NBER WORKING PAPER SERIES

SOCIAL REINFORCEMENT:
CASCADES, ENTRAPMENT AND TIPPING

Geoffrey Heal
Howard Kunreuther

Working Paper 13579
http://www.nber.org/papers/w13579

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2007

We are especially grateful Larry Samuelson for assistance with this paper, and also grateful to Doug Bernheim, Vince Crawford, Rachel Croson, Avinash Dixit, Charley Holt and Alan Kirman for constructive comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Reinforcement: Cascades, Entrapment and Tipping
Geoffrey Heal and Howard Kunreuther
NBER Working Paper No. 13579
November 2007
JEL No. D20,D80,D85,Q59

ABSTRACT

There are many social situations in which the actions of different agents reinforce each other. These include network effects and the threshold models used by sociologists (Granovetter, Watts) as well as Leibenstein's "bandwagon effects." We model such situations as a game with increasing differences, and show that tipping of equilibria as discussed by Schelling, cascading and Dixit's results on clubs with entrapment are natural consequences of this mutual reinforcement. If there are several equilibria, one of which Pareto dominates, then we show that the inefficient equilibria can be tipped to the efficient one, a result of interest in the context of coordination problems.

Geoffrey Heal
Graduate School of Business
616 Uris Hall
Columbia University
New York, NY 10027-6902
and NBER
gmh1@columbia.edu

Howard Kunreuther
Operations and Information Management
The Wharton School
University of Pennsylvania
3730 Walnut Street, 500 JMHH
Philadelphia, PA 19104-6366
and NBER
kunreuther@wharton.upenn.edu

1 Introduction

The choices made by different individuals can sometimes reinforce each other. Network externalities provide an obvious example: your joining a network makes it more attractive for me to join, and vice versa. 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 (Shapiro and Varian [27], Arthur [1]. For reviews of this area see Manski [22] and Economides [10].). 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, Dixit [8] 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.[1] There is an early precedent for the importance of social reinforcement in the 1950 work of Leibenstein on "Bandwagon, Snob and Veblen Effects" (Leibenstein [20]). Citing as a precedent the 1834 work of Rae [24], Leibenstein analyses situations where my 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 again is based on the recognition that other people's actions can reinforce my choices. Schelling's work on tipping exemplifies the same insights ([26]): his iconic example is of a sudden change in the racial composition of a neighborhood. Non-whites 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 in of non-whites is mutually reinforcing, as is the movement out of whites.

In the sociology literature, Granovetter [12] has described similar processes. He discusses the adoption of new behaviors, which he models in terms of individuals' adoption thresholds. 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 high. Again, in this framework an action by one person makes similar actions by others more attractive to them by moving them towards or across their adoption thresholds. We show below that these threshold effects can be modelled by utility-maximizing choices when utility depends on the choices of others. Granovetter 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

[1] The popularity of the QWWERTY keyboard may also 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 relatively expensive. Furthermore it may not be available in other locations in which he is required to type.

through migration choices in third world countries, education choices, and joining strikes or riots. Watts [32] 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. (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.) Brock [3], drawing on results from Brock and Durlauf [4], looks at tipping in the context of dynamical systems, and also builds on the idea of social reinforcement. In his models the payoff to a choice depends on the choices of others and there is a penalty for being "unfashionable," which introduces the social reinforcement element.

Heal [13] discusses an example from the environmental field, the spread of unleaded gasoline. The use of unleaded gasoline required technological innovations in the design of engines and refineries, with significant fixed costs. These innovations were made, and unleaded gas first introduced, in the U.S. Once this had happened, it was easier for other countries to adopt unleaded auto fuel, the fixed costs having been paid. A really nice detail is that Germany adopted unleaded fuel before Italy. As many German tourists visit Italy by car, and their business is important in parts of Italy, unleaded gas was introduced in many areas there before it was legally required. Adoption by the U.S. and Germany made adoption by Italy very straightforward - more so than it had been for the other two countries. This is an example of social reinforcement at the national level leading to a cascade of adopters following an initial adopter. Wagner [31] also finds evidence of a cascading effect as countries ratified the Montreal Protocol on Substances that Deplete the Ozone Layer. From data on the timing of ratifications, he concludes that there was a reinforcing effect: one country ratifying made others more likely to follow suit. It seems likely that something similar will eventually happen with responses to climate change: one country may develop new technologies that reduce the cost of lowering CO2 emissions, and this will make joining a treaty such as the Kyoto Protocol more attractive to others. Another example of this process is the installation of air bags in cars. Mercedes Benz first installed them in a select number of cars at very high average costs. As they and other car manufacturers decided to make this standard equipment the economies of mass production reduced the price significantly, increasing the returns to adoption by others.

The idea of social reinforcement has also been used in the finance literature: Hong, Kubik and Stein [16] 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," where investors buy more of an asset that has recently increased in value (DeLong et al. [7], Barberis and Schleifer [2]). So if some investors buy and raise prices, then others follow suit, leading to just the kind of threshold effects discussed by Granovetter and Watts. As we show below, this behavior can be explained by a model in which social interactions are valued by decision-makers.

In cases of fashion-oriented behavior discussed by Leibenstein, 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 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. Lopez-Pintado and Watts [21] make a similar distinction, but instead of the terms intrinsic and derived use the terms explicit and implicit externalities.

Another area to which our models can be applied is the study of conventions and their evolution. Young [33] models this as a coordination problem. He discusses at length the example of choosing which side of the road to drive on. This can be seen as a non-cooperative game with two Nash equilibria - all drive on the left or all on the right. He gives examples of tipping from one equilibrium to another. Most of his tipping examples are stochastic and dynamic, but one example that he looks at is within our framework: this is an example of multiple linked countries choosing the side of the road to drive on. He shows that a small group of countries or regions can force everyone else to change if they change: in our terms, this group forms a tipping set. In the case of a convention evolving, the central point is of course once again mutual reinforcement: your adopting the convention makes it more attractive for me to do likewise.

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:[2] 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. Apart from Dixit's work, there have been no previous attempts to model these reinforcement phenomena and the associated tipping or cascading in game-theoretic terms. 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 work, though he makes it in the context a dynamic processes rather than Nash equilibria. Although Granovetter does not

[2] The literature on supermodularity goes back to the 1970s, dating to the work of Donald Topkis [28] [29], although it and the literature on tipping have evolved quite separately. The idea of supermodularity was developed further by Xavier Vives [30] and Paul Milgrom and John Roberts [23]: since then it has been used widely in the literatures on game theory and comparative statics. Jeremy Bulow et al. [5] introduced the related idea of strategic complementarity. Supermodularity allows us to identify a class of games for which rather general comparative static results are available, and builds on the idea of increasing differences, which means that the return to a move by one agent can be increased by actions by other agents.

talk about tipping, the idea is implicit in his models of how thresholds interact, and he provides examples of cascades. We also show that if there are two (or more) equilibria, one of which Pareto dominates the other, 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 problem. 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 9/11/01, 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. (For the national security applications see our papers [14] [18], and for computer network applications see Michael Kearns [17].) 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. In trying to understand why tipping occurs we inevitably are drawn to increasing differences, as this is a template of a particular type of interdependence that is well-understood. This interdependence introduces an element of social reinforcement to investments in security.

1.1 Examples

Before we introduce the formal model and present our results, we work with examples of games exhibiting tipping, cascading and entrapment, to provide a perspective on these phenomena. To start with, 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

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

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 : it is more attractive, the more people choose it. In game theoretic terms, choosing 1 rather than 0 shows increasing differences. So this very simple example captures some of the features of the examples discussed above. Clearly $\{0, 0, ...0\}$ is a Nash equilibrium. Likewise $\{1, 1, ..1\}$ is also a Nash equilibrium, which Pareto dominates. Any agent can tip the equilibrium of zeros to that of ones. Thus 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 $\{1, 1, ..1\}$ than at $\{0, 0, ...0\}$, 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 loose 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 players, the trivial tipping coalition. In this case the equilibrium of ones seems stable and that of zeros seems unstable, as the former can be tipped by the action of any agent and the latter requires the action of many. We can easily modify this example to be more like the Schelling tipping examples:

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

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 correpondingly less stable.

Now consider a more complex example. The payoffs are

$$\begin{aligned} u_i(s_i, s_{-i}) &= 0.91i \text{ if } s_i = 0 \\ &= \#(1) \text{ if } s_i = 1 \end{aligned}$$

Again there is social reinforcement in the choice of strategy 1, but now agents are heterogeneous with respect to the payoff to 0. We again assume that $N = 10$. We again have equilibria of all zeros and all ones. In this case agent $N = 10$ can tip the equilibrium of zeros. If 10 changes from 0 to 1 then the payoff to choosing 1 for any other agent is now 1. As $1 > .91$, 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 ones. Agent 10 starts a cascade. 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. If all are choosing 1 and then 10 changes, the payoff to 9 from choosing 1 is 8 and the payoff to choosing 0 is 8.19. The payoff to zero is greater. Now there are two agents choosing zero so for agent 8 the payoff to 1 is 7 and to zero is 7.28, and the payoff to zero is greater. Again the change by 10 initiates a cascade from one equilibrium to another.

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 10 are worse off at the equilibrium of zeros than at that of ones (10 is better off), so when 10 tips an equilibrium of ones to one of zeros, he is making everyone else worse off - even though their best responses are now to choose zero. This illustrates the issue of entrapment discussed in the Dixit paper: in this example, agent 10 can entrap all others at the equilibrium of all zeros. Granovetter's discussion of criminal behavior also seems to fit this framework: talking about the behavior of delinquent boys, he states (page 1435) 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."' So illegal moves by those with low thresholds "entrap" others into following suit even and make them worse off.

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

It was a stochastic version of the second example that lead to our interest in these issues: here it is:

$$\begin{aligned} u_i(0) &= 0.91i \\ u_i(1) &= 10\frac{\#(1)}{N-1} + 0.5\left(1 - \frac{\#(1)}{N-1}\right) \end{aligned}$$

So the outcome of strategy 0 is certain, whereas that of strategy 1 is either 10 or 0.5 with probabilities $\frac{\#(1)}{N-1}$ or $\left(1 - \frac{\#(1)}{N-1}\right)$. 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 who also invest, $\#(1)$ (Heal and Kunreuther [14]). This example has most of the properties of the second (deterministic) example discussed above.

These examples illustrate tipping, cascading and entrapment, showing that they are closely related and arise in very simple and intuitive games when there is social reinforcement of a strategy choice. Next we define these effects in a more general context and relate them to known properties such as increasing differences and strategic complementarity. But first we make a comment on thresholds.

1.2 Thresholds

The concept of a threshold used in the sociological literature (Granovetter [12], Watts [32]) 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

$$\begin{aligned} u_i(0) &= a_i \\ u_i(1) &= \#(1)^{1/2} \end{aligned}$$

So there are social reinforcement effects in the choice of alternative 1, and these display diminishing returns. Clearly she will choose alternative 1 if $\#(1) > a_1^2$ 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 and Watts. 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.

2 Tipping and increasing differences

Consider a game with N players $i \in \{1, 2, ..., N\}$, each choosing a strategy s_i from the discrete set $\{0, 1\}$ and having a utility function $u_i : \{0, 1\}^N \to R$ which 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 ordering on R^N.[3]

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 S of all strategy choices and S_{-i} denotes the vector of choices of all agents other than i, then $\exists\ \epsilon > 0$ such that if $S'_{-i} > S_{-i}$ then

$$u_i\left(1_i, S'_{-i}\right) - u_i\left(0_i, S'_{-i}\right) \geq \epsilon + u_i\left(1_i, S_{-i}\right) - u_i\left(0_i, S_{-i}\right) \qquad (1)$$

This implies that the payoff to agent i from changing from 0 to 1 increases by at least ϵ if another agent changes from 0 to 1. 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, $\{0, 0, .., 0\}$ and $\{1, 1, .., 1\}$. A policy-maker will naturally be interested in ruling out the inefficient equilibrium and ensuring an efficient outcome, as in a coordination problem (Vincent Crawford [6], Walter Heller et al. [15]). 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 by changing their strategies shift the equilibrium from one extreme to the other. 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 by changing strategy shift the equilibrium from $\{0, 0, ...0\}$ to $\{1, 1, ...1\}$. 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 as its only equilibrium $\{1, 1, ..., 1\}$ then we say that T is a tipping set or $T-set$. The key point here is that when agents in T choose

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

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. 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, can force the system to the efficient outcome. If T is a small subset of N then this can be an important policy tool.

Below 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 others 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 measure of the degree to which they 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 an 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 non-trivial tipping set that tips the dominated to the undominated equilibrium.*

It is easy to see intuitively 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. We therefore need enough agents to change to bring the most negative payoffs above zero, and for this to be possible we need enough agents. What number of agents is 'large enough' depends on the parameter ϵ in the definition (1) of increasing differences. The larger is ϵ, the smaller is the critical number. To be precise, we show in the proof of proposition 1 that

$$k > 1 + \frac{Max_i\{u_i\left(0^{N-1}, 1_i\right) - u_i\left(0^{N-1}, 0_i\right)\}}{\epsilon}$$

so that the number of agents has to exceed one plus the ratio of the maximum payoff to a change 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: they 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 a loose and 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.

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

$$\begin{aligned} \Delta_{ij}(0) &= \left[u_j\left(0^{N-2}, 1_i, 1_j\right) - u_j\left(0^{N-2}, 1_i, 0_j\right)\right] - \\ &\quad \left[u_j\left(0^{N-2}, 0_i, 1_j\right) - u_j, \left(0^{N-2}, 0_i, 0_j\right)\right] \end{aligned} \tag{2}$$

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 $\Delta_{ij}(s_{-i-j})$ is independent of s_{-i-j} and so 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. These two conditions of symmetry and independence taken together we call assumption $A1$.

$$\Delta_{ij}(0) = \Delta_{ik}(0) = \Delta_i(0) = \Delta_i \tag{A1}$$

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 $\Delta_1 \geq \Delta_2 \geq .... \geq \Delta_N$. An agent's ability to tip the inefficient equilibrium is measured by its Δ, and we show below that a minimal $T - set$ consists of agents with the greatest Δs.

Proposition 3 *Given A1, 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.

The terms Δ_i are measures of the degree of increasing difference, and assumption $A1$ places a structure on these so that they are symmetric across agents. This structure is necessary for the simplicity of our arguments but not for the basic intuition that increasing differences contribute to tipping, as Proposition 1 shows. Within the structure defined by $A1$ we can say that increasing differences being sufficiently large in the sense of (9) 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 [14].

It is possible to establish results like Proposition 3, but more complex ones, with weaker assumptions than $A1$. Suppose for example that we drop the independence assumption, namely that $\Delta_{ij}(s_{-i-j})$ is independent of s_{-i-j}. In this case in stating and deriving a proposition analogous to Proposition 2 we need to reorder the agents by the size of Δ_i after each selection of a member of the tipping set, as the change of strategy from 0 to 1 by one agent can alter the ranking of the agents still choosing 0 by their $\Delta_i s$. In forming the tipping set at each stage we add the remaining agent whose Δ_i is greatest given the strategies now in place by all other agents, and this gives a more general but less parsimonious version of Proposition 3. If we drop the symmetry condition we are back with the general case of Proposition 1.

2.1 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. Avinash Dixit models this well and we follow his framework [8]. 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 where:

- 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 agent k. The strategy tuple in which agents 1 through k choose 1 and all others choose 0 is a Nash equilibrium.

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 number one to change strategy.

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

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

for all $j\leq k$.

Cascading, like tipping, depends on a game exhibiting "enough increasing difference." A numerical example of cascading from the inefficient to the efficient equilibrium of a supermodular game is given in [14].

3 Applications

3.1 Schelling's work

Schelling [26] 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 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. 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 likewise the number of others who do this, $\#(M)$. 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.

3.2 Computer security

As shown in Kunreuther and Heal [18] 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 itself against a possible virus, when it knows that it may be infected from other agents. The following example adapted from Kearns [17] illustrates this problem. Imagine the user population of a large organization in which each individual has a desktop computer with its own local software and memory, but in which parties also maintain important data files or documents on a shared disk drive accessible to the entire organization.

From the perspective of the organization, the primary security concern is that an intrusion (whether by a piece of malicious software or a human hacker) might erase the contents of this shared hard drive. Each user's desktop computer and its contents—including E-mail, downloaded programs or files, and so on—is a potential point of entry for such intrusion. Each user must implicitly decide about many aspects of their individual security practices: how often they change their password (and how secure those passwords are against dictionary and other common attacks), whether they enable encryption in their web and E-mail communications, their care in not downloading suspicious files and programs, their anti-virus software maintenance, and many other features, each of which takes time and hence is costly. The vulnerability of the shared hard drive is determined by the collective behavior along these dimensions. Hence if an agent invests in rigorous security precautions, her investments can be compromised by a failure to do likewise on the part of just one other. As it may be the case that some other agents do not store valuable data on the shared drive, their incentives to adopt best-practice security measures may be small. So we again have an interdependent security problem, with increasing differences in the associated game and the possibility of tipping from an equilibrium where none take security seriously to one where all do. In fact the formulae derived above for airline security apply with minor modifications to the computer security case.

4 Conclusions

Social reinforcement of choices is widespread, and indeed 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 non-trivial tipping sets, and have characterized them for a class of symmetric cases.

Tipping requires an initial mover or group of movers who begin the process. Sometimes it may be in their interest to do so: for example, in our first example the equilibrium with all ones is Pareto superior to that of all zeros, so that any agent can see that she can tip the system from zeros to ones and that she will gain from doing so. In other cases agents may need an incentive from outside the game(a subsidy or a penalty) to change strategy and tip the equilibrium: this is the case for some of the security examples discussed by Heal and Kunreuther [14]. Regulators can use the existence of a tipping set as a way of coordinating on a socially preferable equilibrium. They only have to persuade the members of the tipping set to change, rather than persuading everyone. Social reinforcement can also lead to entrapment, as discussed by Dixit and illustrated by our second example, where agent 10 can tip from the equilibrium of ones to that of zeros, at which everyone except her is worse off. An interesting point about cascades, and tipping when the tipping set is small, is that it provides a clear rationale for the oft-discussed policy of encouraging early adoption of a desirable behavior pattern. If the conditions for cascading or tipping are met, then subsidies to a set of early adopters could change the overall equilibrium provided that they form a tipping set or the first steps of a cascade.

5 Proofs

5.1 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 where 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

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

and

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

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

$$\Delta_{ij}\left(s_{-i-j}\right) = \Delta_j\left(i=1, s_{-i-j}\right) - \Delta_j\left(i=0, s_{-i-j}\right) \geq 0 \tag{3}$$

That is,

$$\begin{aligned}\Delta_{ij}\left(s_{-i-j}\right) &= \left[u_j\left(s_{-i-j}, 1_i, 1_j\right) - u_j\left(s_{-i-j}, 1_i, 0_j\right)\right] - \\ &\quad \left[u_j\left(s_{-i-j}, 0_i, 1_j\right) - u_j\left(s_{-i-j}, 0_i, 0_j\right)\right]\end{aligned} \tag{4}$$

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 equations (3) and (4) 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:

$$\begin{aligned}\Delta_{ij}(0) &= \left[u_j\left(0^{N-2}, 1_i, 1_j\right) - u_j\left(0^{N-2}, 1_i, 0_j\right)\right] - \\ &\quad \left[u_j\left(0^{N-2}, 0_i, 1_j\right) - u_j, \left(0^{N-2}, 0_i, 0_j\right)\right]\end{aligned} \tag{5}$$

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 zero to one, and which hold by the increasing differences (1) property.

$$u_i\left(0^{N-1}, 1_i\right) - u_i\left(0^{N-1}, 0_i\right) + \epsilon < u_i\left(0^{N-2}, 1_1, 1_i\right) - u_i\left(0^{N-2}, 1_1, 0_i\right) \tag{6}$$

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

$$\begin{aligned}&u_i\left(1_1, 1_2, .., 1_{N-2}, 1_i\right) - u_i\left(1_1, 1_2, .., 1_{N-2}, 0_i\right) + \epsilon < \\ &\quad u_i\left(1_1, 1_2, .., 1_{N-1}, 1_i\right) - u_i\left(1_1, 1_2, .., 1_{N-1}, 0_i\right)\end{aligned} \tag{7}$$

If we take the first inequality

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

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

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

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 (6) we have that

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

Note that

$$u_i\left(0^{N-1}, 1_i\right) - u_i\left(0^{N-1}, 0_i\right) < 0$$

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

$$u_i\left(1_1, 1_2, ..., 1_{N-1}, 1_i\right) - u_i\left(1_1, 1_2, ..., 1_{N-1}, 0_i\right) > 0$$

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 $(k-1)\epsilon - u_i\left(0^{N-1}, 1_i\right) + u_i\left(0^{N-1}, 0_i\right) > 0$ and the first k agents will then form a $T - set$. To be precise, we need k to satisfy

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

or

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

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 is 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.

5.2 Proof of Proposition 2

Let $\left\{0^t, 1^{N-t-1}, 1_k\right\}$ 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 (3) and (4) and (A1) we can write

$$u_j\left(0^{N-K-1}, 1^K, 1_j\right) - u_j\left(0^{N-K-1}, 1^K, 0\right) = u_j\left(0^{N-1}, 1\right) - u_j\left(0^{N-1}, 0\right) + \sum_{i=1}^{i=K-1} \Delta_i \quad (8)$$

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

If $F < N$ is a $T - set$ then for all $j > F$ we must have

$$u_j\left(0^{N-F-1}, 1^F, 1_j\right) - u_j\left(0^{N-F}, 1^F\right) \geq 0$$

By (8) above this is equivalent to

$$\sum_{i=1}^{i=F-1} \Delta_i \geq u_j\left(0^{N-1}, 0\right) - u_j\left(0^{N-1}, 1\right) \forall j > F \quad (9)$$

To construct a minimal $T - set$ we need to find the smallest number F for which (9) holds. Clearly we get this by ranking agents by the size of Δ_i and 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.

5.3 Proof of Proposition 3

For a change by agent 1 to change agent $2's$ strategy we need that

$$u_2(1,1,0..0) - u_2(1,0..0) = u_2(0,1,0..0) - u_2(0,..0) + \Delta_1 > 0$$

or

$$\Delta_1 > u_2(0,..0) - u_2(0,1,0..0)$$

Similarly for a change by agent 2 to change $3\prime s$ strategy

$$u_3(1,1,1,0..0) - u_3(1,1,0..0) = u_3(1,0,1,0..0) - u_3(1,0,..0) + \Delta_2 > 0$$

or

$$\Delta_2 > u_3(1,0,..0) - u_3(1,0,1,0..0)$$

The proposition follows by repeating this argument.

References

[1] Arthur, W.B. 1989 "Competing Technologies, Increasing Returns and Lock-in by Historical Events," *Economic Journal*, 99, 106-131.

[2] Barberis, N. and A. Schleifer, 2003, "Style Investing," *Journal of Financial Economics*, 68:161-199.

[3] Brock, W.A. "Tipping Points, Abrupt Opinion Changes and Punctuated Policy Changes," Department of Economics, University of Madison Wisconsin, May 2004, available at http://www.ssc.wisc.edu/econ/archive/wp2003-28.pdf

[4] Brock, William A. and Durlauf, Steven N., "Discrete Choice with Social Interactions I: Theory" (October 1995). NBER Working Paper No. W5291. Available at SSRN: http://ssrn.com/abstract=225355

[5] Bulow J. I., J. D. Geanakoplos and P. D. Klemperer 1985. "Multimarket Oligopoly: Strategic Substitutes and Complements". *Journal of Political Economy*, vol. 93, no 313, 488-511.

[6] Crawford, V. and H. Haller. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games" *Econometrica* Vol 58: Issue 3 (May 1990) 571 - 595.

[7] DeLong J.B., A. Schleifer, L. Summers and R. Waldmann, 1990, "Positive feedback investment strategies and destabilizing rational speculation," *Journal of Finance*, 45:375-395.9

[8] Dixit A. K. 2002. "Clubs with entrapment," *American Economic Review* Vol. 93 No. 5 pp. 1824-1829.

[9] Easterly, W. 2005 "Empirics of strategic interdependence: the case of the racial tipping point" New York University Development Research Institute Working Paper, October.

[10] Economides, N. "The Economics of Networks," undated, available at http://ssrn.com/abstract=1719

[11] Gladwell M. 2000. *The Tipping Point* Little Brown and Co.

[12] Granovetter, M. "Threshold Models of Collective Behavior," *The American Journal of Sociology*, Vol. 83, No. 6, (May, 1978), pp. 1420-1443.

[13] Heal, G.M. 1994. "The Formation of International Environmental Agreements" in C. Carraro (ed.) *Trade Innovation and the Environment,* pp. 301-322, Boston, Kluwer.

[14] Heal G.M. and H. Kunreuther.(2005). "IDS Models of Airline Security" *Journal of Conflict Resolution* 41:201-17.

[15] Heller, Walter 1986 "Coordination Failure in Complete Markets with Applications to Effective Demand". In *Equilibrium Analysis: Essays in Honor of Kenneth J. Arrow Vol II*, ed. W.P.Heller, R.M. Starr and D.A. Starrett, Cambridge University Press., 1986.

[16] Hong H.G., J.D. Kubik and J.C. Stein, 2001, "Social Interaction and Stock Market Participation," Stanford Business School Working Paper.

[17] Kearns, Michael 2005. "Economics, Computer Science, and Policy". *Issues in Science and Technology,* Winter pages 37-47.

[18] Kunreuther H. and G.M. Heal 2003. "Interdependent Security." *Journal of Risk and Uncertainty*, Special Issue on Terrorist Risks, Vol. 26 No. 2/3 (March/May): 231-249.

[19] Kunreuther H and G.M. Heal 2005 "Interdependencies within an Organization" in B. Hutter and M. Powers (ed.) *Organizational Encounters with Risk* (Cambridge University Press)

[20] Leibenstein H. *"Bandwagon, Snob and Veblen Effects in the Theory of Consumers' Demand,"* The Quarterly Journal of Economics, Vol. 64, No. 2 (May, 1950), pp. 183-207.

[21] Lopez-Pintado, D. and D.J. Watts (2006) "Social Influence, Binary Decisions and Collective Dynamics." Working Paper, Institute for Social and Economic Research and Policy, Columbia University.

[22] Manski, C.F. "Economic Analysis of Social Interactions," *The Journal of Economic Perspectives*, Vol. 14, No. 3, (Summer, 2000), pp. 115-136.

[23] Milgrom P. and J. Roberts. 1990. "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities" *Econometrica* 58: 1255-77.

[24] Rae, J. The Sociological Theory of Capital, London, The Macmillan Company, 1905.

[25] Schelling, T. 1971. "Dynamic Models of Segregation" *Journal of Mathematical Sociology* 1:143-86.

[26] Schelling, T. 1978. *Micromotives and Macrobehavior.* New York: Norton

[27] Shapiro, C. and H. Varian, (1999) Information Rules: A Strategic Guide to the Network Economy, Harvard Business School Press.

[28] Topkis, D. 1978 "Minimizing a Submodular Function on a Lattice" *Operations Research* 26: 305-21.

[29] Topkis, D. 1979 "Equilibrium Points in Nonzero-Sum n-Person Submodular Games" *Siam Journal of Control and Optimization* 17:773-87.

[30] Vives, X.. *Journal of Economic Literature* Vol.. XLIII (June 2005)

[31] Wagner, Ullrich (2007). "Estimating Strategic Complementarities in a Dynamic Game of Timing: The Case of the Montreal Protocol." Draft, October 2007, Columbia Earth Institute.

[32] Watts, D. J. "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences*, April 30 2002, Vol. 99, No. 9, pp. 5766-5771.

[33] Young, H.P. (1996) "The Economics of Convention," *Journal of Economic Perspectives*, Vol. 10 No. 2, Spring 1996, pp.105-122.