Journal of Economic Behavior and Organization 9 (1988) 325–343. North-Holland

ON THE SLIPPERY SLOPE

Conformance vs Defection in a Multi-Party Prisoners' Dilemma

K.K. FUNG*

Memphis State University, Memphis, TN 38152, USA

Received August 1986, final version received July 1987

A slippery-slope slide occurs when minor deviations from an unstable superior solution of total conformance result in an inferior solution of massive defection. Using a multi-party prisoners' dilemma (PD) game in binary choices, this paper examines some possibilities between total conformance and total defection. Specifically, it shows how minor deviations from total conformance can be contained and how total defection can be reversed. The containment of deviations is, however, greatly complicated by technological changes which may precipitate a free slide by reducing the transaction costs of defection, or which may transform a slippery slope into an invisible hand.

1. Introduction

Under a system of universal military draft for able-bodied male youths, it is individually advantageous to get exempted from or to dodge the draft as long as most eligible male youths conform to the system. However, as the number of exemptions and dodgers increases, the chance of being drafted for the remaining conformers noticeably increases. This, in addition to envy, will induce further defections from the system. Thus, minor defections from an unstable but superior solution of total conformance culminate in a stable but inferior solution of total defection. This situation of an unstable superior solution with inherently unstoppable defections is commonly characterized as the slippery slope. In everyday language, whenever we are talking about setting a bad precedent, we are referring to a slippery slope.

This individually advantageous but jointly destructive process of social interactions provides a sharp contrast to the classic laissez-faire model of the invisible hand [Smith (1776)] in which pursuit of self interests promotes collective welfare without any prior organizational design.

The existence of the slippery slope presents at least two challenging research problems. First, the problem of coexistence and interactions between

*Dale Bails and Karen Pickerill commented on the first draft of this paper. Also, they and Julia Heath helped with stylistic revision of the final version.

the invisible hand and the slippery slope. The optimal solution under an invisible hand requires maximum freedom to pursue self interests while the optimal solution on a slippery slope requires voluntary or enforced conformance to a self-denying norm. Thus, the problem boils down to the correct identification of the two situations. Second, the problem of controlling slides down the slippery slope. Can a limited slide be collectively optimal? If so, how can a limited slide be contained? Under what conditions can a free slide be reversed?

These two problems will be analyzed with the help of the diagrammatics of a multi-party prisoners' dilemma (PD) game as developed by Schelling (1978, ch. 7). Specifically, a slippery slope will be seen as a multi-party PD game where the payoff from one action always exceeds the payoff from another action although the pursuit of the higher-paying action leads to declining payoffs for both actions. Using the same diagrammatics, but varying the directions, slopes, and positions of the payoff curves of binary choices, the interactions between the invisible hand and the slippery slope, and the problem of controlling slides down the slippery slope will be illustrated.

2. The multi-party prisoners' dilemma and the slippery slope

The central characteristic of a slippery slope is the tendency for minor defection from an unstable optimal solution to develop into massive defection. This is also the central characteristic of a prisoners' dilemma (PD) game.

The conventional two-party two-strategy PD game has been useful in modelling the instability between total conformance and total defection. However, in order to analyze points in between these two extremes, we need to have a multi-party PD game. We will illustrate the difference between a two-party and a multi-party PD game in fig. 1.

Fig. 1 depicts a multi-party two-strategy PD game. The horizontal axis from left to right represents the number of people (n in all) choosing strategy or action A, and from right to left the number of people choosing non-A (i.e., B). In other words, at any point on the horizontal axis, the number of people choosing A plus the number of people choosing B is equal to n. For example, at the right end of the horizontal axis, the number of people choosing A is n, and the number of people choosing B is zero. At the left end of the horizontal axis, the numbers of people choosing A and B are reversed. The vertical axis represents per capita payoffs *gross of transaction costs* for A and B. The per capita payoff curves are linear and non-intersecting as are commonly assumed in a two-party PD game. For simplicity, all parties choosing A or B are assumed to have equal per capita payoffs within each action group.

A conventional two-party-two-strategy PD game is concerned only with

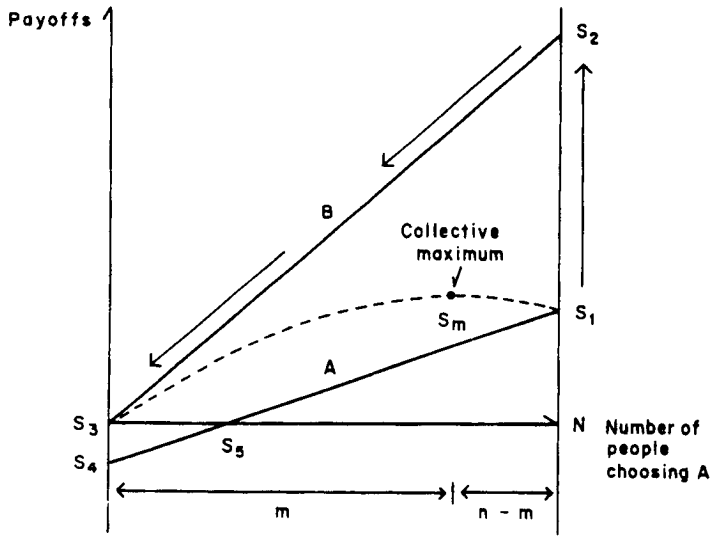


Figure 1

SLIPPERY SLOPE TO COMPLETE DEFECTION

the four end points of the two payoff curves (i.e., S_1, S_2, S_3, and S_4) because only two persons are involved. Specifically, S_1 is the Pareto optimal[1] solution with both parties choosing A. But S_1 is an unstable solution because there is constant temptation to defect from A. Departure from S_1 will improve the payoff of the defector from A at the expense of the conformer to A. The advantage of defection from A over conformance to A is indicated by the gap between S_2 and S_1. Since defection from A decreases the payoff of the conformer to A, defection of one party from A will be counteracted with defection from the other party. Therefore, departure from S_1 will inevitably lead to total defection at S_3. Although S_3 is Pareto inferior[2] to S_1, it is a stable solution. Defection from B will hurt only the defector. The disadvantage of defection from B is measured by the gap between S_3 and S_4.

However, a multi-party two-strategy PD game as depicted in fig. 1 allows us to analyze the points between S_1 and S_3 along the payoff curves for A and B. Specifically, we can show that minor defection from A, if containable,

[1]A certain state is said to be Pareto superior to another if some parties can gain without other parties' losing from the move. A state is, therefore, Pareto optimal if no more Pareto superior moves exist.

[2]A certain state is said to be Pareto inferior to another if some parties lose without other parties gaining from the move.

may lead to a higher collective optimum[3] at S_m than total conformance to A at S_1 [Schelling (1978, ch. 7)]. S_m is, of course, not Pareto superior to S_1 because the conformers to A (m in number) are worse off than at S_1. But the gain obtained by those who defect from A ($n-m$ in number) is potentially high enough to more than compensate for the loss suffered by those who still conform to A. Therefore, concern for justice and collective welfare may well justify minor defections from total conformance to A.

A multi-party PD game also makes it possible to model a situation where total defection is not an inevitable result of minor defection. In fact, the assumption of an unstoppable free slide is valid only if the payoff curve for B is always higher than that for A, as depicted in fig. 1. This means that B is always the unconditional own preference and A is always the unconditional other preference. In other words, one always prefers B for oneself, but prefers the other parties to choose A. We can certainly think of examples where the unconditional own preference always conflicts with the unconditional other preference. For example, in the tragedy of the commons, everyone prefers

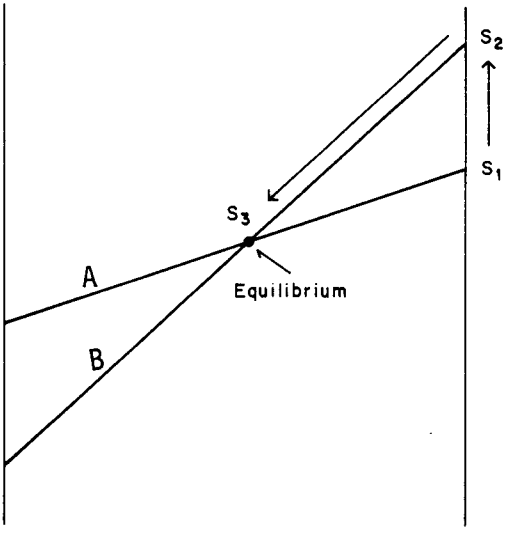


Figure 2

SELF-ARRESTING SLIPPERY SLOPE

[3]A collective optimum represents the highest sum total of payoffs. Where a Pareto optimum does not coincide with a collective optimum, a move from the former to the latter will not be Pareto efficient. That is, some parties will gain while others lose even though the collective sum of payoffs is higher. In fig. 1, S_m represents the sum of the per capita payoff from A at m weighted by m/n and the per capita payoff from B at m weighted by $(n-m)/n$. Other points of the dotted line are computed similarly.

grazing more livestocks of one's own while wishing others to exercise self restraint [Hardin (1968)].

But if the payoff curves intersect, the end of a slippery slope need not be a total defection from A. It may well be a stable non-corner solution as depicted by S_3 in fig. 2. In fig. 2, the slippery slope S_1–S_2–S_3 falls short of total defection from A because beyond S_3 simple pursuit of narrow self interests through an invisible-hand process would guarantee a return to the Pareto superior S_3 (see section 6 below). There are many examples of self-arresting slippery slopes. For example, in a general assembly, although some level of attendance is worthwhile, there is no need for everyone to attend. Minor deviations from full attendance do not generally lead to total absence.

3. Transaction costs

The above discussion of the slippery slope and the invisible hand has concentrated on the payoffs of alternative actions *gross of transaction costs.* When transaction costs are taken into account, the net advantage between alternative actions may be completely offset. For our present purposes, it is helpful to separate transaction costs into migration costs, expected loss, and application costs.

Migration costs are those involved in defecting from A to B. They are high if migration: (a) necessitates complicity from other parties; (b) requires commitment of a large amount of private resources; (c) is physically difficult to accomplish; and/or (d) arouses overwhelming shame or guilt in the potential defector. This sense of shame or guilt felt by the potential defector is the by-product of an established norm that stabilizes an unstable Pareto optimal solution to a slippery slope. These norms have been referred to as PD norms [Ullmann-Margalit (1977, ch. 2)].

Expected loss is the penalty upon detection discounted by the probability of detection per unit of enforcement effort. Expected loss tends to increase with migration costs. When migration costs are high, detection is easier because there are fewer violations. Thus given the penalty, easier detection results in higher expected loss. Where migration costs are increased by the imposition of PD norms, detection is made even easier by the moral outrage of conformers who are more likely to report violations. Given the migration costs, the probability of detection also varies with the visibility of the defectors. For example, detection of computer software piracy is much easier against corporate users than against home users even though the cost of pirating (a migration cost) is the same for both groups. The many more violators among home users and their lower visibility make it more difficult to catch a given number of violations. Also, disgruntled corporate employees are more likely to be whistle blowers than family members of home pirates.

Furthermore, a higher penalty is more likely to be imposed on corporate pirates because of the deep pockets of their employers. Thus, easier detection and higher penalty result in higher expected loss for the corporate pirates. When expected loss is high because the probability of detection is high, we can also say that the costs of enforcing the PD norm is low since enforcement costs are inversely related to the probability of detection.[4]

When the benefit of defection does not come or is not secure until the defection is formally exempted, potential defectors must prove to the proper authorities that they satisfy the criteria for exemption. Or when prior approval is not possible, the defectors must be able to defend their defections when later challenged. The expenses involved in obtaining exemptions can be called application costs. For example, refugees from Central America must prove that they are political refugees rather than economic refugees before they are granted permanent residence in the U.S. Thus, although migration costs may be low, application costs are high.

4. Total conformance

Because transaction costs are barriers to defection, they can be adjusted to achieve a desired level of conformance. If the collective maximum coincides with the Pareto optimal solution at the point of total conformance to the PD norm, transaction costs should then be set high enough to offset any net advantage of defection over conformance.

In diagrammatic terms, the imposition of transaction costs to achieve total conformance amounts to a downward shift of the B payoff curve to completely eliminate the advantage of B over A at the point of total conformance to A. In fig. 3, the B payoff curve is transformed into $B-C_1$, where C_1 is the net advantage of B over A, gross of transaction costs. This downward shift of B converts A into a dominant action. A stable Pareto optimal solution with total conformance to A is thus guaranteed.

On most slippery slopes, however, migration costs in the absence of moral restraint tend to be low. To be successful in containing slides, a PD norm must, therefore, be able to instill a sense of shame or guilt into the potential defector. For example, the code of silence among members of the Mafia and sanctions against military desertions both carry tremendous moral weights with potential defectors. Members of these groups are extremely reluctant to break the respective PD norms because a matter of honor is involved. The close-knit nature of the groups also facilitates detection. And to further discourage defection in case of wavering, the penalty is extremely high.

Total conformance is easier to enforce in societies where the value system is binary. That is, things are either black or white with no grey areas in

[4]Karen Pickerill helped me clarify this point.

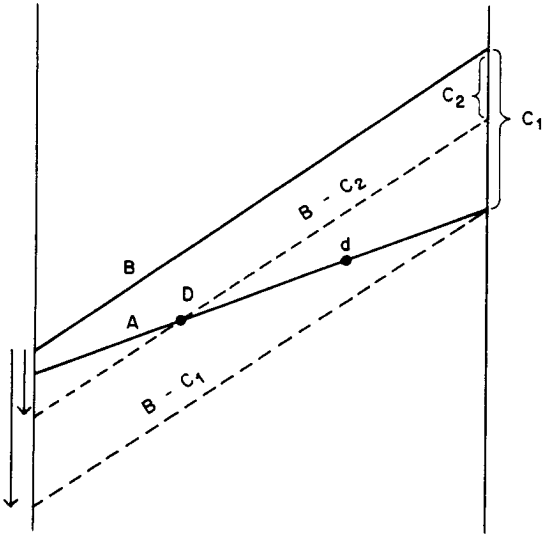


Figure 3

TOTAL CONFORMANCE VS SEGMENTATION

between. When people are not trained to see any grey areas, deviations from the norm can thus be easily detected. Furthermore, a fall from the desirable category to the undesirable category evokes no sympathy. The universal condemnation accompanying such a fall facilitates the enforcement of PD norms. In such societies, not only will deviations from established norms be severely punished, but any emergence of grey areas between the binary categories due to natural mutations or deliberate innovations will also be ruthlessly suppressed.

5. Segmentation

When partial conformance produces a higher collective payoff than total conformance, exemptions from total conformance should be permitted to exploit collectively superior opportunities. The challenge of partial conformance is to contain defections within some allowable limits. This is most easily achieved when the benefit from exemptions cannot be secured without formal approval from the proper authorities. Since partial conformance involves segmenting the target population, we can call these methods of containment 'segmentations'.

Segmentation can be achieved by creating a sub-category of an appro-

priate size. The characteristic of this sub-category should ideally be related to deserts (see section 9). In other words, the criterion that is the highest on the priority list of deserts should be chosen to determine who is eligible for exemption. But if the number of those who deserve exemption from the norm exceeds the optimal number for the collective maximum, then secondary constraints, which may have little to do with deserts, must be added to obtain the allowable number of deviations. Or the next criterion on the deserts list can be chosen instead if it produces the optimal number of containable exemptions. Whatever the criteria chosen, they should be tight enough to prevent artificial migration from unqualified aspirants. It is easy to imagine physiological criteria for segmentation. For example, young men with flat feet may be exempted from military draft without any fear of unqualified aspirants' trying to artificially flatten their normal feet. Segmentation through inborn physical criteria has the obvious advantage of high migration costs. If the physical criteria are also highly visible, they have the additional advantage of low enforcement costs. In diagrammatic terms, segmentation amounts to a downward shift of the B payoff curve to partially eliminate the net advantage of B over A at the point of total conformance to A (see $B-C_2$ in fig. 3).

Although visible boundaries often serve to contain defections, minor defections are just as often contained by limiting the application of potentially precedential information. There are at least two ways this may be implemented. First, the existence of exemptions can be concealed. Ignorance on the part of potential candidates automatically reduces the number of applicants. Second, the criteria on which exemptions are based can be concealed. This allows the criteria to be adjusted to suit individual situations without fear of setting any precedents. By treating each exemption in an ad hoc manner, artificial migration is prevented because no meaningful efforts can be made to qualify for unknown criteria. These ad hoc segmentations are effective only in cases where the benefit of migration cannot be secured without formal approval from the proper authorities. In effect, ad hoc segmentations increase application costs, but they work only if the number of sources from which exceptions can be obtained is limited. If the number of sources is not controlled, these sources may independently grant limited exemptions that collectively total in excess of the global limit of permissible deviations.

The jury system in court trials is an example of ad hoc segmentation. The jury does not have to announce its decision criteria. Furthermore, jury decisions have no precedential values until they are challenged and upheld in the appellate courts. Because a jury is an *a*responsible agency with no permanent tenure, it also need not concern itself with the danger of setting bad precedents. As a result, an otherwise effective method for ad hoc segmentation can easily be turned into an agency for initiating a free slide down the slippery slope when there is a consistent ideological bias among

the population from which the juries are selected [Calabresi and Bobbitt (1978, ch. 3)]. The large size and increasing frequency of punitive damages awarded by the jury in U.S. liability suits illustrate this pitfall (*Wall Street Journal*, April 21, 1986, p. 21).

A committee with limited tenure suffers from the same drawbacks as an agency for ad hoc segmentation. Studies have shown that committees tend to avoid issues of general principles in order to reach agreement on specific policies. While the absence of agreement on any general principles in the policy recommendations certainly limits the general applicability of these policies outside the specific situation considered by the committee, concern with only specific policy recommendations may induce incremental creep towards a sudden free slide down the slippery slope [Mendeloff (1985)]. Each incremental extension may seem harmless by itself, yet the sum of individual extensions may exceed what is collectively optimal.

In other words, ad hoc segmentation attempts to contain defection to an indefensible limit by allowing inconsistency among: (1) the criteria for exemptions, (2) the transaction costs of effecting the exemptions, and (3) the number of permissible exemptions. This inconsistency exists because precedential information is concealed. If the criteria for exemptions are openly announced, the transaction costs associated with the criteria would result in more exemptions. Similarly, if the number of permissible exemptions is openly announced, more stringent criteria for exemptions would be required at the existing level of transaction costs. In diagrammatic terms, ad hoc segmentation is intended to limit exemptions to points (such as d) that fall far short of the intersection between A and $B-C_2$ (see fig. 3). Since the intersection D represents the limit of exemptions that would result if the general criteria for exemptions are unambiguously and openly announced, ad hoc segmentation is inherently unstable.

Any achieved stability must, therefore, depend on manipulating the costs of applying the ad hoc decision to new cases. If each new application of an old decision to similar situations involves large application costs, very few applicants will come forth. Minor deviations can thus be easily contained. For example, if the application costs for punitive damages are increased by limiting punitive damages or forbidding contingent fee arrangements, the number of liability suits would be greatly decreased. However, if new extensions of applicability are inexpensive, minor deviations can easily precipitate a slide down the slippery slope to a logical resting point (i.e., from d to D on $B-C_2$ in fig. 3). The political tendency in the U.S. is to lower application costs to potential applicants for exemptions. In general, the burden of application costs has increasingly been shifted to parties that have the most resources to bear it. Contingent fees, joint and several liability, class action suits, and employee rights are examples of this trend. As a result, most ad hoc segmentations have been rendered ineffective in arresting free slides down the slippery slope.

6. The critical mass to reverse slippery-slope slides

In general, slippery-slope slides are not a serious problem unless massive or total defection results. Massive or total defection is possible only when B dominates A over most or all of the target population. Even if such dominance exists, segmentation usually ensures that massive or total defection does not happen. When total or massive defection does occur, the payoffs from A and B and/or the transaction costs must have been drastically altered. Such a drastic alteration of transaction costs and payoffs in turn stabilizes total defection.

The stability of total defection is usually attributed to the dominance of B over A at the Pareto inferior solution. But a more important consideration may be the size of the minimum critical effort required to reverse the slide. This minimum critical effort is usually far larger than the effort needed to stabilize total conformance under stable transaction costs. With near total conformance, enforcement costs are limited to catching only the few violators. With near total defection, a critical mass has to be induced to change their behavior.

The size of this critical mass depends on behavioral assumptions and the relative position and dominance of B and A. When B dominates A over the entire population as in fig. 1, a pseudo critical mass exists if people have become less envious and more *Pareto altruistic* after they have experienced the deprivation of total defection. People are not envious if they evaluate their own well-being solely on their absolute level of payoffs and not on their payoffs relative to other people. People are Pareto altruistic if they prefer a state where other people's absolute payoffs are higher as long as their absolute payoffs stay the same. Then S_3–S_5 in fig. 1 represents the minimum number of people that must dissociate from total defection in order for Pareto altruism to take effect. At S_5, S_3–S_5 of the people choosing A are at least as well off as at S_3 while $n-S_5$ of the people choosing B are better off than at S_3. But S_3–S_5 is not a true critical mass because it does not have any catalytic effect. That is, at S_5, pursuit of self interests alone still cannot induce a higher level of conformance to A. A higher level of conformance must again await another substantial dose of minimum critical efforts even if Pareto altruism still applies to the right of S_5. That means this level of conformance to A (S_5) is always at the brink of collapse to total defection. But just as defection from A at S_1 can be contained by imposing a tax on B, so dissociation from B at S_3 can be encouraged by granting a subsidy to A to achieve a pseudo critical mass. This subsidy could be in the form of reducing the organizing costs of creating such a mass. The size of this subsidy depends, of course, on the net advantage of B over A. The greater the advantage is, the larger the critical minimum subsidy must be.

A more interesting case of critical mass involves directional payoff

reversals. So far, we have assumed that the payoffs from B always exceed the payoffs from A regardless of the direction of migration. This means that the dominance of B over A is not affected by the number of people or the experience from choosing A and B. But if people consistently underestimate the payoffs from the action they have chosen and overestimate the payoffs from the action they have not chosen, and if people have a more favorable evaluation of the payoffs from the action they have abandoned after they have experienced the alternative action, there may be directional reversals in the payoffs from A and B. The fact that people do frequently romanticize untried alternatives [Fung (1979)] and regret their choices means that directional payoff reversals are quite common.

We can illustrate directional payoff reversals with the example of sexual chastity (A) and promiscuity (B). From a state of total chastity, the availability of effective birth control devices and antibiotics for sexually transmitted diseases may precipitate a migration to promiscuity which is sustained by a dominance of B over A. On the other hand, from a state of total promiscuity, a change in social values or the discovery of a hidden cost (such as AIDS) may induce a migration towards chastity which is sustained by a dominance of A over B. For example, as promiscuity becomes widespread, people may question the value of promiscuous physical intimacy and long for the emotional intimacy that may come from only a long-term monogamous relationship.

Although payoff reversals can explain the reversibility of slides, they in turn present problems of instability. If A's payoffs completely dominates B's payoffs as soon as total defection is reached, total defection would be unstable. In other words, total defection would be reversed as soon as it is achieved. This problem of extreme instability can be solved if the payoff reversal begins only some distance short of total defection. A payoff reversal at less than total defection can also explain why a critical mass must be achieved before slippery-slope slides can be reversed.

In diagrammatic terms (see fig. 4), from a state of total chastity at S_1, the payoff curve for A consists of the solid portion of the payoff curve A from S_1 to S_3, the dotted portion (A') from S_3 to S_4, and the solid portion of A from S_4 to S_6. From a state of total promiscuity at S_5, the payoff curve for A is the entire solid line labelled A. For simplicity, we assume the payoff curve for B is unchanged by the direction of migration. Before the new technology that reduces transaction costs is introduced, society may be at an unstable solution S_1. There is, however, no danger of massive defection as the slippery slope is self-arrested at S_3. The new technology, by lowering transaction costs, reverses the dominance of A over B between S_3 and S_4 and induces further defection from A at S_3. As S_4 is reached, the payoffs between A and B are reversed again, making A the dominant action between S_4 and S_3. However, the slide would not stop at S_4. Instead, the momentum of the slide

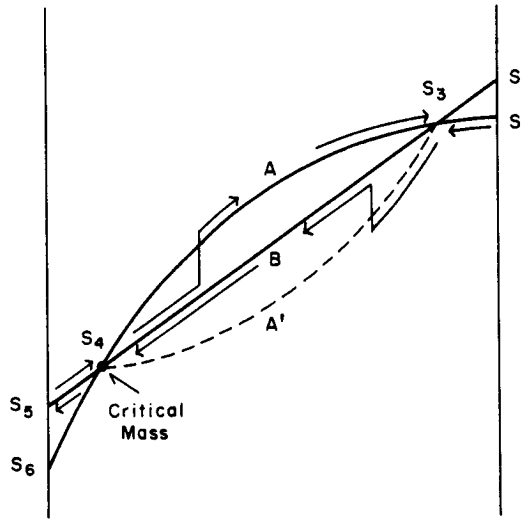


Figure 4

PAYOFF REVERSALS AND CRITICAL MASS

down B is reinforced by the continued dominance of B over A from S_4 to S_5.

A free slide beyond S_4 to S_5 even when the reversed payoffs justify a return movement to S_3 can be ascribed to a loss of critical mass. If at least S_5–S_4 number of people is firmly committed to A at S_4, a self-sustained conformance to A (between S_4 and S_3 on payoff curve A) through the invisible-hand process would follow.

Most discussions of slippery slopes assume the non-existence of a self-sustaining critical mass. That is, a free slide down the slippery slope must lead to total defection, and at the point of total defection, a return to conformance is impossible. But if a reversal of payoffs is typical before a free slide down the slippery slope is completed, then the problem facing a policy maker is how to preserve a critical mass for a comeback to conformance. At S_4 when A again becomes dominant between S_4 and S_3 after a slide down B, there is no net advantage for further defection from A. A little foresight and some minimal resources are all that is needed to preserve a critical mass for a return to S_3 via the invisible-hand process. If action is not taken until total defection from A is reached at S_5, more resources are required to rebuild a critical mass back to S_4. The minimum amount of subsidy needed is equal to the triangular area S_4–S_5–S_6. Many formal and informal seed grants are examples of subsidy to create a critical mass. If we drop the assumption of equal payoffs for members within each action group, we can rely on diversity to generate a critical mass. The more diverse a society is in its values, the

more likely it is that some segments still retain values which are different from those of the mainstream and needed to serve as a core for a critical mass. History has shown that there has never been any shortage of martyrs and exemplary figures. And one can think of many critical turning points in history that were affected by the catalytic acts of a few far-sighted individuals.

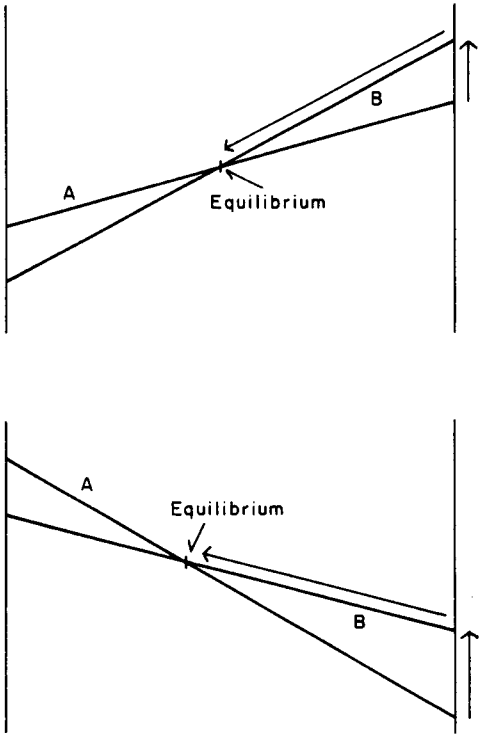
Although payoff reversals may reverse the slide down the slippery slope, there is no guarantee that the payoff reversals will always happen in time to be exploited. If a slide to S_4 has done irreparable damage to the system, a payoff reversal at S_4 will be too late to be useful. For example, a nuclear armaments race may precipitate a global holocaust at S_4 or before S_4, rendering a return to S_3 impossible.

7. False alarm?

One of the major contentions about slippery slopes is our inability to predict the direction of the underlying slope. This inability is made more serious by the short-term individual advantage of defection and the additional uncertainty introduced by technology. Under changing technology, the direction of the slope may well have been reversed. The same behavior that may have precipitated a free slide down the slippery slope under the old technology (fig. 5a) may instead pioneer an invisible-hand process under the new technology (fig. 5b).

Software piracy is a case in point. The migration costs of becoming a pirate have become so low that it is all but impossible to resist infringement of copyrights. The material costs consist of an inexpensive floppy disk which can be reused if the pirated product turns out to be unsuited to the pirate's needs. The enforcement costs are high because the act of piracy is not highly visible and the pirated product can be used in privacy. More importantly, the act of piracy is not clearly perceived as wrong because similar acts with respect to other private property are not perceived as wrong. In other words, computer technology has realized the potential of information as a public good because the material embodiment itself no longer presents a serious barrier to the dissemination of information.

Although piracy on the surface appears to reduce legitimate sale, the extent of reduction is highly debatable. First, not all pirated products are actually used. Copies may have been made simply for reference. Second, not all users of pirated products will buy their legitimate versions even if piracy can be stamped out. They may simply buy a cheaper legitimate product that fits their budgets. Third, sale of legitimate products to those who would not or cannot afford to use pirated products may actually increase through personal recommendations from pirates. Indeed, the most pirated products



are also the ones with the highest sales to legitimate users (*InfoWorld*, May 5, 1986, p. 11). Fourth, pirates can reduce the support costs of software houses because they are the most available support experts. Fifth, pirates can increase the sale of hardwares through their software demonstration. And higher sale of hardwares will inevitably spill over to higher sales of legitimate softwares.

The low cost of piracy for most non-corporate individual users means that it is practically impossible to enforce complete legitimacy for all users. The only question remaining is whether legitimate purchases alone can continue to compensate innovative products. If they can, then there is no serious incentive problem. Piracy (*B*) will both result in Pareto superior solution and stop short of total defection from legitimacy (fig. 5b). If they cannot, and free access to softwares is indeed potentially Pareto optimal, then software writers just have to be compensated by other means than private sale.

Even when a slippery slope is positively identified, there is no need to be alarmed at all defections. Failure to appreciate that defection on a slippery slope may stop far short of total defection even without artificially increasing transaction costs means that a lot of the concern for moral erosion may be

misplaced (see fig. 5a). For example, it is difficult to imagine that legalizing homosexuality will lead to total homosexuality and legalizing abortion will lead to total abortion. Thus, given a necessarily limited budget for law enforcement, it would be more productive to be concerned only with slippery slopes that are not easily self-arresting.

8. Searching for defensible segmentations

Unless migration costs and application costs are zero, most minor defections do not necessarily lead to a free slide. When minor defections persist or expand, it may not be because they cannot be stopped but because a consensus regarding a global limit is absent. This lack of consensus leads to indefensible segmentations in which the criteria for exemptions, the transaction costs of effecting exemptions, and the total number of permissible exemptions are not mutually consistent.

The absence of a global limit creates both opportunities and problems. In terms of opportunities, this absence means that various ad hoc solutions can be tried without having to adhere to any general principles on a global limit.[5] Some of these solutions may prove to be the basis for a later consensus on general principles.

But ad hoc solutions also create problems because of the limited context in which incremental changes are justified. For example, ad hoc solutions are typical of organ transplants in the U.S. because there are not yet any global limits. The sentiments of the general public and some legislators seem to favor public funding or private insurance coverage for *all* organ transplants as they become technically feasible. There is little concern for budget limit or cost effectiveness. The executive branch's excuse for not covering some organ transplants is that they are still experimental. This means that if the transplants were no longer experimental, the government would have to think of other excuses not to cover them. In other words, nobody is willing to put a global limit on organ transplants that is based on some sound, though politically unpopular, general principles.

It is obvious that technical feasibility can never be the general principle for setting global limits on any slippery slopes in an age of rapid technological changes. Technical feasibility is a convenient excuse only because it happens to be the most immediate and politically acceptable constraint. When

[5]This process of ad hoc solutions is similar to the formation of constitutive norms through normalization. Constitutive norms are those which are made up and changed in the process of social interactions. In interactions from which withdrawal of participation is difficult, such as among members of a family or a close-knit group, individuals may unilaterally change the rules, or the rules may, by group interactions, be shifted to a new ground. Behavior which is unacceptable under old rules (deviations) is normalized under the new rules. Normalizations will continue as long as the satisfactions from interactions are higher than those sacrificed by continued normalization [Lemert (1967)].

technical feasibility is no longer a constraint, organ availability may become the next constraint. As long as this constraint is binding, politicians can be generous about public funding of all technically feasible transplants. They know very well that the specter of unlimited transplants is not going to haunt them for some time. For example, the federal government now pays practically all the costs of kidney transplant operations. But at any moment, 7,000 to 10,000 people are on dialysis awaiting new kidneys (*Wall Street Journal*, September 25, 1985, p. 31).

While organ availability is still a binding constraint, the strength of the next constraint is already being tested. For example, the California legislature decided in 1983 to cover liver transplants for almost every one who could not afford them. Since California already devotes 10% of its budget to health care, the coverage of liver transplants means that other health services have to be trimmed. Predictably, the California legislature also decided that it could no longer afford to pay for health care for more than 250,000 low-income adults, for services except those 'designed to prevent serious disability or death', or for care at dozens of high-cost facilities. As liver transplants are the most costly organ transplants (costing up to $300,000 each), the prospect of providing more and more health care for fewer and fewer patients poses an imminent threat to many of the medically indigent who happen to suffer from run-of-the-mill diseases that are not politically hot (*Wall Street Journal*, April 12, 1984, p. 1).

Technical feasibility, organ availability, and transplant costs are in the nature of migration costs. In the absence of government funding and private insurance coverage, the number of transplants can be easily contained because of high migration costs. Only those who can afford them receive transplants. Those who can't afford them may receive transplants only at very high application costs associated with mass media appeals. But once government funding and private insurance coverage is available, some general principles governing global limits must be agreed upon if unlimited transplants are not to ensue. If there is no agreement on general principles, access to funding would have to be limited once again by high application costs. That means additional funding would have to be provided as long as the mass media are willing to come up with politically attractive anecdotes. Fortunately, even mass media and politicians do not have an unlimited reserve of humanitarian energy for one particular special interest, especially after the novelty wears out. And when potential organ donors realize that organ transplants are squeezing out other essential health care services, they may be less willing to donate organs until some general principles on organ allocation are agreed upon.

Absence of general principles on global limits is typical of situations where there is ignorance about the location of defensible segmentations. Minor deviations serve to explore the location of defensible segmentations and help

to define general principles. More importantly, absence of general principles on global limits may simply reflect the lack of political support for radical solutions. For example, at the initial stage of organ transplant technology, it is most unlikely that people will accept a priority for those who have led a healthful lifestyle or for young people who are willing to be sterilized so that the hereditary diseases that make the transplants necessary will not be passed on to the next generation. Such political support is more likely to be forthcoming after a sense of crisis is produced by a free slide between widely-spaced defensible segmentations.

A consensus on general principles governing global limits is particularly likely if simultaneous efficiency-reducing slides are occurring on several slippery slopes. A return to defensible segmentations could then be clearly seen as Pareto superior for all the special interests [Buchanan et al. (1980, pp. 366–367)]. The recent tax reform in the U.S. to swap tax loopholes and deductions for lower income tax rates is such a case in point.

Defensible segmentations are of course subject to erosion. Technology can drastically alter their locations and precipitate a free slide. But technology may also shift the whole landscape so that a free slide becomes less significant, or a slippery slope may be transformed into an invisible hand.

9. Minor deviations and tragic choices

When minor deviations are allowed from a regime of total conformance, the choice of whom to exclude and include can have severe consequences for those adversely affected. If these consequences impact on our incommensurable core values, they may be regarded as tragic [Calabresi and Bobbitt (1978)]. For example, educational exemptions from military draft are acceptable if they are randomly distributed among various ethnic and income groups. They are not acceptable if they are solely received by the rich. This is because in the former case the incommensurable core values of life and equality are not in conflict but in the latter case they are. Tragic choices would not have been necessary in an invisible-hand context because the feasibility of market pricing presupposes value commensurability. Even on a slippery slope, a global limit on permissible deviations would not induce a sense of tragedy if only incommensurable, but not highly prized, values are affected [Jones (1984?)]. For example, those drivers who can afford radar detectors are less likely to be caught speeding on the highway than those speeders who cannot afford them. Although here again two incommensurable values, i.e., equality before the law and the desire for speed, conflict, such a conflict is seldom regarded as tragic.

10. Summary and conclusion

A slippery slope can be formalized as a multi-party prisoners'

dilemma (PD) game where minor deviations from an unstable Pareto superior solution in pursuit of self interests lead to a stable Pareto inferior solution of total defection. The prediction of total defection is based, however, on the assumption that the payoff from one action (B) always exceeds the payoff from another action (A) although pursuit of the higher-paying action leads to declining payoffs for both actions. If the dominance of B over A reverses as defection proceeds, slides on the slippery slope will be self-arrested short of total defection. Even with total dominance of B over A, slides can still be arrested by adjusting the transaction costs of defection to eliminate all or part of the net advantage of B over A. When minor deviations can be thus contained, they may well be collectively though not Pareto optimal

Total defection, therefore, does not usually occur unless transaction costs are drastically reduced. When total defection does occur, reversibility of the slide depends on the size of the minimum critical effort required to revive a critical mass of conformance. This critical mass is easier to achieve if people are less envious and more Pareto altruistic after they have experienced the deprivation of total defection. But as long as the dominance of B over A persists, this critical mass is not likely to be self-sustaining. If the dominance of B over A is reversed on the return trip to conformance, then a self-sustaining critical mass is possible as soon as the threshold dominance of B over A at total defection is overcome.

The greatest challenge of containing defection is the uncertainty about the underlying game-theoretic situations. This uncertainty is particularly serious when technology may transform a slippery slope into an invisible hand or vice versa. Under an invisible hand, the correct strategy is incremental adjustment towards the optimum. On the other hand, the correct strategy on a slippery slope is to stick to a defensible segmentation. When one is not sure whether one is under an invisible hand or on a slippery slope, incremental changes represent an optimal search strategy. This is especially true under constant technological changes when migration costs and enforcement costs are continually being altered. Incremental changes as a search strategy does not require consensus over general principles but only ad hoc agreement over specific policies. Incremental changes will settle down to defensible segmentations when huge discontinuities produce a free slide. The resulting sense of crisis will create the political consensus for a global limit.

Limiting slides on a slippery slope is, of course, the concern of only members of the group who benefit from total conformance. A Pareto superior solution of total conformance, while beneficial to the group members, may well be detrimental to the interest of non-group members. Indeed, the antitrust laws in the U.S. are specifically enacted to prevent collusion at the expense of helpless consumers. However, total conformance is not necessarily immoral even though non-group members may be hurt. For

example, the cooperative action of non-confession by political prisoners fighting a tyrannical regime may well be regarded as an act of moral courage [Ullmann-Margalit (1977, pp. 41–44)]. Thus, the morality of conformance or defection on a slippery slope depends on whose interests we identify with.

References

Buchanan, James M., 1980, Reform in the rent-seeking society, in: James M. Buchanan, Robert D. Tollison and Gordon Tullock, eds., Toward a theory of rent-seeking society (Texas A&M University Press, College Station, TX).

Calabresi, Guido and Philip Bobbitt, 1978, Tragic choices (Norton, New York).

Fung, K.K., 1979, On romantic love: An analysis of open system economic behavior, Policy Sciences 11, 179–186.

Hardin, Garett, 1968, The tragedy of the commons, Science 162, 1243–1248.

InfoWorld, 1986, Publisher to drop copy protection, May 5, 11.

Jones, W.T., 1984(?), Impossible choices, Working Papers (Centre European de Difussion Scientifique).

Lemert, E.M., 1967, Social structure, social control, and deviation, in: E.M. Lemert, ed., Human deviance, social problems, and social control (Prentice-Hall, Englewood Cliffs, NJ).

Mendeloff, John, 1985, Politics and bioethical commissions: 'Muddling through' and the 'slippery slope', Journal of Health Politics, Policy and Law 10, 81–92.

Schelling, Thomas C., 1978, Micromotives and macrobehavior (W.W. Norton & Company, New York).

Smith, Adam, 1776, An inquiry into the nature and causes of the wealth of nations, 1937 edition (Modern Library, New York).

Ullmann-Margalit, Edna, 1977, The emergence of norms (Oxford University Press, Oxford).

Wall Street Journal, 1984, Transplants increase, and so do disputes over who pays bills, April 12, 1.

Wall Street Journal, 1985, Rising success in organ transplants strains hospitals and governments, September 25, 31.

Wall Street Journal, 1986, Recent awards in Florida reflect scope of nation's liability woes, April 21, 21.