

Social Reinforcement: Cascades, Entrapment, and Tipping[†]

By GEOFFREY HEAL AND HOWARD KUNREUTHER*

The actions of different agents sometimes reinforce each other. Examples are network effects and the threshold models used by sociologists as well as (Harvey) Leibenstein's "bandwagon effects." We model such situations as a game with increasing differences, and show that tipping of equilibria, cascading, and clubs with entrapment are natural consequences of this mutual reinforcement. If there are several equilibria, one of which Pareto dominates, then the inefficient equilibria can be tipped to the efficient one, a result of interest in the context of coordination problems. We characterize the smallest tipping set. (JEL C72, D80, D85, Z13)

The choices made by different individuals can reinforce each other. Network externalities provide an obvious example. Explanations for Microsoft's domination of the operating system market, eBay's domination of the online auction market, and the permanence of the QWERTY keyboard layout have been based on network effects.¹ The rapid growth in popularity of social networking sites such as Facebook and MySpace also illustrates the importance of social reinforcement. People join because their friends have already done so. In his theory of entrapment, Avinash Dixit (2003) exploits reinforcing effects of people's choices to show how some may be entrapped into joining a club in spite of the fact that its existence makes them worse off.² There is an early precedent for the importance of social reinforcement in the work of Harvey Leibenstein (1950). Citing, as a precedent, the work of John Rae (1834), Leibenstein analyses situations in which one's demand for a good increases with the number of others also buying it, using the term "bandwagon effects" to describe such situations. This is an early theory of fads and fashions, and is based on the recognition that other people's actions can reinforce one's own choices. Thomas Schelling's (1978) work on tipping exemplifies the same insights. His iconic example

*Heal: Columbia Business School, 616 Uris Hall, 3022 Broadway, New York, NY 10027-6902 (e-mail: gmh1@columbia.edu); Kunreuther: Wharton School of the University of Pennsylvania, Philadelphia, PA (e-mail: kunreuther@wharton.upenn.edu). We are especially grateful to Larry Samuelson for assistance with this paper, and also grateful to Doug Bernheim, Vince Crawford, Rachel Croson, Avinash Dixit, Charley Holt, Alan Kirman, and a referee for constructive comments. Partial support from National Science Foundation Grant CMS-0527598 is gratefully acknowledged.

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at <http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.1.86>.

¹ See Carl Shapiro and Hal R. Varian (1999) and W. Brian Arthur (1989). For reviews of this area see Charles F. Manski (2000), and Nicholas Economides (1995).

² The popularity of the QWERTY keyboard may be an illustration of entrapment, in the sense that a redesigned keyboard would enable everyone to type much faster than they currently do. However, there is no incentive for any one individual to learn the new system on his own because the keyboard he needs to purchase will have to be specially made and, hence, will be relatively expensive. Furthermore, it may not be available in other locations in which he is required to type.

is of a sudden change in the racial composition of a neighborhood (Schelling 1971). Nonwhites move into an originally white neighborhood, and when the proportion reaches a critical level the neighborhood tips, and the remaining whites move out. Underlying this is an assumed (and hopefully outdated!) preference for neighbors of one's own color, so that the movement into the neighborhood of nonwhites is mutually reinforcing, as is the movement out of whites. Recent work by David Card, Alexandre Mas, and Jesse Rothstein (2008) tests for tipping of the racial composition of residential neighborhoods in the United States over the period 1970–2000, and finds strong confirmation that once-white neighborhoods have tipped to all-minority neighborhoods once the minority population passes a relatively low threshold.

Mark Granovetter (1978) has described similar processes in the sociology literature. He discusses the adoption of new behaviors, which he models in terms of individuals' adoption thresholds.³ Again, an action by one person makes similar actions by others more attractive by moving them toward or across their adoption thresholds. These threshold effects can be modelled by utility-maximizing choices when utility depends on the choices of others. Granovetter (1978) gives an interesting example of entrapment into criminal behavior by groups of young males and cites many other examples of mutually reinforcing choices, from the adoption of birth-control practices in Korea to migration choices in third world countries, as well as education choices, and joining strikes or riots. Duncan J. Watts (2002) presents an analysis of cascades in a network of people all of whom show threshold effects in their behavior. Each agent is most influenced by those to whom she is nearest in the network, and Watts presents results on the probability of a cascade when the network is a random graph.⁴ William A. Brock (2004), drawing on results from Brock and Steven N. Durlauf (1995), looks at tipping in the context of dynamical systems, and also builds on the idea of social reinforcement. In his models, the payoff of a choice depends on the choices of others; and there is a penalty for being "unfashionable," which introduces the social reinforcement element.

The idea of social reinforcement has also been used in the finance literature. Harrison G. Hong, Jeffrey D. Kubik, and Jeremy C. Stein (2001) develop and test empirically the idea that stock market participation is affected by social interactions, and that a person's chances of investing are greater if most of his peers also invest. Another application in the area of finance is to "positive feedback trading" in which investors buy more of an asset that has recently increased in value (Bradford J. DeLong et al. 1990 and Nicholas Barberis and Andre Shleifer 2003). So if some investors buy and raise prices, then others follow suit, leading to just the kind of threshold effects discussed by Granovetter (1978) and Watts (2002). As we will show, this behavior can be explained by a model in which social interactions are valued by decision makers. Abhijit V. Banerjee (1992) also develops a model of herd behavior that appears in some ways similar. People are influenced by the choices of others on the assumption (not always correct) that these choices are based on private information.

³ A person's adoption threshold is the number of others she must see engaging in a new behavior before she too adopts that behavior. For early adopters this number is low and for late adopters it is high.

⁴ A cascade is the movement of the group from one pattern of behavior to another by a sequence of individual changes, just like the classic image of a sequence of dominoes falling.

In cases of fashion-oriented behavior discussed by Leibenstein (1950), there is a perceived intrinsic value to being like others. In cases such as securities trading or adopting new habits, the value of following others is not intrinsic but rather is derived. Seeing the others make a choice gives the impression that this choice is less risky than it would otherwise appear to be. In trading models, transactions may reveal private information and allow this to diffuse through the market. Dunia López-Pintado and Watts (2006) make a similar distinction, but instead of the terms intrinsic and derived, they use the terms explicit and implicit externalities.

In all of these diverse situations, individuals' choices can reinforce each other. Someone else choosing X over Y makes that choice more attractive to me. In game-theoretic terms, these are all games that show the increasing differences property that is associated with supermodularity. Agent i 's payoff to a choice increases if j makes that choice as well. One can also think of this as strategic complementarity.

Here, we formulate such reinforcing situations in game-theoretic terms, and model tipping, cascading, and entrapment as properties of the Nash equilibria of games. Using a static game-theoretic framework, we show that the potential for tipping and cascading will be widespread when games display increasing differences or social reinforcement. This means that a subset of the participants, sometimes a very small subset, can shift the system from one equilibrium to another just by changing their choices. This is the point of Schelling's (1978) work, though he makes it in the context of dynamic processes rather than Nash equilibria. We also show that if there are two (or more) equilibria, one of which Pareto dominates, then under very general conditions there is a coalition of agents who can tip the inefficient equilibrium to the efficient one. This is an interesting insight into the resolution of certain types of coordination problems. Finally, we give a characterization of tipping sets for a family of symmetric games.

Our work originated from research on mutual reinforcement in the context of national security. Originally motivated by a desire to understand the impact of interdependence in airline security after the September 11, 2001 terrorist attacks, it has evolved to a more general model of how interdependence and reinforcement affects the incentive to invest in protective measures for any kind of network, including electronic networks such as computer networks.⁵ One of our early findings was that many networks exhibit a tipping phenomenon with respect to investments in security. For certain values of state variables few agents invest and the system is vulnerable and insecure. A small change can lead to everyone investing with a massive increase in security.

A. Examples

Consider a game with N agents in which each agent i has two strategies: $s_i = 0$ or 1. We use s_{-i} to denote the choices of all agents other than i and take $N = 10$. The payoffs are

$$u_i = 0.5 \quad \text{if } s_i = 0 \quad \text{and} \quad \#(1) \quad \text{if } s_i = 1,$$

⁵ For the national security applications see Heal and Kunreuther (2005) and Kunreuther and Heal (2003), and for computer network applications see Michael Kearns (2005).

where $\#(1)$ is the number of 1s in s_{-i} , i.e., the number of other agents who choose 1. There is social reinforcement in the choice of strategy 1. The more people choose it, the more attractive it is. In game theoretic terms, choosing 1 rather than 0 shows increasing differences. Clearly $\{0, 0, \dots, 0\}$ is a Nash equilibrium. Likewise, $\{1, 1, \dots, 1\}$ is also a Nash equilibrium, which Pareto dominates. Any agent can tip the equilibrium of zeros to that of ones. Starting from this equilibrium, if agent 1 changes from 0 to 1 then the payoffs to all other agents from choosing 1 are now $1 > 0.5$, so that 1 is their best response. As every agent is better off at $\{1, 1, \dots, 1\}$ than at $\{0, 0, \dots, 0\}$, it may seem obvious that any agent would tip the latter to the former equilibrium. But the fact remains that when everyone else plays 0, any agent's best response is to play 0 too. To make it rational for an agent to tip, we would have to think of a multi-round version of the game. In the first round, agent i would play 1 in response to everyone else's 0, and would lose from doing so, but in the next round, all others would respond with 1 and, if this situation were to be maintained, agent i would compensate for her first-round loss.

Note that in contrast the equilibrium of ones can be tipped to that of zeros only by a coalition of all but one player, the trivial tipping coalition. In this case, the equilibrium of ones seems stable and that of zeros seems unstable. We can easily modify this example to be more like the Schelling tipping examples:

$$u_i = 4 \quad \text{if } s_i = 0 \quad \text{and} \quad \#(1) \quad \text{if } s_i = 1.$$

In this case, we need five people to choose 1 to tip the equilibrium of zeros, which is now more stable. The equilibrium of ones is correspondingly less stable.

Now, consider a more complex example. The payoffs are

$$u_i(s_i, s_{-i}) = 0.91i \quad \text{if } s_i = 0 \quad \text{and} \quad = \#(1) \quad \text{if } s_i = 1.$$

Again, there is social reinforcement in the choice of strategy 1, but now agents are heterogeneous with respect to the payoff to 0. Again, we assume that $N = 10$. Again, we have equilibria of all zeros and all ones. In this case, agent $N = 10$ can tip the equilibrium of zeros.⁶ Note that no agent other than 10 can tip the equilibrium of zeros. Also, note that agent 10 can tip the equilibrium of ones back to that of zeros.⁷ So, in this case, with $N = 10$, there is only one agent who can tip, and he can tip in either direction, from all zeros to all ones or vice versa. He alone can determine which equilibrium is chosen. In a certain sense he is a dictator.

All players except player 10 are worse off at the equilibrium of zeros than at the equilibrium of ones (player 10 is better off), so when player 10 tips an equilibrium of ones to one of zeros, he is making everyone else worse off, even though their best

⁶ If agent 10 changes from 0 to 1, then the payoff to choosing 1 for any other agent is now 1. As $1 > 0.91$, agent 1 will change too. But now the payoff to any other agent from choosing 1 is 2, and as $2 > 1.82$, agent 2 will change also. This logic continues until all agents have changed, so that the only Nash equilibrium consistent with N choosing 1 is all 1s. Agent 10 starts a cascade.

⁷ If all are choosing 1 and then agent 10 changes, the payoff to 9 from choosing 1 is 8, and the payoff to choosing 0 is 8.19. Now there are two agents choosing zero, so for agent 8 the payoff to 1 is 7 and to 0 is 7.28. The change by agent 10 initiates a cascade from one equilibrium to another.

responses are now to choose 0. This illustrates the issue of entrapment discussed in Dixit (2003). In this example, player 10 can entrap all others at the equilibrium of all zeros.⁸

This second example can be thought of as a threshold model along the lines discussed by Granovetter (1978). Strategy 0 is the status quo, from which agent i will move if the payoff from moving exceeds $0.91i$. Essentially this means that agent 1 has a threshold of one, agent 2 has a threshold of two, etc. An interpretation of this example is that a uniform distribution of thresholds can generate cascading.

It was a stochastic version of the second example that led to our interest in these issues:

$$u_i(0) = 0.91i \quad \text{and} \quad u_i(1) = 10 \frac{\#(1)}{N-1} + 0.5 \left(1 - \frac{\#(1)}{N-1} \right).$$

So, the outcome of strategy 0 is certain, whereas that of strategy 1 is either 10 or 0.5 with probabilities $[\#(1)/(N-1)]$ or $[1 - (\#(1)/(N-1))]$. This structure arises in our earlier analysis of airline security problems, where the payoff to investing in security (strategy 1) depends on and increases with the number of other airlines that also invest, $\#(1)$ (Heal and Kunreuther 2005). This example has most of the properties of the second (deterministic) example discussed above.

B. Thresholds

The concept of a threshold used in the sociological literature (Granovetter 1978; Watts 2002) can be modelled by interactions that display social reinforcement (and increasing differences) in the utility functions. Consider an agent who has to choose between 0 and 1, the payoffs to which are

$$u_i(0) = a_i \quad \text{and} \quad u_i(1) = \#(1)^{1/2}.$$

So there are social reinforcement effects in the choice of alternative 1, and these display diminishing returns. Clearly she will choose alternative 1 if $\#(1) > a_i^2$, and this is agent i 's threshold for choosing 1 over 0. This, in essence, is Leibenstein's model, and underlies the discussions of Granovetter (1978) and Watts (2002). Note that the social reinforcement represented by the term $\#(1)^{1/2}$ could arise, as in Leibenstein's case, from the intrinsic merits of being similar to others, or could reflect informational gains from seeing others adopting choice 1 and prospering from it.

⁸ Granovetter's (1978) discussion of criminal behavior also seems to fit this framework. Talking about the behavior of delinquent boys, he states that "Most did not think it 'right' to commit illegal acts or even particularly want to do so. But group interaction was such that none could admit this without loss of status: in our terms, their threshold for stealing cars is low because daring, masculine acts bring status, and reluctance to join, once others have, carries the high cost of being labelled a 'sissy.'" (Granovetter 1978, 1,435) So illegal moves by those with low thresholds "entrap" others into following suit and make them worse off.

I. Tipping and Increasing Differences

Consider a game with N players $i \in \{1, 2, \dots, N\}$, each choosing a strategy s_i from the discrete set $\{0, 1\}$ and having a utility function $u_i : \{0, 1\}^N \rightarrow R$ that depends on the choices of all agents. We have a natural order on the set of strategy vectors $\{0, 1\}^N$ given by the standard vector ordering⁹ on R^N .

We assume that each agent's payoff function u_i shows what we term *uniform strict increasing differences* in the choices of strategies by other agents. Formally, this means that if 0_i or 1_i denotes a 0 or 1 in the i^{th} position of the vector \mathbf{S} of all strategy choices, and \mathbf{S}_{-i} denotes the vector of choices of all agents other than i , then $\exists \epsilon > 0$ such that if $\mathbf{S}'_{-i} > \mathbf{S}_{-i}$, then

$$(1) \quad u_i(1_i, \mathbf{S}'_{-i}) - u_i(0_i, \mathbf{S}'_{-i}) \geq \epsilon + u_i(1_i, \mathbf{S}_{-i}) - u_i(0_i, \mathbf{S}_{-i}).$$

This implies that the payoff to agent i of changing from zero to one increases by at least ϵ if another agent changes from zero to one. This is our formalization of "social reinforcement." All of the examples discussed above satisfy this condition. In order to talk about tipping, we shall assume that the game has (at least) two Nash equilibria, $\{0, 0, \dots, 0\}$ and $\{1, 1, \dots, 1\}$. A policy-maker will naturally be interested in ruling out the inefficient equilibrium and ensuring an efficient outcome, as in a coordination problem (Vincent P. Crawford and Hans Haller 1990; Walter P. Heller 1986). We study conditions under which it is possible to do this by changing the strategies of a subset of the players. This is the tipping problem: a "tipping set" of agents can shift the equilibrium from one extreme to the other by changing their strategies. This "tipping set" is the set that can "entrap" other agents, using Dixit's term.

Let T be an arbitrary subset of players. We are going to investigate whether agents in the set T can "tip" the system, i.e. can shift the equilibrium from $\{0, 0, \dots, 0\}$ to $\{1, 1, \dots, 1\}$ by changing strategy. To do this, we define the T -game as the above game with the additional constraint that for all players in T the only permissible strategy choice is $s_i = 1$. If the T -game has $\{1, 1, \dots, 1\}$ as its only equilibrium, then we say that T is a tipping set or T -set. The key point here is that when agents in T choose strategy 1, this is also the best response for all agents not in T . So those in T can lead others to change strategy by changing their own strategies.

A set is a *minimal T-set* if it is a T -set, and no subset is a T -set. Clearly if T is a T -set then, getting the members of T to adopt strategy 1 will rule out the equilibrium of zeros. Members of the set T can tip the equilibrium and can force the system to the efficient outcome. If T is a small subset of N , then this can be an important policy tool.

Now, we show that in certain cases a minimal T -set can be formed by a simple algorithm in which we rank agents by a very natural characteristic and then pick the first $K \leq N$ in this ranking. Intuitively the characteristic is a measure of the changes in other agents' payoffs that result when an agent changes her strategy from 0 to 1. It is a measure of the externalities that an agent generates, and a

⁹ We use $A > B$ to show that A exceeds B in at least one component and is no less in all components, and $A \geq B$ to show that it is at least as great in all components.

measure of the degree to which she can reinforce the choices of others. Next, we give conditions for the existence of a tipping set and note that such a set always exists at an equilibrium that is Pareto dominated by another. All proofs are in the Appendix.

PROPOSITION 1: *Under assumption 1 of uniform strict increasing differences, and with a large enough number of agents, there exists a tipping set with less than $N - 1$ agents that tips the equilibrium with all 0s to that with all 1s.*

COROLLARY 2: *If there are two equilibria, one of which Pareto dominates the other, then with uniform strict increasing differences, and with a large enough number of agents, there is a nontrivial tipping set that tips the dominated to the undominated equilibrium.*

It is easy to see the need for the condition that the number of agents is ‘large enough.’ Each time any agent changes strategy from 0 to 1, the payoff to every other agent from such a change increases. For some agents this payoff is initially negative. For the system to tip, the payoff from the change has to be positive for every agent. Therefore, we need enough agents to change to bring the most negative payoffs above zero, and for this to be possible we need enough agents. The number of agents k that is ‘large enough’ depends on the parameter ϵ in the definition of increasing difference in equation (1). The larger ϵ is, the smaller the critical number. To be precise, we show in the proof of Proposition 1 that¹⁰

$$k > 1 + \frac{\text{Max}_i \{u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i)\}}{\epsilon},$$

so that the number of agents has to exceed one plus the ratio of the maximum payoff to a change in strategy from 0 to 1 when all others choose 0, to the minimum reinforcement effect. If the reinforcement effect is large relative to the payoff change, the number of agents needed is small, and with a small reinforcement effect, many agents are needed.

These results have implications for coordination problems. These may be solvable if we can identify tipping sets. To be interesting, this requires that these sets be significantly smaller than the population as a whole. The examples have shown that this can be the case, and the next proposition sheds some more light on the nature of tipping sets. If agents $T \subset N$ form a tipping set and can shift the equilibrium from all zeros to all ones and gain in the transition, then, in an intuitive sense, it is rational for them to coordinate and change the equilibrium. But this statement can only make sense outside of the context of the original Nash game.

Let $(0^{N-2}, 1_i, 1_j)$ denote a vector of strategy choices consisting of zeros in all positions other than i and j with 1s in the i -th and j -th positions. In order to provide

¹⁰ $(0^{N-1}, 1_i)$ is a vector with $(N-1)$ zeros and a 1 as the i -th coordinate, see the discussion preceding equation (A1) in the Appendix.

a simple characterization of a T -set, we focus on the special case in which the difference $\Delta_{ij}(0)$:

$$(2) \quad \Delta_{ij}(0) = [u_j(0^{N-2}, 1_i, 1_j) - u_j(0^{N-2}, 1_i, 0_j)] \\ - [u_j(0^{N-2}, 0_i, 1_j) - u_j(0^{N-2}, 0_i, 0_j)]$$

is independent of the index j , i.e., the effects of i 's change of strategy are symmetric over other agents. In addition, we assume that $\Delta_{ij}(s_{-i-j})$ (see equation (A1) in the Appendix) is independent of s_{-i-j} and does not depend on the strategies chosen by others. So the process of social reinforcement is symmetric. This rules out close personal ties, such as arise if my payoff is affected more by the behavior of friends and colleagues than by that of people not known personally to me. We call these two conditions of symmetry and independence, taken together, condition SI.

$$(SI) \quad \Delta_{ij}(0) = \Delta_{ik}(0) = \Delta_i(0) = \Delta_i$$

A sufficient condition for SI is that an agents's payoff depends on the number of other agents choosing each strategy and not on their identities. It is implied by an anonymity condition.¹¹ For each agent i , Δ_i is the alteration in the payoff that all other agents get from switching strategy from 0 to 1 as a result of agent i changing from 0 to 1, a uniform externality that i , by changing strategy, imposes on others when they change strategy.

Given this, agents can be ranked unambiguously by the values of their Δ_i functions, and we assume without loss of generality that they are numbered so that $\Delta_1 \geq \Delta_2 \geq \dots \geq \Delta_N$. An agent's ability to tip the inefficient equilibrium is measured by its Δ , and we show that a minimal T -set consists of agents with the greatest Δ s.

PROPOSITION 3: *Given SI, if a minimal T -set exists, then for some integer F it consists of the first F agents when agents are ranked by the value of Δ_i .*

Proposition 3 shows that the agents that are most capable of changing the game's equilibrium are those that generate the largest externalities to others or play the greatest role in social reinforcement. They may be people who are perceived as leaders in their community. Within the structure defined by SI, we can say that increasing differences being sufficiently large¹² is necessary and sufficient for tipping of the inefficient equilibrium. A numerical example of tipping at the inefficient equilibrium of a super modular game is given in Heal and Kunreuther (2005).

Cascading.—A cascade is a sequence of events in which a change of strategy by one agent leads another to change, the changes of the two together lead a third to change, and so on. It is a version of the classic domino effect. Dixit (2003) models

¹¹ Such conditions are common in social choice problems. For an illustration of such a condition in a similar context to ours, see Todd Sandler (1992).

¹² In the sense of (A5) in the Appendix.

this well, and we follow his framework. In our context, a cascade will begin from an equilibrium where all agents choose $s_i = 0$. A cascade of length k is a situation in which if 1 were to change from 0 to 1 but all others were to remain at 0, then 2's best response would be 1. If 1 and 2 were to choose 1 and all others 0, then 3's best response would be 1. If 1, 2, and 3 were to choose 1 and all others 0, then 4's best response would be 1, and so on up to and including agent k . The strategy tuple in which agents 1 through k choose 1, and all others choose 0, is a Nash equilibrium, and k may be less than N .

If we think of the game as one in which moves are made sequentially by players in ascending order, if the first mover chooses 1 (perhaps as a result of factors outside the game as we have defined it, such as policy inducements), then the second mover chooses 1 and so on up to and including the k -th mover, and thereafter all will choose 0 and the outcome will be an equilibrium. If the only equilibria involve either all zeros or all ones, then the outcome of such a cascade will be the equilibrium with all 1s, and this will be attained from the equilibrium of zeros by persuading agent 1 to change strategy.

Formally we have a cascade of length k at the Nash equilibrium $\{0, 0, \dots, 0\}$ if agents can be numbered so that agent 2's best response to $\{1, s_2, 0, \dots, 0\}$ is $s_2 = 1$, agent 3's best response to $\{1, 1, s_3, 0, \dots, 0\}$ is $s_3 = 1$, and for all agents j for $j \leq k$ the best response to $\{1, 1, \dots, s_j, 0, \dots, 0\}$ is $s_j = 1$, and in addition $\{1, 1, \dots, s_k = 1, 0, \dots, 0\}$ is a Nash equilibrium. Using the framework and assumptions of the previous section, we can set out sufficient conditions for a cascade of length k .

We can give a formal characterization of the conditions for a cascade of length k as follows.

PROPOSITION 4: *A cascade of length k occurs if for all $2 \leq j \leq k$,*

$$\Delta_{j-1} \geq u_j \left(\overbrace{1, \dots, 1}^{j-2}, 0, \dots, 0 \right) - u_j \left(\overbrace{1, \dots, 1}^{j-2}, 0, 1, 0, \dots, 0 \right),$$

and for $k < j \leq N$, the opposite inequality holds.

The proof is obvious. Cascading, like tipping, depends on a game exhibiting "enough increasing difference."

II. Applications

A. Schelling's Work

Schelling (1978) provides a number of examples of the role of a critical mass in inducing tipping. In these examples individuals make a decision about being part of a process or group based on what they see others doing. A key example is given by individuals' decisions about whether to reside in a neighborhood, which they do if there are enough others like themselves who are already there. Schelling's most famous example of racial segregation in residential districts was essentially dynamic

with a sequence of changes evolving over time (Schelling 1971). However it is possible to capture much of what was interesting in and essential to that model with a static formulation identical to that used above.

Consider a population of P people of a certain type living in a neighborhood. Each has two possible strategies: stay S or move M . The payoff to each depends on how many others of the group do the same. The payoff to staying is the number of others who stay, $\#(S)$, and the payoff to moving is the number of others who move, $\#(M)$. Clearly all staying or all moving are both Nash equilibria, and if $\#(M) > \#(S)$, then the best response of all who have not moved is to move, so that we have the possibility of tipping. This game displays increasing differences, as the payoff to changing from S to M increases with the number of people who have already changed.

B. Security

As shown in Kunreuther and Heal (2003) there is a wide range of security-related problems that exhibit features of supermodularity and in which tipping could occur. One area that naturally falls into this class is computer security. Here, the central issue is the incentive each agent has to invest in protecting himself against a possible virus, when he knows that he may be infected from other agents. There is an interesting example in Kearns (2005) that illustrates this problem in the context of a shared hard drive.

III. Conclusions

Social reinforcement of choices is widespread, and may even be the norm in many areas of behavior. It can be modelled by games showing increasing differences, and naturally generates a propensity for tipping and cascading. We have proven that for a wide range of such situations there are nontrivial tipping sets, and have characterized them for a class of symmetric cases.

APPENDIX

PROOF OF PROPOSITION 1:

Key to our analysis is the effect on agent j 's payoff of changing strategy from 0 to 1, and how this effect changes when another agent, say i , also changes from 0 to 1. How much does i 's move reinforce j ? By the increasing differences property (1), we know that the change by i will increase the effect on j 's payoff of the change by j . Let s_{-i-j} , 1_i , 0_j denote the vector of strategies in which all agents k other than i, j are choosing $s_k \in s_{-i-j}$ and i, j are choosing 1 and 0, respectively. Define

$$\Delta_j(i = 0, s_{-i-j}) = u_j(s_{-i-j}, 0_i, 1_j) - u_j(s_{-i-j}, 0_i, 0_j)$$

and

$$\Delta_j(i = 1, s_{-i-j}) = u_j(s_{-i-j}, 1_i, 1_j) - u_j(s_{-i-j}, 1_i, 0_j).$$

These are the returns to agent j from changing from 0 to 1 when agent i chooses either 0 (in the first line) or 1 (in the second line) and everyone else chooses s_{-i-j} . The difference between these returns is

$$(A1) \quad \Delta_{ij}(s_{-i-j}) = \Delta_j(i = 1, s_{-i-j}) - \Delta_j(i = 0, s_{-i-j}) \geq 0.$$

This is the increase in the return to j 's change of strategy as a result of i 's change of strategy, and from the condition of increasing differences (1) we know that this is positive. It is a measure of the positive externalities or social reinforcement generated by a change of i 's strategy, such reinforcement being guaranteed by increasing differences. As more agents i change their strategy from 0 to 1 there will be a greater increase in utility for the other agents j in the system.

We focus on equation (A1) when all agents other than i and j are choosing strategy 0 so as to derive conditions for tipping the equilibrium of zeros to that of ones:

$$(A2) \quad \Delta_{ij}(0) = [u_j(0^{N-2}, 1_i, 1_j) - u_j(0^{N-2}, 1_i, 0_j)] \\ - [u_j(0^{N-2}, 0_i, 1_j) - u_j(0^{N-2}, 0_i, 0_j)],$$

where 0^{N-2} indicates that there are $N - 2$ zeros in the positions other than i and j . Note that if all ones is a Nash equilibrium, then if all agents other than i choose strategy 1, i 's best response must be 1, so that $N - 1$ agents form a trivial tipping set. For a tipping set to be interesting, it must contain fewer than $N - 1$ agents.

Consider the following sequence of inequalities, which link the equilibrium with all zeros to that with all ones in a series of steps in each of which an additional agent changes strategy from 0 to 1, and which hold by the increasing differences property of equation (1):

$$(A3) \quad u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i) + \epsilon < u_i(0^{N-2}, 1_1, 1_i) - u_i(0^{N-2}, 1_1, 0_i) \\ u_i(0^{N-2}, 1_1, 1_i) - u_i(0^{N-2}, 1_1, 0_i) + \epsilon < u_i(0^{N-3}, 1_1, 1_2, 1_i) - u_i(0^{N-3}, 1_1, 1_2, 0_i) \\ u_i(1_1, 1_2, \dots, 1_{N-2}, 0_j, 1_i) - u_i(1_1, 1_2, \dots, 1_{N-2}, 0_j, 0_i) + \epsilon \\ < u_i(1_1, 1_2, \dots, 1_{N-1}, 1_i) - u_i(1_1, 1_2, \dots, 1_{N-1}, 0_i).$$

If we take the first inequality,

$$u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i) + \epsilon < u_i(0^{N-2}, 1_1, 1_i) - u_i(0^{N-2}, 1_1, 0_i),$$

we see that the payoff to agent i from a strategy change is raised by at least ϵ when agent 1 also picks strategy 1. The second inequality,

$$u_i(0^{N-2}, 1_1, 1_i) - u_i(0^{N-2}, 1_1, 0_i) + \epsilon < u_i(0^{N-3}, 1_1, 1_2, 1_i) - u_i(0^{N-3}, 1_1, 1_2, 0_i),$$

shows that the payoff to i from the change from 0 to 1 is again increased by ϵ when agent 2 changes from 0 to 1. The inequalities repeat this process changing one additional agent's strategy each time. Working back from a typical inequality in (A3) we have that

$$u_i(0^{N-k}, 1_1, 1_2, 1_i) - u_i(0^{N-k}, 1_1, 1_2, 0_i) > (k-1)\epsilon + u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i).$$

Note that $u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i) < 0$ as the vector of all zeros is a Nash equilibrium, so 0 is a best response for i . Note also that, to the contrary, the last difference $u_i(1_1, 1_2, \dots, 1_{N-1}, 1_i) - u_i(1_1, 1_2, \dots, 1_{N-1}, 0_i) > 0$ is positive as the vector of all ones is also a Nash equilibrium, and now 1 is a best response. As the sequence of differences starts negative and ends positive it must change sign: for N sufficiently large there will be a $k < N-1$ such that $(k-1)\epsilon - u_i(0^{N-1}, 1_i) + u_i(0^{N-1}, 0_i) > 0$, and the first k agents will then form a T -set. To be precise, we need k to satisfy

$$k > 1 + \frac{u_i(0^{N-1}, 1_i) - u_i(0^{N-1}, 0_i)}{\epsilon} \text{ for all } i.$$

Thus, k has to exceed one plus the maximum over all agents of the ratio of the change in agent i 's payoff from changing from 0 to 1 when all others choose 0 to the parameter ϵ , which indicates the minimum magnitude of the reinforcement effects between agents. The larger the reinforcement effects, the smaller the value of k needed. Once the first k agents have chosen 1 as a strategy, this is the best response of all other agents. This proves that a T -set exists and is not the trivial tipping set of all agents but one.

PROOF OF PROPOSITION 2:

Let $\{0^t, 1^{N-t-1}, 1_k\}$ denote the following vector: the k th coordinate is 1, t coordinates are zero, all others (of which there are $N-t-1$) are 1, and the first $N-t-1$ coordinates are 1 if $k > N-t$, and the first $N-t$ are 1 otherwise.

From (A1) and (A3) and (SI), we can write

$$(A4) \quad u_j(0^{N-K-1}, 1^K, 1_j) - u_j(0^{N-K-1}, 1^K, 0) = u_j(0^{N-1}, 1) - u_j(0^{N-1}, 0) + \sum_{i=1}^{i=K-1} \Delta_i.$$

Hence, finding a t such that $u_i(0^{N-t}, 1_1, \dots, 1_t, 1_i) - u_i(0^{N-t}, 1_1, \dots, 1_t, 0_i) > 0$ is the same as finding a t such that $u_j(0^{N-1}, 1) - u_j(0^{N-1}, 0) + \sum_{i=1}^{i=t-1} \Delta_i > 0$ or $\sum_{i=1}^{i=t-1} \Delta_i > u_j(0^{N-1}, 0) - u_j(0^{N-1}, 1)$.

If $F < N$ is a T -set, then for all $j > F$ we must have $u_j(0^{N-F-1}, 1^F, 1_j) - u_j(0^{N-F}, 1^F) \geq 0$. By (A4), this inequality is equivalent to

$$(A5) \quad \sum_{i=1}^{i=F-1} \Delta_i \geq u_j(0^{N-1}, 0) - u_j(0^{N-1}, 1) \quad \forall j > F.$$

To construct a minimal T -set, we need to find the smallest number F for which equation (A5) holds. We get this by ranking agents by the size of Δ_i and first choosing first those with the largest value of Δ_i , i.e., those that create the largest externalities or that exhibit increasing differences to the greatest degree.

REFERENCES

- Arthur, W. Brian.** 1989. "Competing Technologies, Increasing Returns, and Lock-in by Historical Events." *Economic Journal*, 99(394): 116–31.
- Banerjee, Abhijit V.** 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107(3): 797–817.
- Barberis, Nicholas, and Andrei Shleifer.** 2003. "Style Investing." *Journal of Financial Economics*, 68(2): 161–99.
- Brock, William A.** 2004. "Tipping Points, Abrupt Opinion Changes, and Punctuated Policy Change." <http://www.econ.wisc.edu/archive/wp2003-28.pdf>.
- Brock, William A., and Steven N. Durlauf.** 1995. "Discrete Choice with Social Interactions I: Theory." National Bureau of Economic Research Working Paper 5291.
- Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer.** 1985. "Multimarket Oligopoly: Strategic Substitutes and Complements." *Journal of Political Economy*, 93(3): 488–511.
- Card, David, Alexandre Mas, and Jesse Rothstein.** 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics*, 123(1): 177–218.
- Crawford, Vincent P., and Hans Haller.** 1990. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." *Econometrica*, 58(3): 571–95.
- DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann.** 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, 45(2): 379–95.
- Dixit, Avinash.** 2003. "Clubs with Entrapment." *American Economic Review*, 93(5): 1824–36.
- Easterly, William.** 2005. "Empirics of Strategic Interdependence: The Case of the Racial Tipping Point." New York University Development Research Institute Working Paper 5.
- Economides, Nicholas.** 1995. "The Economics of Networks." <http://www.stern.nyu.edu/networks/94-24.pdf>.
- Gladwell, Malcolm.** 2000. *The Tipping Point: How Little Things Can Make a Big Difference*. New York, NY: Little, Brown and Company.
- Granovetter, Mark.** 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology*, 83(6): 1420–43.
- Heal, Geoffrey, and Howard Kunreuther.** 2005. "IDS Models of Airline Security." *Journal of Conflict Resolution*, 49(2): 201–17.
- Heller, Walter P.** 1986. "Coordination Failure Under Complete Markets with Applications to Effective Demand." In *Essays in Honor of Kenneth J. Arrow: Volume II, Equilibrium Analysis*, ed. Walter P. Heller, Ross M. Starr, and David A. Starrett, 155–75. Cambridge, UK: Cambridge University Press.
- Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein.** 2001. "Social Interaction and Stock-Market Participation." National Bureau of Economic Research Working Paper 8358.
- Kearns, Michael.** 2005. "Economics, Computer Science, and Policy." *Issues in Science and Technology*, 21(2): 37–47.
- Kunreuther, Howard, and Geoffrey Heal.** 2003. "Interdependent Security." *Journal of Risk and Uncertainty*, 26(2–3): 231–49.
- Kunreuther, Howard, and Geoffrey Heal.** 2005. "Interdependencies within an Organization." In *Organizational Encounters with Risk*, ed. Bridget Hutter and Michael Power, 190–208. Cambridge, UK: Cambridge University Press.
- Leibenstein, Harvey.** 1950. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand." *Quarterly Journal of Economics*, 64(2): 183–207.
- López-Pintado, Dunia, and Duncan J. Watts.** 2006. "Social Influence, Binary Decisions and Collective Dynamics." http://www.feem-web.it/ctn/papers/07_lopez-pintado.pdf.
- Manski, Charles F.** 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*, 14(3): 115–36.
- Rae, John.** 1834. Republished as *The Sociological Theory of Capital*. London: The Macmillan Company 1905.

- Sandler, Todd.** 1992. *Collective Action: Theory and Applications*. Ann Arbor, MI: University of Michigan Press.
- Schelling, Thomas.** 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology*, 1(2): 143–86.
- Schelling, Thomas.** 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton & Company.
- Shapiro, Carl, and Hal R. Varian.** 1999. *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press.
- Watts, Duncan J.** 2002. "A Simple Model of Global Cascades on Random Networks." *Proceedings of the National Academy of Sciences*, 99(9): 5766–71.