marwell and Oliver, 1993). For example, the benefits of clean rivers may be reduced for those who worked to clean them up if the landscape becomes overcrowded with a flood of new tourists.

In sum, the least hospitable conditions for collective action are situations with zero jointness, where every additional person reduces the value of the collective good for the others. The utopia is a jointness of one, where the population can grow to infinity without affecting people’s interests in contributing. Figure 4 demonstrates how the size of the critical mass, *n*, changes with population size as a function of these changing conditions.

The top line indicates zero jointness. As Marwell and Oliver (1993: 43–44) note, "Olson's group-size argument is clearly correct when the good has zero jointness of supply… larger groups are much less likely to be provided with the good than smaller groups." Figure 4 shows that with zero jointness, the required critical mass increases rapidly, growing from about eight people for $N$ of 100, to about 40 people for $N$ of 300. And, for $N$ of 500, the critical mass approaches 100 people. For larger $N$, the size of the critical mass observes a strict regularity of the form $n = .5N$; in other words, the critical mass required is half the size of the population, regardless of $N$. This relationship holds for all four distributions of $V$, which overlap to form a single line in Figure 4.

This, too, seems disparaging for large group collective action. However, small deviations away from zero jointness have a large impact on the size of the critical mass required for mobilization. For a jointness of 0.1, the size of the critical mass observes the ratio $n = 0.19N$. And, for a jointness of 0.2, the ratio drops to $n = 0.07N$.[20] More generally, Figure 5 shows that the relative size of the critical mass ($n/N$) observes a very well-behaved exponential decay with the jointness of supply, well approximated by the curve $n/N = .5e^{kJ}$, where the decay constant $k \approx -9.697$. The open circles in Figure 5 show the ratio of $n/N$ (for $N > 1000$), for increasing values of the jointness of supply ranging from $0 \leq J \leq 0.5$; the solid line plots the exponential decay curve. As shown in the inset in Figure 4, this means that for a collective action with a jointness of 0.5, the critical mass necessary to mobilize a population of 20,000 people is a coalition that ranges in size from 8 to 25 people, depending on the distribution of $V$.

## The dynamics of mobilization

The foregoing analyses show that for any population ($N$, $\bar{V}$, $\sigma$) and collective good ($J$), we can determine if there is a critical mass, $n$, that can initiate collective action. If there is, then the paramount factor for successful mobilization is whether $n$ individuals can form a coalition. This has the interesting implication that distinct collective action problems that have different characteristics, but which have the same $n$ – e.g. ($N = 1000$, $V = 600$, $\sigma = 200$, $J = 0.3$) and ($N = 100$, $V = 500$, $\sigma = 150$, $J = 0$) both have $n = 8$ – will have the same basic mobilization dynamics. If they can form coalitions of size $n$, collective action will succeed.

Once a coalition forms to initiate collective action, bandwagon dynamics take over. As more people cooperate, even more people become willing, and the collective action grows (Finkel et al., 1989; Granovetter, 1978; Marwell

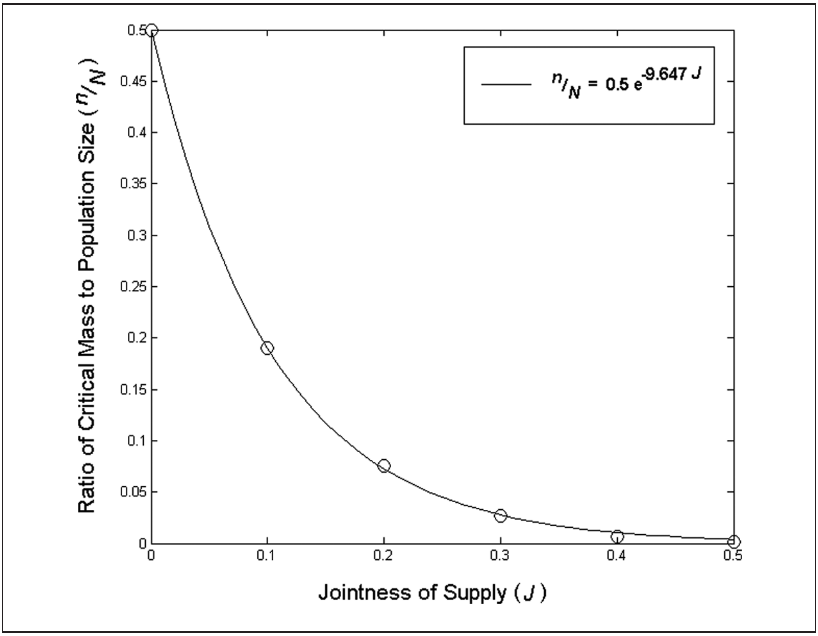


**Figure 5.** Exponential decay in critical mass size. For the four distributions shown in Figure 2, and for population sizes of $N > 1000$. The ratio of the critical mass to the population size, $^{n}/_{N}$ (open circles), decays exponentially with linear increases in the jointness of supply, $J$. The solid line plots the exponential decay curve with a decay constant $k = -9.647$.

and Oliver, 1993; Schelling, 1978). Figure 6 uses the strategy space diagram to show how these collective dynamics unfold through changes in the strategic interests of the population. Once the critical mass cooperates, changes in the level of public goods create new incentives that move actors upward in the strategy space. Actors who had been facing a prisoner's dilemma or an assurance game now find themselves in the privileged game, willing to cooperate as individuals.

While this upward movement through the strategy space is qualitatively the same for the class of production functions that generate critical mass dynamics (*cf.* Dodds and Watts, 2004; Heckathorn, 1996), the long-term trajectory of the population will differ based on the subsequent shape of the production curve. For the S-shaped curve shown in Figure 1, after 50% cooperation the curve starts to decelerate – reducing the incentives for further cooperation. This causes the population to follow an arc-shaped

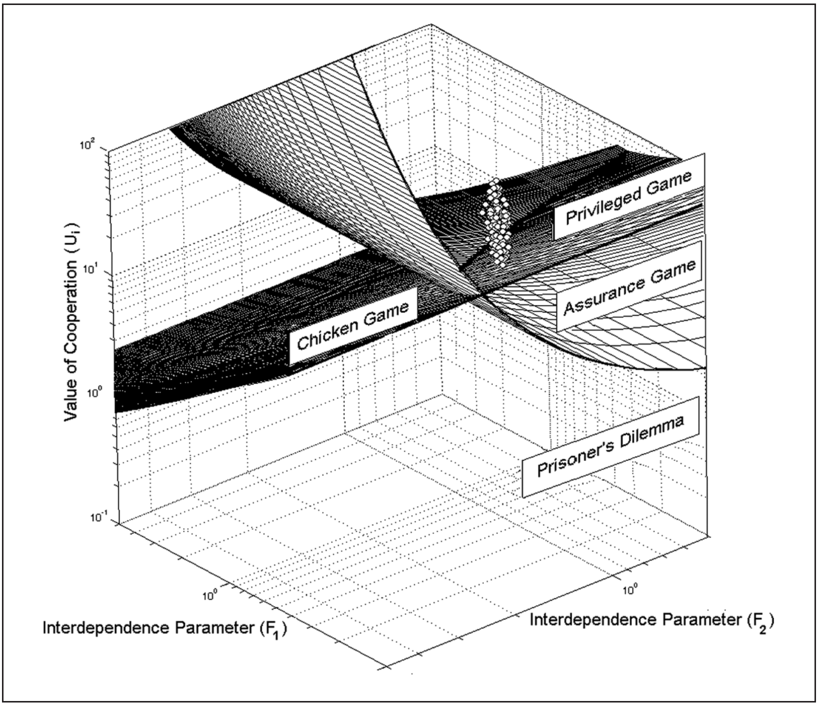


**Figure 6.** Bandwagon dynamics with local coalitions ($N$ = 1000, $J$ = 0.3, $\bar{V}=600$, $\sigma=200$, $n$ = 8). The contributions of the critical mass in Figure 2 create new incentives for other actors to cooperate, moving the population upward through the strategy space.

trajectory, which moves later players down into the chicken game, and results in less than full participation at equilibrium (this is described in detail in Heckathorn, 1996). By contrast, purely convex functions (Heckathorn, 1993) create a continuous upward movement resulting in complete cooperation at the equilibrium (Heckathorn, 1993; Marwell and Oliver, 1993).

## Homophily

The success of the mobilization process illustrated in Figures 3 and 6 raises an important question: Just because there are a sufficient number of individuals willing to cooperate, does this necessarily mean that these people can actually find one another to form a critical mass?

Research on collective action has long emphasized the importance of solidarity (Fireman and Gamson, 1979; Hechter, 1987), collective identity (Collins, 1993; Ferree, 1992), and shared commitment (Heckathorn, 1990; Polletta, 1998) for the emergence of social movements, civic change organizations, and revolutionary actions. More formally, this idea has been expressed as the ability of likeminded people to coordinate in organizing for the collective good (Heckathorn, 1993; Macy, 1990; Marwell and Oliver, 1993). Thus, it is not surprising that homophily – the tendency of people to interact with others who are similar to them – may be an important factor in the mobilization of collective action (Centola, 2011; Kitts et al., 1999; McPherson et al., 2001).

However, recent research on bandwagon dynamics in collective action argues that while moderate amounts of homophily are beneficial, very high levels of homophily can actually impede critical mass dynamics. "[T]he optimal distribution of [interests] across networks would be a pattern where agents associate with a certain proportion of others with similar [interests] while keeping a certain level of friendship with others of discrepant [interests]" (Chiang, 2007: 67). The intuition behind these findings, which have been repeated in other formal studies of norms and cooperation (Boyd and Richerson, 2002; Goyal, 1996; Stark, 1996), comes from a network-centered approach to social dynamics in which the benefits and costs of cooperation are evaluated solely in terms of one's network neighbors. These findings show that excessive homophily can trap collective action in a single region of the network, preventing local initiatives from successfully spreading across the social space.

To see if these findings generalize for critical mass mobilization in large groups, I embed the actors from the above coalition-based model into a social network, and then incrementally alter the level of homophily in the network.[21] At the same time, I systematically alter the size of the critical mass, which allows us to identify the effects of homophily on critical mass mobilization across a range of different start-up problems. I compare the smallest possible coalitions ($n$ = 2) with larger coalitions ($n$ = 4, 6, 8, and 12). This provides me with a means for identifying the interaction effects of coalition size with homophily.

To isolate the effects of homophily, I alter the size of the critical mass, $n$, using the valuations of the public good ($\bar{V}$, $\sigma$), while keeping the population size fixed at $N = 1000$, and the jointness of supply fixed at $J = 0.3$. As shown in Figure 4, the critical mass does not vary greatly with changes to the distribution of $V$. To investigate the interaction of homophily with the size of the critical mass over the interval $n$ = [2, 12], I examine distributions over the parameter range ($400 \leq \bar{V} \leq 1500$, $50 \leq \sigma \leq 400$).

To create the social network, actors are embedded into a simple lattice structure with 'Moore neighborhoods', in which each actor's 'neighborhood' consists of its eight neighbors on a two-dimensional grid, four on the rows and columns (North, South, East, and West) and four on the diagonals. Moore neighborhoods create a spatially constrained social network, with high levels of clustering (i.e., network transitivity), which is useful for studying the dynamics of local interaction (Centola et al., 2005; Centola and Macy, 2007). The number of neighbors that each person has, $z$, can then be increased from 8 to 24 to 48 (and so on) by increasing the neighborhood radius $r$, where $z = 4r(r+1)$. For the following analysis, I use $r = 2$, which gives each node $z = 24$ neighbors. While the public good is shared over the entire population (i.e., the production function is 'global'), actors can only build coalitions using their 24 direct social contacts. Thus, in general, $z$ must be greater than or equal to the minimum coalition size (i.e., $z \geq \min(\pi)$) in order for actors to be able to build successful coalitions.

Homophilous interaction is modeled by controlling the correlation of actors' valuations of the public good, $V$, in the social network. When homophily is zero, there is no correlation between neighbors' valuations of the public good ($\rho \approx 0$). At the other extreme, when homophily is very high, neighbors have highly correlated values of $V$, meaning that every actor's neighbors have valuations of the public good that are similar to hers ($\rho \approx 0.9$).

Figure 7 shows the frequency with which critical mass collective action is successful for five different coalition sizes, ranging from the smallest possible coalitions of $n$ = 2 to larger coalitions of $n$ = 12 (averaged over 100 runs of the model with identical settings).[22] The results show that there is an increasing, positive effect of homophily on mobilization. For every coalition size, the more homophilous that ties are, the greater the likelihood of successful collective action.

As the size of the minimum coalition increases, mobilization becomes increasingly dependent upon homophily. This is because as coalition size increases, it becomes more unlikely that $n$ individuals with strong interests in the collective action will be randomly located within the same part of the social network. For coalitions of $n$ = 2, randomly finding actors to coordinate with is difficult, but not impossible. But, without homophily, randomly finding seven or eight other actors to coordinate with is nearly impossible. The larger the coalition size, the greater the importance of homophily in facilitating the formation of critical mass.[23]

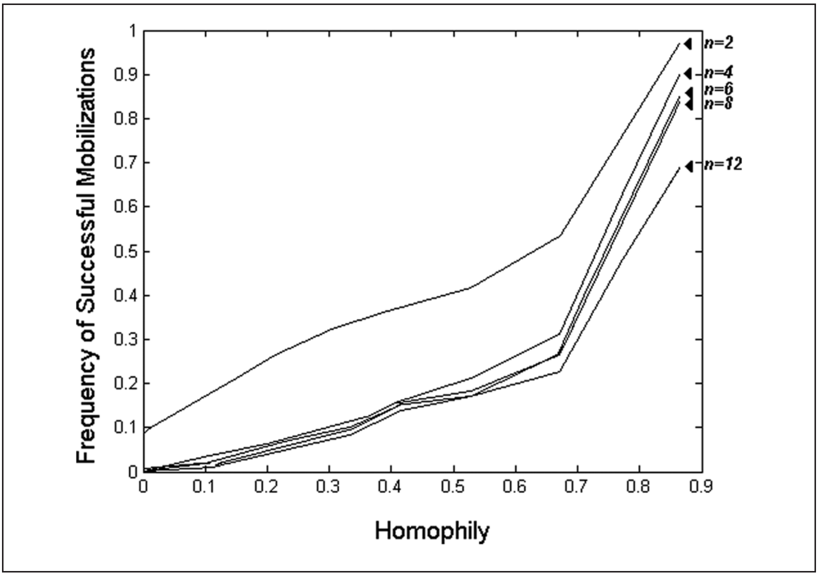


**Figure 7.** Positive effects of homophily ($N$ = 1000, $J$ = 0.3, averaged over 100 realizations). Parameter settings correspond to five different critical mass sizes ($n$ = 2, $\bar{V} = 1000$, $\sigma = 250$), ($n$ = 4, $\bar{V} = 700$, $\sigma = 260$), ($n$ = 6, $\bar{V} = 500$, $\sigma = 280$), ($n$ = 8, $\bar{V} = 500$, $\sigma = 240$), ($n$ = 12, $\bar{V} = 250$, $\sigma = 280$). Frequency of successful mobilizations are shown for increasing levels of homophily. As homophily increases, so does the success of critical mass collective action.

The reason that these results present a noticeable departure from previous network studies of homophily in bandwagon dynamics (Boyd and Richerson, 2002; Chiang, 2007; Chwe, 1999) is that we are studying large group collective action. In this study, the collective good is divided over the entire population. By contrast, in many network-based accounts of collective action (Centola and Macy, 2007; Chiang, 2007; Chwe, 1999; Kim and Bearman, 1997; Macy, 1990), models assume that individuals evaluate their contributions as if the entire collective good were shared exclusively among the local neighborhood of $z$ actors.

This assumption has two consequences. First, it makes the size of the entire population irrelevant, and thus the collective action problem easier to solve. Since the good is only divided over the local neighborhood, the population can increase without bound without affecting actors' willingness to contribute. However, this formalization of collective action fails to address the original concern voiced by Olson (1965), and later by Kim and Bearman

(1997), that the larger the group is, the smaller any given individual's contribution is.

Second, this assumption limits the value of homophily. Because the collective good is not shared over the entire population, actors cannot find out about people's contributions unless they are recruited through their social networks. Since there is no public awareness of the collective good, mobilization depends upon new recruits having local contact with network neighbors who have contributed. Consequently, if high-interest people only interact with other high-interest people, their excessive homophily prevents news of their efforts from helping to recruit lower-interest people in the network. Thus, too much homophily causes the collective action to get stuck in one region of the network, from which it cannot escape. This is the reason why previous formal studies of collective action in social networks have found negative effects of high levels of homophily.

By contrast, while large group collective action is a harder problem to solve, it also benefits more from homophily. In large group collective action, the public good is shared equally over the population. Because of this, there is public awareness of what people are doing, and how it affects the collective good. If grassroots organizers put together a campaign to clean up the neighborhood, the activists working in the streets will let others know that their efforts are under way (Gould, 1993). Similarly, efforts to get out the vote (Sandler, 1992), to ban smoking from public restaurants (Heckathorn, 1993), or to start a strike (Klandermans, 1988), are not limited to people's immediate social contacts. Once a core group of activists mobilize behind a cause, they publicize their efforts, letting everyone know that they are contributing to the collective good, and that others' contributions will make a difference (Marwell and Oliver, 1993; Opp and Gern, 1993; Schelling, 1978). Because of this public signal, the efforts of a homophilous few can be successful at mobilizing a large majority of people. And, the more homophilous the activists are, the greater the chances that they can coordinate to initiate a mass action.

## Network structure

Beyond homophily, another important second way that neighborhood composition can be altered is by changing the topology of the social network.[24] The above homophily model uses highly clustered lattice networks, in which social interaction is limited by spatial proximity. Some empirical research has shown that clustered, spatially constrained social networks can be very important for the mobilization of social movements (Centola, 2010; Gould, 1996; Hedstrom, 1994). However, a large literature on collective action (Chwe, 1999; Granovetter, 1973; McAdam and Paulsen, 1993; Macy, 1991) and collective behavior (Newman, 2000; Watts, 1999, 1999; Watts and Strogatz, 1998) has shown that weak ties, which disrupt the local structure of spatial networks by creating links between otherwise remote actors, can also be beneficial for the spread of cooperation.

To test the effects of network structure on the dynamics of collective action, I extend the above model to study how clustered networks versus weak ties affect critical mass formation. To do this, I use a random rewiring model (Centola and Macy, 2007; Watts and Strogatz, 1998), in which randomly selected ties in the spatial network are broken and then re-attached to randomly selected members of the population, creating 'shortcuts' across the lattice. The rewiring model uses the parameter $q$ ($0 \leq q \leq 1$) to determine the probability that ties will be randomly rewired. For $q = 0$, no ties are rewired and the network is a regular lattice with Moore neighborhoods, while for $q = 1$ every tie is rewired, making the social network into a random graph (Watts and Strogatz, 1998). As the parameter $q$ increases, there are more weak ties in the social network, which reduces local structure and transforms the social topology into a small world network (Centola and Macy, 2007; Watts, 1999).

The results presented in Figure 8 show the three-way interaction between coalition size, homophily and network structure. I studied this interaction by testing the five different coalition sizes used above ($n = 2$, 4, 6, 8, and 12) both with homophily ($\rho \approx 0.9$), and without homophily ($\rho \approx 0$), over the full range of network structures, ranging from a clustered spatial lattice to a random network with abundant weak ties ($0 \leq q \leq 1$). As above, the coalition size is determined by selecting corresponding distributions of $V$. Using a population of 1000 actors and a jointness of supply of $J = 0.3$, each of these distributions place approximately $n$ actors in the assurance game region of the game space. Figure 8 shows the effects of network structure on the mobilization process for coalitions with homophilous (solid lines) and non-homophilous (dotted lines) ties. The *y axis* in Figure 8 shows the frequency of successful mobilization attempts, while along the *x axis* $q$ increases from 0 to 1, indicating that the network structure moves from clustered neighborhoods to abundant weak ties.

For very small coalitions with homophilous ties ($n = 2$, solid line), increasing the number of weak ties does not dramatically affect the success of critical mass mobilization. Across the range of network structures, actors with homophilous relationships are able to find at least one other actor with a similar incentive structure, giving a high probability of success to small

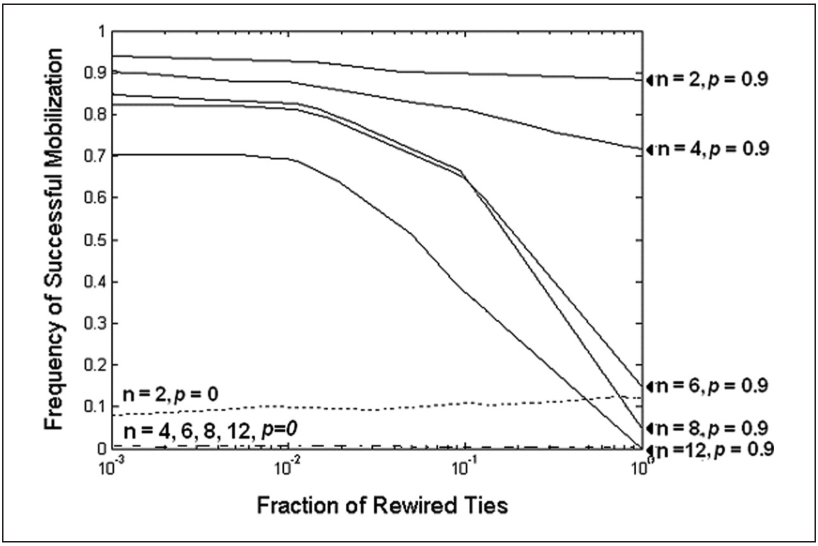


**Figure 8.** Effects of network structure, coalition size, and homophily on critical mass ($N$ = 1000, $J$ = 0.3, averaged over 100 realizations). The frequency of successful mobilizations shown for the same distributions of $V$ and coalition sizes shown in Figure 6 ($n$ = 2, $n$ = 4, $n$ = 6, $n$ = 8, and $n$ = 12), both with homophily ($\rho = 0.9$) and without ($\rho = 0$). The $x$ axis shows the structure of the social network, ranging from $q \approx 0$ (regular lattice) to $q$ = 1 (random network). For the smallest coalitions with homophily ($n$ = 2, $\rho = 0.9$, solid line), there is no significant effect of rewiring the network on the success of mobilization. However, as coalition size increases ($n \geq 4$, $\rho = 0.9$, solid lines), there is a negative effect of increasing the fraction of weak ties on mobilization. When there is no homophily, larger coalitions cannot mobilize at all ($n \geq 4$, $\rho = 0$, dotted lines). Very small coalitions can have modest success even in the absence of homophily ($n = 2$, $\rho = 0$, dotted line). Increasing weak ties can slightly improve the success of mobilization in these cases.

coalitions. For slightly larger coalitions ($n$ = 4), network rewiring has a more noticeable negative effect on the probability of success. And, for even larger coalitions ($n$ = 6), as there are more weak ties in the network we observe a significant effect. Adding random ties reduces the frequency of successful mobilizations. For coalitions of $n$ = 8, success rates drop from more than 80% success in the spatially structured network to nearly zero in the network with abundant weak ties. This striking decline in mobilization frequency comes from the fact that as the network is randomized, transitivity is reduced, making it less likely that assurance game players who are neighbors with each other, will also have neighbors in common. While homophily increases the likelihood that actors with cooperative interests will have similar neighbors, increasing $q$ makes it less likely that a critical mass of these actors will all be socially connected to one another. As the number of weak ties increases, the network becomes too diffuse to ensure the local clustering necessary to mobilize larger coalitions.

For non-homophilous ties, larger coalitions are even more difficult to mobilize. When there is no homophily, mobilization fails entirely for coalitions with $n > 2$, regardless of the network structure. This is because in a large population, it is unlikely that the small fraction of interested cooperators will be randomly connected with each other. However, for the smallest coalitions ($n = 2$, dotted line), each person only needs to find a single partner to organize a collective action, allowing these coalitions to succeed even without homophilous ties. As the spatial structure of the local neighborhoods is perturbed by the addition of weak ties, the only effect is that as $q$ approaches 1 the mobilization frequency increases slightly. Thus, when the coalitions are small – only requiring a single other person to initiate the action – having more shortcuts across the social space provides actors with more opportunities to find another player whose interests correspond to their own (Granovetter, 1973; Macy, 1991). However, these results do not generalize to larger coalitions. The larger the coalitions, the more that mobilization depends upon homophilous interaction in clustered networks.

These results indicate that the structure of the social network interacts with both homophily and coalition size in the dynamics of critical mass mobilization. The implications for the 'strength of weak ties' are divided. For the smallest coalitions, where only a single contact is required in order to make both individuals willing to act, weak ties seem to either not affect mobilization, or to weakly improve the mobilization process. These results are consistent with the findings of Granovetter (1973) and Macy (1991). However, these effects are not very pronounced. The stronger effect is for larger coalitions. For greater $n$, the results weigh in favor of empirical findings stressing the importance of spatial networks (Gould, 1996; Hedstrom, 1994) and strong ties (McAdam and Paulsen, 1993) for local organizing and recruitment to collective action. This also provides formal support for Opp and Gern's (1993) findings that homogenous, clustered networks, in which trust can be established, can play a significant role in the mobilization of collective action. Taken together, these results suggest that the required size of the critical mass has significant implications for the structural conditions that are necessary to support mobilization. Consistent with recent formal and empirical work on 'complex contagions' (Centola, 2010; Centola and Macy, 2007), these results suggest that the more difficult the collective action problem is, the more that mobilization depends upon clustered social networks.

## Conclusion

Formal models of collective action have traditionally used an *N*-person framework for studying large group collective action (Bonacich et al., 1976; Hamburger, 1973; Hardin, 1982; Olson, 1965), in which it is unlikely that an individual's solitary contribution will make a sufficient impact to make cooperation rational. This has led to a by-product theory (Gould, 1993; Olson, 1965) of social cooperation in which the use of selective incentives, and not an interest in the public good, is the principle mechanism for explaining cooperative behavior. Critical mass theory (Marwell and Oliver, 1993) offers an alternative to by-product theory, but requires an explanation of how critical mass mobilization can emerge endogenously among equal resource actors who are mired in the suboptimal equilibrium of universal defection. Network-based approaches have attempted to solve this problem, but most of these models eliminate group size from the individual calculus, and thus avoid the large group problem altogether.

I present a coalition-based model of collective action that solves the problem of critical mass in large groups. My results show that the formation of coalitions can change the strategic incentives of individuals mired in an *N*-person prisoner's dilemma. Self-organization into coalitions can create cooperative incentives that make the formation of a critical mass rational, allowing interested actors to initiate bandwagon dynamics that produce successful collective action. Further, my results show that high levels of homophily can significantly benefit critical mass mobilization, and that weak ties can be detrimental to mobilization as the required coalition size increases. These results help to clarify conflicting claims about the roles of social networks (Gould, 1996; Granovetter, 1973; Hedstrom, 1994; Kitts, 2000) and homophily (Boyd and Richerson, 2002; Chiang, 2007; Kitts et al., 1999; McPherson et al., 2001) in social movement dynamics, and suggest that there is no generic 'ideal network' for mobilizing mass action. The advantages of various structural resources, such as clustered neighbors or weak ties, depend upon the needs of the particular mobilization.

### Funding

I gratefully acknowledge the support of the James S. McDonnell Foundation.

## Notes

1. There is some debate about whether there is a 'critical mass theory'. Marwell and Oliver refer to their (1993) book on critical mass collective action as 'a micro-social theory'. In that spirit, I refer to the idea that critical mass dynamics can explain voluntary collective action in large groups as 'critical mass theory'. The definition of critical mass used here is the same as that used by Schelling (1978), Granovetter (1978), Rogers (1995), Dodds Watts (2004, 2005), and sometimes by Marwell and Oliver (1993); *viz*., the critical mass is the number of people that need to adopt a behavior in order for that behavior to keep spreading. Related work that explores a similar concept of critical mass, but falls outside the scope of this study, includes work following from Macy's (1991) learning-theoretic account, in which critical mass represents a level of collective action that is sufficient to make cooperation self-sustaining. By these definitions, below the critical mass, any initial adopters quickly drop out and there is no bandwagon dynamic. However, Marwell and Oliver (1993) have another definition of critical mass, not used here: the critical mass is a group of high-resource contributors who can produce all or most of the collective good for everyone.
2. Marwell and Oliver (1993) show that an organizer agent can solve this problem when the distribution of ties, interests, and resources coincide to give the person with the most individual ties access to the greatest resources in the population. This solution emphasizes the importance of resource heterogeneity (including social, human, and financial capital) for successful collective action.
3. Thank you to an anonymous reviewer for clarifying this point about resources.
4. Homophily is the principle that people with similar traits are more likely to be socially connected than people with dissimilar traits. Different mechanisms, such as selection and influence, can both play a role in producing this outcome. I will not focus on the particular mechanism here, only on the likelihood of connected actors having similar valuations of the public good.
5. The value of weak ties for mobilization is based on their structural role as 'long-distance' ties in the network, which reduce clustering and connect remote regions of the social space (Centola and Macy, 2007; Granovetter, 1973). Consistent with Watts and Strogatz's (1998) small world model, and Centola and Macy's (2007) complex contagions model, in the present discussion I use the term 'weak ties' to refer to structurally 'long ties'. Affect is not manipulated in the present study.
6. The production function maps levels of contribution onto levels of goods production. This is discussed in detail below. (See Marwell and Oliver, 1993, for an extensive discussion of production functions.)
7. It is also possible for some people to have thresholds that are so high that they would never cooperate. The *N*-person assurance game works the same even in populations with these non-participants, but the equilibrium level of cooperation is less than 100%.

8. Work by game theorists such as Schofield (1975), Ellickson (1973), Buchanan (1965) and others operate outside of the narrow one-versus-all framework, and address the use of coalitions to solve collective action problems. However, these approaches either use side payments within coalitions to organize cooperation (Schofield, 1975), or restrict cooperative behavior to a small subset of the population, such as jurisdictions (Ellickson, 1973) or clubs (Buchanan, 1965), thereby introducing excludability.
9. For robustness, I also explored these dynamics for alternative production functions, and found no significant deviations from these results.
10. $\beta > 1$ is a shape parameter for the production curve. For the results presented below, I use $\beta = 10$ (Heckathorn, 1996; Macy, 1990, 1991)
11. More technically, Dodds and Watts (2005) define the necessary conditions for critical mass as the need for the production function to have a positive second derivative near the origin; that is, returns accelerate as more people the join the action (*cf.* Marwell and Oliver, 1993). This condition holds when payoffs for cooperation are determined by the marginal impact that each person's cooperation has on the public good. This is the standard model of returns used in collective action (Heckathorn, 1996; Marwell and Oliver, 1993; Olson, 1965).
12. The expression $V_iL$ in equation (2) can be expanded to read $V_iL - L(0)$, where $L(0)$ is the level of public goods before the coalition acts. The abbreviated form shown in equation (2), however, is appropriate at start-up, when $L(0) = 0$.
13. For some distributions of $V$ that I study, a small fraction of actors may have negative valuations of the public good. This does not change the dynamics, but it does imply that the equilibrium level of cooperation will be less than 100% (see Note 6).
14. This corresponds to a 'weakest link' structure because every coalition member must prefer mutual cooperation to unilateral defection in order for the collective player to prefer mutual cooperation; that is, the weakest link breaks the chain.
15. Figure 1 also shows that participation is less attractive at the end of the production curve when so many people have participated that additional contributions will not have much impact (see discussion below). See also Marwell and Oliver (1993) and Heckathorn (1996) on 'suboptimal' public goods production.
16. See Heckathorn (1996) for a proof of the completeness of the strategy space.
17. An individual's location in a region of the strategy space does not indicate that the individual is in a symmetric game, but rather that the individual's incentive structure corresponds to an ordering of outcomes associated with a given game type.
18. Subsequent analyses examine the effects of network structure on mobilization.
19. Other model parameters, such as network structure, do not affect the minimum coalition size. As discussed below, the distribution of interests can cause some modest variation in coalition size, but the dominant effect comes from the values of $N$ and $J$.
20. Figure 4 also shows that while these ratios are approximate for smaller values of $N$, they converge in the limit of large $N$.

21. Homophily is implemented using a global smoothing function (García-Ojalvo and Sancho, 1999), which iterates over the population and attempts to create a smooth, single-peaked distribution of values in the network (i.e., all high valued individuals in the same neighborhood). The algorithm iterates until the correlations between neighbors reach the level that is specified by the Pearson's correlation coefficient ($\rho$).
22. Coalition sizes are determined by selecting the corresponding distributions of $V$: ($n$ = 2, $\bar{V}=1000$, $\sigma=250$), ($n$ = 4, $\bar{V}=700$, $\sigma=260$), ($n$ = 6, $\bar{V}=500$, $\sigma=280$), ($n$ = 8, $\bar{V}=500$, $\sigma=240$), ($n$ = 12, $\bar{V}=250$, $\sigma=280$). Successful coalition formation is harder to achieve as $n$ approaches $z$. For large, sparse networks (i.e., $z << N$), altering the size of the neighborhoods does not qualitatively affect these results.
23. These results require clustering in the network, but do not require that the clustered network be a lattice.
24. Earlier work on social networks in critical mass dynamics (Marwell et al., 1988) shows that: (1) network density and centrality allow high-resource individuals to be more easily mobilized; (2) high-influence individuals can be more effective in mobilization (Gould, 1993; Kim and Bearman, 1997); and (3) imitation and homophily can affect mobilization in complete graphs (Kitts et al., 1999). The present study assumes actors have equal resource and influence, and are embedded in large networks with interactions limited to immediate neighbors. I focus on how the overall topological structure of the network (e.g., clustering and path length) facilitates or inhibits the emergence of critical mass. See Centola (2010), Centola et al. (2007), Centola and Macy (2007), Watts and Strogatz (1998).

## References

Bonacich P, Shure GH, Kahan JP, et al. (1976) Cooperation and group size in the N-person prisoner's dilemma. *Journal of Conflict Resolution* 20: 687–706.

Boyd R and Richerson P (2002) Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology* 215: 287–296.

Buchanan J (1965) An economic theory of clubs. *Economica* 32: 1–14.

Centola D (2010) The spread of behavior in an online social network experiment. *Science* 329: 1194–1197.

Centola D (2011) An experimental study of homophily in the adoption of health behavior. *Science* 334: 1269–1272.

Centola D, Eguiluz V and Macy M (2007) Cascade dynamics of complex propagation. *Physica A* 374: 449–456.

Centola D and Macy M (2007) Complex contagions and the weakness of long ties. *American Journal of Sociology* 113: 702–734.

Centola D, Willer R and Macy M (2005) The emperor's dilemma: A computational model of self-enforcing norms. *American Journal of Sociology* 110: 1009–1040.

Chiang Y-S (2007) Birds of moderately different feathers: Bandwagon dynamics and the threshold heterogeneity of network neighbors. *Journal of Mathematical Sociology* 31: 47–69.

Chong D (1991) *Collective Action and the Civil Rights Movement*. Chicago, IL: University of Chicago Press.

Chwe M (1999) Structure and strategy in collective action. *American Journal of Sociology* 105: 128–156.

Collins R (1993) Emotional energy as the common denominator of rational action. *Rationality and Society* 5: 203–230.

Dodds P and Watts D (2004) Universal behavior in a generalized model of contagion. *Physical Review Letters* 92: 218701.

Dodds P and Watts D (2005) A generalized model of social and biological contagion. *Journal of Theoretical Biology* 232: 587–604.

Ellickson B (1973) A generalization of the pure theory of public goods. *American Economic Review* 63: 417–432.

Ferree MM (1992) The political context of rationality: Rational choice theory and resource mobilization. In: Morris A and Mueller C (eds) *Frontiers in Social Movement Theory*. New Haven, CT: Yale University Press, pp.29–52.

Finkel SE, Muller EN and Opp K-D (1989) Personal influence, collective rationality, and mass political action. *American Political Science Review* 83: 885–903.

Fireman B and Gamson WA (1979) Utilitarian logic in the resource mobilization perspective. In: Zald MN and McCarthy JD (eds) *The Dynamics of Social Movements*. Cambridge, MA: Winthrop, pp.8–45.

García-Ojalvo J and Sancho JM (1999) *Noise in Spatially Extended Systems*. New York: Springer-Verlag.

Gould RV (1993) Collective action and network structure. *American Sociological Review* 58: 182–196.

Gould RV (1996) *Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune*. Chicago, IL: University of Chicago Press.

Goyal S (1996) Interaction structure and social change. *Journal of Institutional and Theoretical Economics* 152(3): 472–494.

Granovetter M (1973) The strength of weak ties. *American Journal of Sociology* 78: 1360–1380.

Granovetter M (1978) Threshold models of collective behavior. *American Journal of Sociology* 83: 1420–1443.

Hamburger H (1973) N-person prisoner's dilemma. *Journal of Mathematical Sociology* 3: 27–48.

Hardin R (1982) *Collective Action*. Baltimore, MD: Johns Hopkins University Press.

Hechter M (1987) *Principles of Group Solidarity*. Los Angeles, CA: University of California Press.

Heckathorn DD (1990) Collective sanctions and compliance norms: A formal theory of group-mediated social control. *American Sociological Review* 55: 366–384.

Heckathorn DD (1993) Collective action and group heterogeneity: Voluntary provision versus selective incentives. *American Sociological Review* 58: 329–350.

Heckathorn DD (1996) Dynamics and dilemmas of collective action. *American Sociological Review* 61: 250–277.

Hedstrom P (1994) Contagious collectivities: On the spatial diffusion of Swedish trade unions. *American Journal of Sociology* 99: 1157–1179.

Kim H and Bearman P (1997) The structure and dynamics of movement participation. *American Sociological Review* 62: 70–93.

Kitts J (2000) Mobilizing in black boxes: Social networks and social movement organizations. *Mobilization: An International Journal* 5: 241–257.

Kitts J (2006) Collective action, rival incentives, and the emergence of antisocial norms. *American Sociological Review* 71: 235–259.

Kitts J (2008) Dynamics and stability of collective action norms. *Journal of Mathematical Sociology* 32: 142–163.

Kitts J, Macy M and Flache A (1999) Structural learning: Attraction and conformity in task-oriented groups. *Computational and Mathematical Organization Theory* 5: 129–145.

Klandermans B (1988) Union action and the free-rider dilemma. In: Kriesberg L and Misztal B (eds) *Research in Social Movements, Conflict and Change: Social Movements as a Factor of Change in the Contemporary World*, vol. 10. Greenwich, CT: JAI Press, pp.77–92.

McAdam D (1986) Recruitment to high-risk activism: The case of freedom summer. *American Journal of Sociology* 92: 64–90.

McAdam D (1988) *Freedom Summer*. New York: Oxford University Press.

McAdam D and Paulsen R (1993) Specifying the relationship between social ties and activism. *American Journal of Sociology* 99: 640–667.

McPherson JM, Smith-Lovin L and Cook J (2001) Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–444.

Macy MW (1990) Learning theory and the logic of critical mass. *American Sociological Review* 55: 809–826.

Macy MW (1991) Chains of cooperation: Threshold effects in collective action. *American Sociological Review* 56: 730–747.

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

Marwell G, Oliver P and Prahl R (1988) Social networks and collective action: A theory of the critical mass. III. *American Journal of Sociology* 94: 502–534.

Newman ME (2000) Models of the small world: A review. *Journal of Statistical Physics* 101: 819–841.

Oliver P (1993) Formal models of collective action. *Annual Review of Sociology* 19: 271–300.

Oliver P and Marwell G (1988) The paradox of group size in collective action: Theory of critical mass. II. *American Sociological Review* 53: 1–8.

Oliver P and Marwell G (2001) Whatever happened to critical mass theory? A retrospective and assessment. *Sociological Theory* 19(3): 292–311.
Oliver P, Marwell G and Teixeira R (1985) A theory of the critical mass, I. Interdependence, group heterogeneity, and the production of collective goods. *American Journal of Sociology* 91: 522–556.
Olson M (1965) *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.
Opp K-D and Gern C (1993) Dissident groups, personal networks, and spontaneous cooperation: The East German revolution of 1989. *American Sociological Review* 58: 659–680.
Polletta F (1998) ‘It was Like a Fever’: Narrative and identity in social protest. *Social Problems* 25: 137–139.
Prahl R, Marwell G and Oliver P (1991) Recruitment for collective action under conditions of interdependence: A theory of the critical mass. V. *Journal of Mathematical Sociology* 12: 137–164.
Rasler K (1996) Concessions, repression, and political protest in the Iranian revolution. *American Sociological Review* 61: 132–153.
Rogers EM (1995) *Diffusion of Innovations*. New York: Free Press.
Runge CF (1984) Institutions and the free rider: the assurance problem in collective action. *Journal of Politics* 4: 154–181.
Sandler T (1992) *Collective action: Theory and Applications*. Ann Arbor, MI: The University of Michigan Press.
Schelling T (1978) *Micromotives and Macrobehavior*. New York: W.W. Norton & Co.
Schofield N (1975) A game theoretic analysis of Olson’s game of collective action. *Journal of Conflict Resolution* 19: 441–461.
Stark R (1996) *The Rise of Christianity*. Princeton, NJ: Princeton University Press.
Watts D (1999) *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ: Princeton University Press.
Watts DJ and Strogatz SH (1998) Collective dynamics of ‘Small-World’ networks. *Nature* 393: 440–442.
Wood E (2003) *Insurgent Collective Action and Civil War in El Salvador*. New York: Cambridge University Press.
Yin C-C (1998) Equilibria of collective action in different distributions of protest thresholds. *Public Choice* 97: 535–567.
Young P (2009) Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review* 99: 1899–1924.

## Appendix 1

### *Formal definition of the model*

***Production function and payoffs.*** Equations (1)–(6) governing the production function and the individual payoffs are stated in the text. They are repeated here for completeness. They are as follows:

Production function:

$$L(\pi) = \frac{1}{1 + e^{(.5 - \frac{\pi + \delta}{N^{1-J}})\beta}} \tag{1}$$

Individual payoff:

$$U_i = V_i L - C_i K \tag{2}$$

Payoff for mutual cooperation:

$$R_i = V_i L(n) - K \tag{3}$$

Payoff for unilateral cooperation:

$$S_i = V_i L(1) - K \tag{4}$$

Payoff for unilateral defection:

$$T_i = V_i L(n - 1) \tag{5}$$

Payoff for mutual defection:

$$P = 0 \tag{6}$$

*The strategy space.* The strategy space shown in Figures 2, 3, and 6 has three axes. The *z*-axis is based on the value $U_i$ from equation (2). The *x* and *y* axes correspond to the changing shape of the production function in terms of two parameters, $F_1$ and $F_2$ (*cf.* Heckathorn, 1996). The changes in the values of $F_1$ and $F_2$ indicate the changing interdependence of actors in a coalition (Heckathorn, 1993; Marwell and Oliver, 1993). $F_1$ indicates the effect of the first member's decision whether to contribute, on the decisions made by the rest of the coalition members.

$$F_1 = \frac{\ln\left(\frac{L(n)-L(1)}{L(n)-L(0)}\right)}{\ln\left(\frac{n-1}{n}\right)} \tag{8}$$

$F_2$ is the complement of $F_1$. $F_2$ reports the effect of the decisions made by the first $n$–1 players in the coalition on the final player's decision. Taken together, $F_1$ and $F_2$ describe the interdependence between individual and group decision-making within a coalition.

$$F_2 = \frac{\ln\left(\frac{L(n)-L(n-1)}{L(n)-L(0)}\right)}{\ln\left(\frac{1}{n}\right)} \tag{9}$$

The planes in Figures 2, 3, and 6 show where two outcomes are equally valued ($S = P$, $T = R$, and $R = P$), consequently each of the five regions bordered by the planes corresponds to a unique ordinal ranking of $T$, $R$, $P$, and $S$. Since this ordinal ranking determines the structure of individual and collective interests (or 'game type'), each region of the strategy space in these figures corresponds to a unique strategic situation.

*The decision model.* As discussed in the text, an active individual $i$ plays a 2x2 game against the remaining $n$–1 members of the coalition aggregated into an alter $j$. I use a 'weakest link' rule for determining $j$'s strategy, which aggregates the $n$–1 players' interests as shown in equations (10) and (11).

$$R_j = \sum_{k=1}^{n-1} H(R_k - T_k),\ k \neq i \tag{10}$$

In order for $j$ to prefer mutual cooperation to unilateral defection ($R_j = n–1$), every member of $j$ must prefer mutual cooperation to unilateral defection. Similarly, every member of the coalition must have $S_k > 0$ in order for $j$ to prefer unilateral cooperation ($S_j$=$n$–1) over mutual defection ($S_j$<$n$–1), as shown in equation (11).

$$S_j = \sum_{k=1}^{n-1} H(S_k),\ k \neq i \tag{11}$$

The active player, $i$'s, decision whether to cooperate or defect is based the payoffs for individual and mutual cooperation or defection, as described in the text. The decision algorithm is provided in equation (12).

$$C_i = \begin{cases} 0 \quad \text{if } T_i > R_i \text{ and } P_i > S_i \\ 1 \quad \text{if } R_i > T_i \text{ and } S_i > P_i \\ \text{if } R_i > T_i \text{ and } P_i > S_i : \\ \begin{cases} 1 & \mathit{if}\ (R_j = n-1) \\ 0 & \text{otherwise} \end{cases} \\ \text{if } T_i > R_i \text{ and } S_i > P_i : \\ \begin{cases} 0 & \mathit{if}\ (S_j = n-1) \\ 1 & \text{otherwise} \end{cases} \end{cases} \tag{12}$$

*The simulation model.* Each run of the simulation model specifies a number of agents ($N$), a jointness of supply ($J$), and a distribution of valuations of the public good ($\overline{V}, \sigma$). The model initializes each agent by drawing a random value $V_i$ from the distribution of $V$, and setting the agent's state to $C_i$ = 0. The value of $n$ (coalition size) is the same for all agents, and it is determined by the calculation given in Appendix 2.

The model iterates as follows:

1. Select an agent $i$ at random from the population.
2. Agent $i$ randomly selects $n$–1 individuals from the population.

In the fully connected network, agents are randomly chosen from the population.
In the local network model, agents randomly select the $n$–1 players from their immediate network neighborhood.

3. Each agent iterates through the calculation of their individual best options using equations (1) through (6), and (10) through (12). Each agent decides their best option for cooperation or defection.
4. The active agent $i$ then calculates the outcome of each player's decision.

If all $n$ players including $i$ agree to cooperate, then the outcome of the game is R, mutual cooperation.
If only $i$ agrees to cooperate, then $i$ cooperates and everyone else defects.
If everyone else agrees to cooperate, and $i$ does not, then everyone else cooperates and $i$ defects
If no one agrees to cooperate, then everyone defects.

5. Once this action is taken, the model iterates again from step 1.

# Appendix 2

## *Solution to the size of the critical mass*

Given a production function $L(\pi)$, and a distribution of interests in participation $f(V)$, we are required to find the minimum coalition size, $\pi$, such that for $\pi$ individuals mutual cooperation, $R_i = V_iL(\pi) - K$, dominates unilateral defection, $T_i = V_iL(\pi - 1)$.

For an individual $i$, the cooperation point where $R_i > T_i$ occurs when

$$V_iL(\pi) - K > V_iL(\pi - 1) \tag{13}$$

which can be rewritten as

$$L(\pi) - L(\pi - 1) > \frac{K}{V_i} \tag{14}$$

For a population of size $N$, the size of the minimum coalition to create critical mass is given by

$$\min(\pi)\ s.t.\ \left[\sum_{i=1}^{N} H\left(L(\pi) - L(\pi - 1) - \frac{K}{V_i}\right)\right] \geq \pi \tag{15}$$

where the Heaviside function $H(x)$ is given by

$$H(x) = \begin{cases} 0 \ \ if \ \ x \leq 0 \\ 1 \ \ if \ \ x > 0 \end{cases} \tag{16}$$

Equation (15) states that the minimum coalition $\pi$, is the smallest coalition that is sufficient to give $\pi$ individuals incentives for mutual cooperation over unilateral defection.

To provide an example, let us take the production function given in equation (1).

$$L(\pi) = \frac{1}{1 + e^{(.5 - \frac{\pi + \delta}{N^{1-J}})\beta}}$$

We set $\delta = 0$ since there is zero cooperation at the start, which gives us the equation:

$$\min(\pi)\ s.t.\ \left[\sum_{i=1}^{N} H\left(\frac{1}{1 + e^{(.5 - \frac{\pi}{N^{1-J}})\beta}} - \frac{1}{1 + e^{(.5 - \frac{\pi - 1}{N^{1-J}})\beta}} - \frac{K}{V_i}\right)\right] \geq \pi \tag{17}$$

Assuming, for example, $N = 1000$, $J = 0.3$, $\beta = 10$, and a normal distribution of thresholds with $\bar{V} = 600$, $\sigma = 200$, this gives us a minimum coalition size of $\pi = 8$. Given the stochastic nature of the distribution, this value is approximate, changing slightly with different realizations of the threshold distribution. As $N$ becomes large, this variation between realizations of distribution becomes smaller, and the prediction becomes more exact.