# Road Rage: Boundedly Rational Learning and Enforcement via Simulated Annealing

by Roger A. McCain

Classical game theory assumes that agents always choose optimal strategies and assume that other agents in the game also do so. However, a good deal of experimental evidence (as well as common sense) suggests that people are not so perfectly rational and only learn by experience to choose optimal strategies, if they do so at all.[1] Much evidence from experimental game theory is consistent with the idea that agents choose strategies probabilistically, but assign to each strategy a probability commensurate with the success the agent has had with that strategy in the past.[2] This would lead to a probabilistic Nash equilibrium, (Mailath) in which each agent chooses strategies with probabilities commensurate with the success of those strategies, given the probabilities with which other agents choose strategies.

This invites computer simulation, for three reasons: first, because the problem is complex enough that an analytic solution is likely to be difficult or too abstract to be useful; second, because transient phenomena may be important determinants of long-term tendencies, and simulation admits of studying the transient processes in detail, while analytic solutions for final equilibria do not, and third, because there is a large literature and experience in simulating probabilistic processes that we are able to draw on.

However, agent-based simulation models in economics tend to follow the same broad tradition, assuming that agents maximize utility, even if they do it imperfectly. Examples may be found in simulations that assume boundedly rational and imitative learning. A number of papers have studied imperfect imitative learning as a utility-increasing activity (e.g. Dawid, McCain 2000). Some studies of imitative learning have taken account of the tendency of people to imitate others who are "near" them in some sense (e.g. McCain 2000, Bala and Goyal). As Axtell observed, however, imitation may not be utility-increasing and may be motivated quite differently. Indeed, as McCain (1992) observed, imitation may lead to destructive and self-destructive behavior, as in road rage.

Experimental evidence suggests that real decisions may arise from a variety of motives, including altruism (Levine) simple conformism (Bikhchanandi, Hirschliefer, and Welsh), a sense of common identity (McCain 1992; van der Heiden et. al.), and fairness (Rabin) in the sense of reciprocity (Sugden; Berg, Dickhaut, and McCabe; McCabe, Rassenti, and Smith; Hoffman, McCabe and Smith; Shefrin and Triest; Fehr, Kirchsteiger, and Reidl). This multiplicity of motives lends itself to an "impulse-filtering" model (McCain, 1992) which would generate a probabilistic choice function (Mailath; Erev and Roth; Chen, Friedman, and Thisse). It seems appropriate that agent-based simulation studies might incorporate some of these more "realistic" aspects of human behavior, and probabilistic choice models, in which the probabilities may be influenced by more than one sort of motivation, supply a simple framework for such a pluralistic approach.

In many studies, hypotheses based on rational (utility-increasing) motives give rise to positive statistical tests, but other sorts of hypotheses, such as epidemiological ones, also fit some of the data well. McCain (1992) proposes a model of choice behavior designed to reconcile these observations. In that model, decisions are set in motion by random impulses. The decision to act on an impulse or not are determined by a series of probabilistic filters, any one of which may suppress or transform the impulse. Rational decisions arise from the action of a "cognitive filter," but there may be other filters, such as filters of conformity, reciprocity, or obsession. Moreover, the cognitive filter is fallible, and may be less effective if e.g. the individual is aroused.

This paper reports simulations of a population of semi-rational agents playing a simple aggression-retaliation game. The agents are modeled along lines suggested by the impulse-filtering model, and in particular arousal is allowed for. The agents learn optimal behavior, if at all, only through experience with boundedly rational learning. Simple as this model is, it may be used for policy assessment. To illustrate this, we explore the rationale and implications of a law enforcement strategy to reduce the number of road rage incidents.This policy -- implemented by the Washington State Police in 1998 (Watson) -- focuses primarily on penalizing aggressors, not retaliators. Simulations of such a policy with and without arousal are compared to the no-enforcement case.

## Learning Optimal Strategy Choice and Simulated Annealing

Thus, we propose to simulate choices made probabilistically, with probabilities influenced but not determined by considerations of increasing utility. Simulated annealing (Goffe, Shaffer and Small, Wu and Wang) has promise here. Simulated annealing has been used in numerical optimization where the objective surface is complex so that conventional optimization algorithms (such as Newton-Raphson methods) may become "trapped" in local optima that are considerably inferior to the global optima. As such, simulated annealing is considered an alternative method to genetic algorithms.

The simulated annealing algorithm is simple, and is also known as the "metropolis process." The system is initialized in a randomly chosen state. At each step the current state is compared with a randomly chosen alternative state and $\Delta E$, the change in the value to be maximized or minimized, is calculated. (For this economic application we focus on maximization). If $\Delta E$ is positive, the system is shifted to the new state which becomes the current state, as in any hill-climbing algorithm. However, if $\Delta E$ is negative, then the system is shifted to the new state with probability $\exp(\Delta E/T)$ where T is a parameter, and kept in the current state without change with a probability of $1-\exp(\Delta E/T)$. This (generally small) probability of jumping to a state that is inferior to the current one is what makes it possible for the algorithm to escape from local optima and to find the global optimum with good reliability.

As the name suggests, simulated annealing is based on annealing in mechanics, a process that roughly minimizes the energy (irregularity) of a metal. Thus, the parameter T may be read as the temperature; T is also known as the resist. In all cases of annealing -- simulated or actual -- the temperature or resist is gradually reduced, so that the stability of the system increases as it approaches its optimal state. But our knowledge of the appropriate rate of decrease of temperature is empirical, and the optimal control of T is not understood. Figure 1 shows the probability of a negative step as it varies with the size of the step for T values of 100 (light gray) 10 (dark gray) and 1 (black).

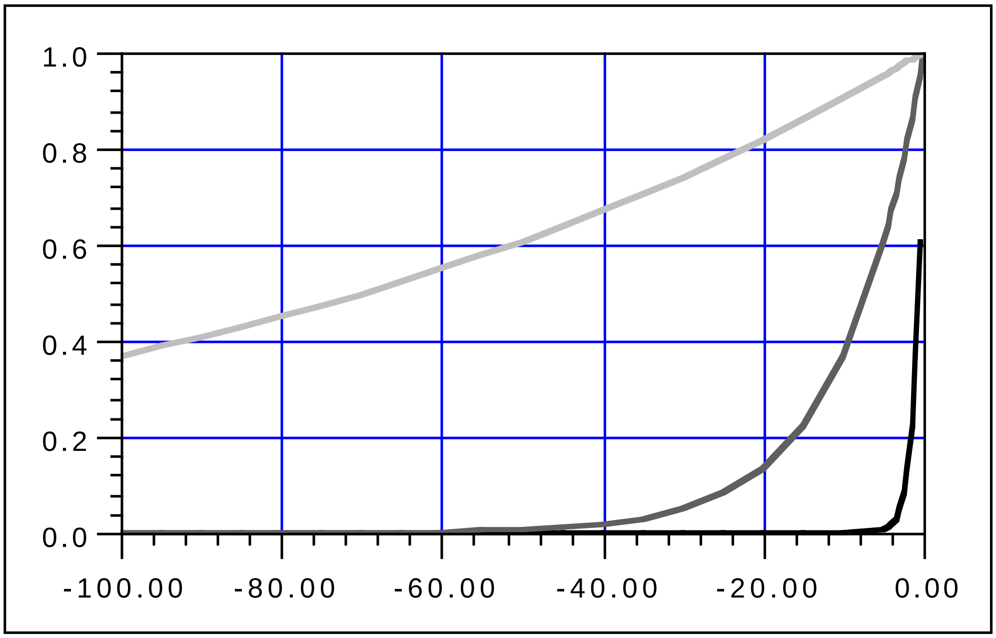


**Figure 1. Probabilities of Utility-Reducing Choice with Three Resist Values**

For present purposes, however, simulated annealing needs to be adapted in several ways.

a) Learning by experience. In the optimization applications, the objective function value for each state is, in effect, known in advance. By contrast, boundedly rational human beings must learn by experience the payoffs from the different strategies that they may choose.

b) Errors of two kinds. The metropolis process allows the agent to choose a new state with a lower objective value with some probability, but never to fail to choose the new state if it has a higher objective value. This probably makes for more reliable optimization, but real human beings do sometimes fail to adopt potential improvements, so the adapted process should allow for this.

c) Non-self-interested motives. In that $\Delta E$ is a difference of utility or payoffs, some changes in the process may be required to introduce non-self-interested motivations such as altruism, reciprocity, and conformism.

To meet some of these issues, we modify the process by allocating probabilities among them in such a way that $p_1$ is the probability that alternative 1 will be chosen, $p_2$ is the probability that alternative 2 will be chosen, for application to 2x2 game examples. The probabilities are calculated in two steps. In the first step, the agent recalls his past average utility experience for alternative i and compares it with his past average utility experience with each of the other. Then E is computed as the difference between past average utility for i and the past average utility of the alternative. In place of $\exp(\Delta E/T)$ we than compute $\exp(\Delta E/T+\ln(g/(1-g)))$. The parameter g is the probability that a new state will be chosen if there is no difference between the values of the objective function in the new state and the old. This gives rise to a probability curve like Figure 2, which should be contrasted with Figure 1. In Figure 2, g is 0.9 and the "temperature," or resist, is 0.5. Thus the probability of choosing state 1 is no less than 0.9 if the state 1 is better than state2.

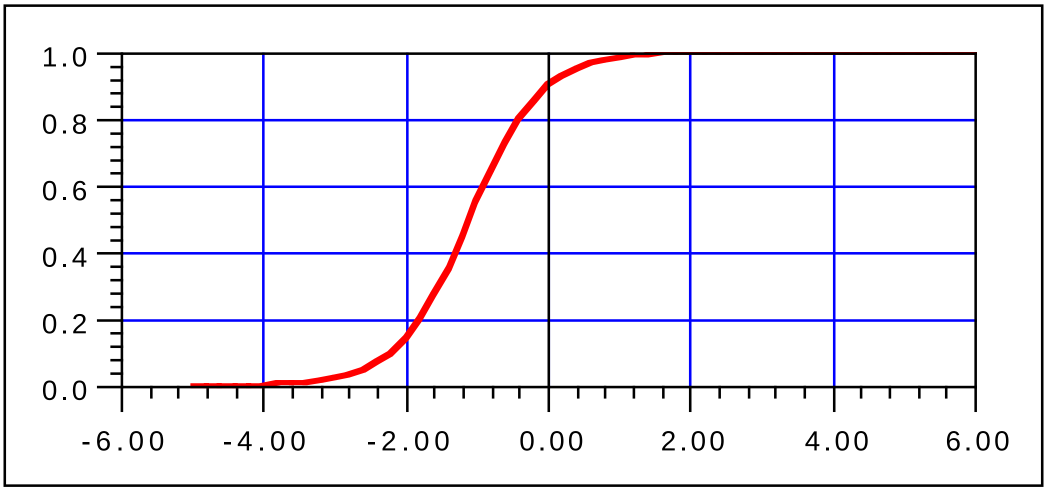


**Figure 2. Probabilities for a Model Allowing Two Kinds of Errors**

Clearly, for the large value of the resist, the probabilities are roughly equal over a wide range of Δ, while for the smallest value of the resist, the choice is nearly determinate except for very small Δ. For a dichotomous choice the default value of g would be 0.5, so that, with no experience of difference between the outcomes of the two alternatives, the agent would choose them equiprobably.

## A Measure of Specialization

We will be interested in the tendency of agents to specialize in the "better" choice or strategy. Accordingly, we will need a measure of agent specialization. Since the choice is seen as being probabilistic, it lends itself to an entropy measure. The entropy of a probability distribution is $N=-\Sigma p_i \ln(p_i)$ where $p_i$ is the probability that portal i will be chosen and ln denotes the natural logarithm. (Since the logarithms are negative, the overall measure is positive). Entropy is an inverse measure of concentration of the probabilities, and thus of specialization. Entropy is at its maximum when each alternative is chosen with the same probability and is zero when one portal is chosen with probability one. For a dichotomous choice maximum entropy is 0.69. In the simulations, we will measure specialization by inverse entropy, 1/N, which will take a minimum value of 1.44 for a dichotomous choice and a maximum value of infinity.

This model of probabilistic choice and learning was tried in several baseline simulations, involving non-interactive polychotomous choice and several well-known two-by-two games in normal form. These simulations confirm that optimal strategies are learned and the learning evolves in ways that differ slightly according to the mixture of motives in the game.

## Law Enforcement and Subgame Perfection

"Road rage" is a kind of aggression-retaliation "game" and susceptible to analysis in terms of game theory. Since it involves a sequence of action (by an aggressor) and retaliation (by the victim) "road rage" calls for analysis in terms of sequential games and subgame perfection, which is the standard of "rational" behavior in such games. These concepts will be sketched intuitively in terms of the simplified "road rage" game shown in Figure 3.

The figure shows that the potential aggressor, A, first decides whether to aggress or not. If he does not, we follow the arrow indicated by "don't," and the aggressor gets a payoff of 5 while the (potential) victim's payoff is 10. If the potential aggressor chooses "aggress," the victim chooses whether to retaliate or not. If the victim does not retaliate, then the aggressor gains with a payoff of 10 versus 5 for the victim. However, if the victim chooses "retaliate," both parties loose 20. These numbers are realistic only in the qualitative sense that the aggressor has something to gain if there is no retaliation, but that the losses in case of retaliation are greater than anything anyone has to gain.

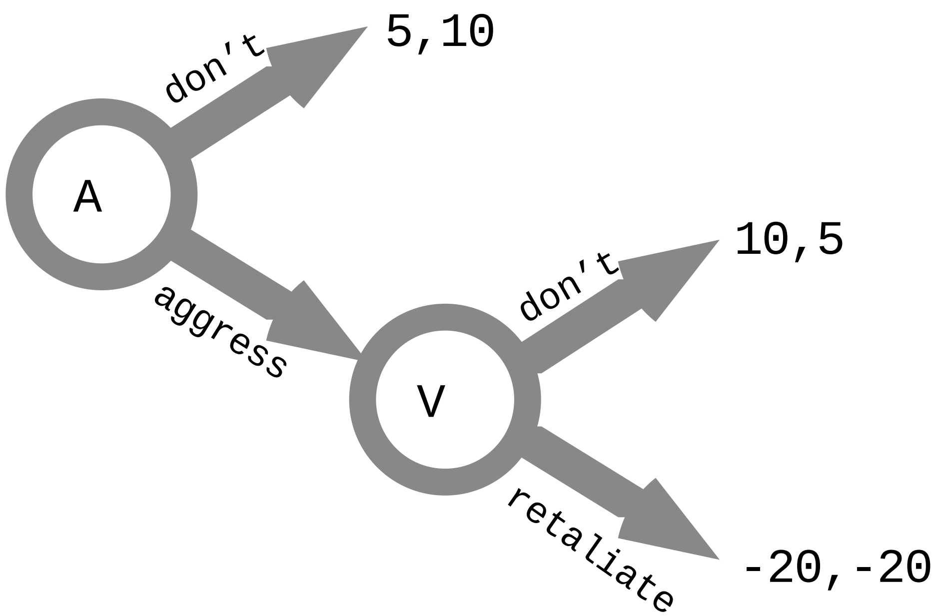


**Figure 3. "Road Rage" in Extensive Form**

The criterion of "rationality" in such a game is "subgame perfect equilibrium. (Selten) "Subgame perfection" reflects the common idea that the past cannot be changed, and that a rational person will make decisions that reflect this fact. Thus, at each stage, the decision-maker chooses the highest payoff, taking whatever has gone before as given and unchangeable. Thus, in particular, the victim of aggression must choose between a payoff of +5, for nonretaliation, and a payoff of -20 for retaliation. The rational choice is nonagression. (If the participants do not expect to interact repeatedly in the future, as in the anonymous face-offs of highway traffic, there is no benefit from gaining a reputation as a "retaliator.") But the potential aggressor can anticipate this choice, and thus knows that the payoff of aggression (against a rational victim) is 10 rather than 5. Thus aggression is the rational choice and (agress, don't retaliate) is the "subgame perfect equilibrium" for this game.

It is a bit more complex than this if some participants are not rational or if rationality must be learned. In that case, the potential aggressor has to judge whether the probability of facing an irrational or naive victim who will retaliate is offset by the gains from aggression in case the victim acts rationally. In this game, with a frequency of irrational retaliation less than 16.67%, aggression remains the subgame perfect strategy at the first stage of the game.

"Rationality" is important from the point of view of enforcement. In this game, aggression per se raises no efficiency issues[3] although it may raise issues of justice. However, retaliation leads to efficiency losses. Thus, prevention of retaliation is a particular focus of enforcement. However, legal penalties for retaliation may not be effective. Penalties shift the payoffs -- the rational determinants of actions -- to make retaliation more costly. However, retaliation by its very nature is costly enough that it will not be rationally chosen, which suggests that the penalties are likely to be ineffective. This is a rationale for the Washington State policy. A sufficient penalty on aggression would make non-aggression the subgame perfect strategy at the first stage, and thus eliminate the occasion for socially costly retaliation. Penalizing aggressive behavior is the enforcement strategy that corresponds to the assumption of universal rationality. In a population of mixed "rational," naive and irrational agents, we have a reasonable hope that the rational component of aggressive behavior at least could be eliminated, thus reducing the frequency of both aggression and retaliation.

The Road Rage Game Simulation

The Road Rage game was initially tested in a simple random-matching model with a population of 1000 and without spacial differentiation, arousal, or non-self-interested motivations. On each round of play 100 matches took place. In each round, an agent chosen as random was a potential aggressor and chose between the strategies "aggress" and "don't aggress." A potential victim was also chosen at random but had to choose a strategy only if the potential aggressor chose "aggress." In that case the victim chose between "retaliate" and "don't retaliate." While clearly the subgame perfect equilibrium is "aggress, don't retaliate," some experience would be required to learn this equilibrium. Initially, with no experience to draw on, agents choose both strategies 50-50. Moreover, after every 100 matches, agents were selected at random with a 0.03% probability, and re-initialized -- forgetting whatever experience they might have had and resetting the agent's resist value to the starting value.

We do observe some convergence toward equilibrium behavior on the part of the victims of aggression. Figure 4 shows the evolution of the proportion of victims who retaliated in three hundred sessions of 100 matches each, in simulations with five random number seeds. The five random number seeds were color-coded as shown in Figure 5.

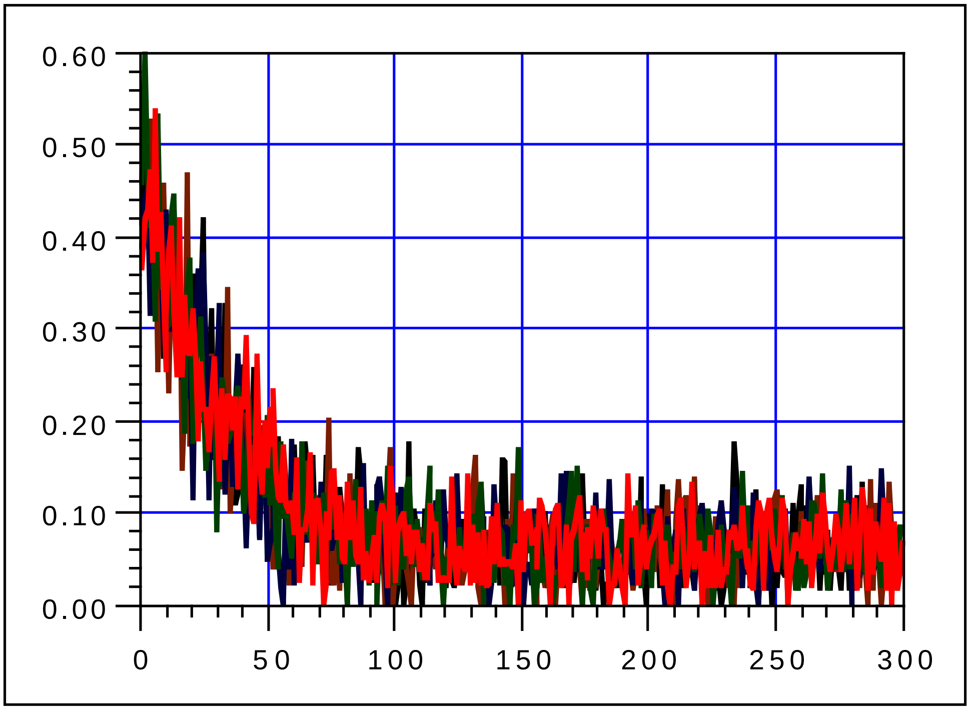


**Figure 4. The Proportion of Victims who Retaliated in 5 Families of Simulations**

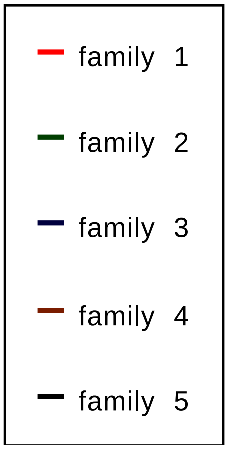


**Figure 5. Color-Coding for the 5 Families**

We see that the proportion of retaliators declines over the first 100 sessions to about 0.06, and thereafter remains steady on the average despite a good deal of fluctuation. The average for all five simulations for the sessions from round 101-300 was 6.13% of incidents of aggression, with a standard deviation of 3.34%. This incidence of retaliation is entirely a consequence of forgetting: in simulations without forgetting, the rate of retaliation drops to zero. However, with forgetting at 1%, the rate of retaliation from 101-300 was 14.4% of retaliators, with a standard deviation of 5.05%. The rate of retaliation that balances the expected value of aggression against that of nonaggression is 0.1667, so with forgetting at 1%, retaliation closely approximates the balanced game -- the incentive to choose either an "aggressive" or "nonaggressive" strategy is slight, in such a case.

With retaliation at about 6%, however, it would seem that there is plenty of incentive for a rational agent to aggress. Looking at the results we see some, but a distinctly limited tendency in that direction. Figure 6 shows the proportion of agents choosing "aggress" over the 30000 plays.

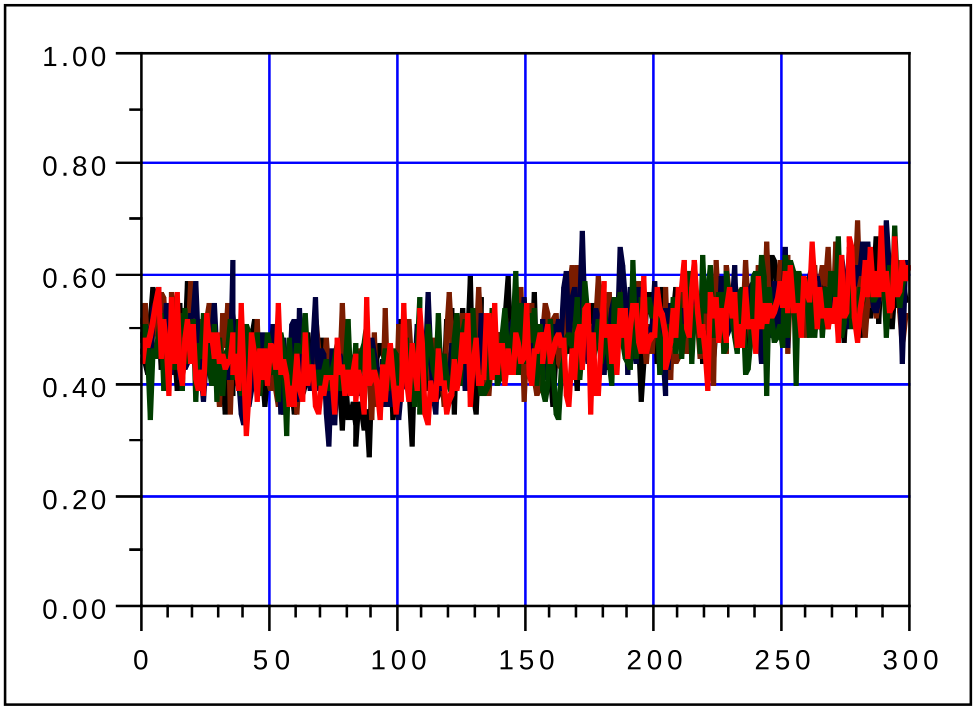


**Figure 6. The Proportion of Aggressors**

The different random seeds are coded as before. We see that, after an initial dip, the proportion of aggressors rises so that a distinct majority are aggressors -- 0.5428 of agents are aggressors in the last 100 sessions, with a standard deviation of 0.0543 -- but it is clear that many agents are failing to learn the subgame perfect equilibrium, which calls for 100% aggression at the first stage. This datum is not highly sensitive to the rate of forgetting. Table 1 shows the average proportion of retaliators and of aggressors for selected rates of forgetting.

**Table 1. Aggression, Retaliation, and Forgetfulness**

| | retaliators | S. D. | aggressors | S. D. |
|---|---|---|---|---|
| forgetting = .003 | 6.13% | 3.34% | 54.28% | 5.43% |
| forgetting= .005 | 8.75% | 4.03% | 53.67% | 5.31% |
| forgetting= .01 | 14.4% | 5.05% | 48.85% | 4.86% |

Nevertheless, there is a clear utility gain for the first-stage player, the potential aggressor. The evolution of the average utility for the 300 sessions of 100 games each is shown in Figure 7, with, once again, forgetting at .003. Enough equilibrial interactions take place so that the utility of the average first-stage player rises above the payoff to nonaggression. Indeed, the average net payoff to aggression is usually positive. It is shown for the five simulations in Figure 8. This is the difference between the expected payoff to aggression and that to non-aggression by agents. While the differential reward is at its highest in the first 50 sessions, it remains positive on the average and averages 2.88 for the last 200 sessions in all five simulations. Why, then, is there so little aggression? The differential reward is an average of two kinds of outcomes -- successful aggression, with a payoff of 10, and aggression with retaliation, with a payoff of -20. Agents who try aggression early and acquire a negative experience may never again try the experiments that would lead to a growth in the experience of positive payoffs that would reverse their experience.

Accordingly, we look at the proportion of agents who, having been chosen as potential aggressors, have experienced negative utility from aggression on the average, in the past. This is shown if Figure 9. The random seeds are coded as before. We see that, after the first 50 sessions, a majority of agents have negative average utility from earlier adventures in aggression. For the last 200 sessions, the average proportion of first-stage agents with negative prior experience is 60.56% with a standard deviation of 4.88%.

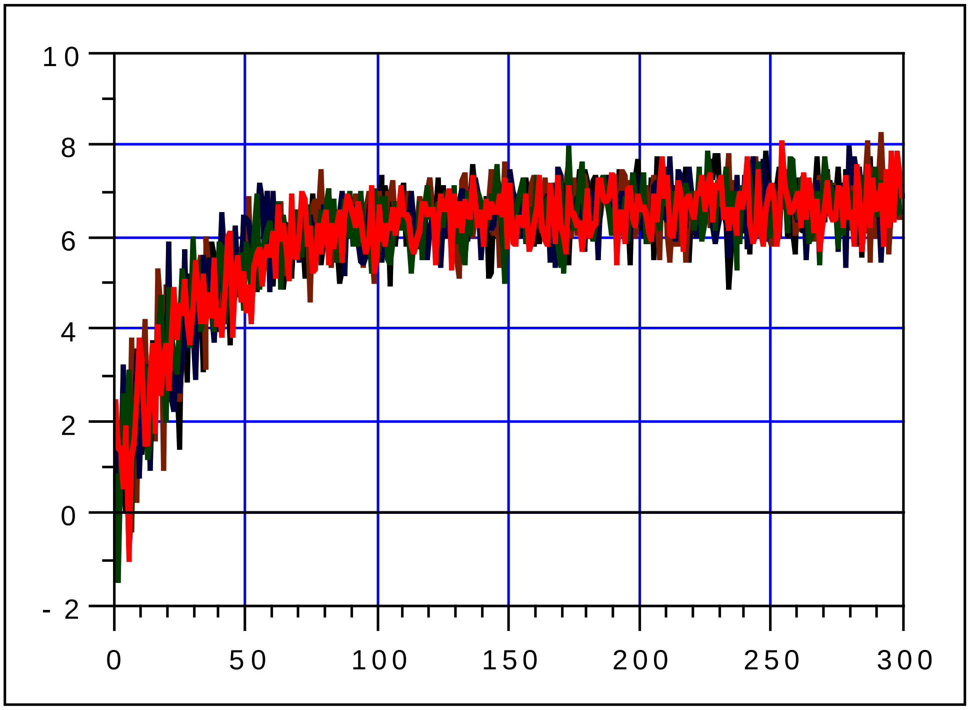


**Figure 7. Average Utility**

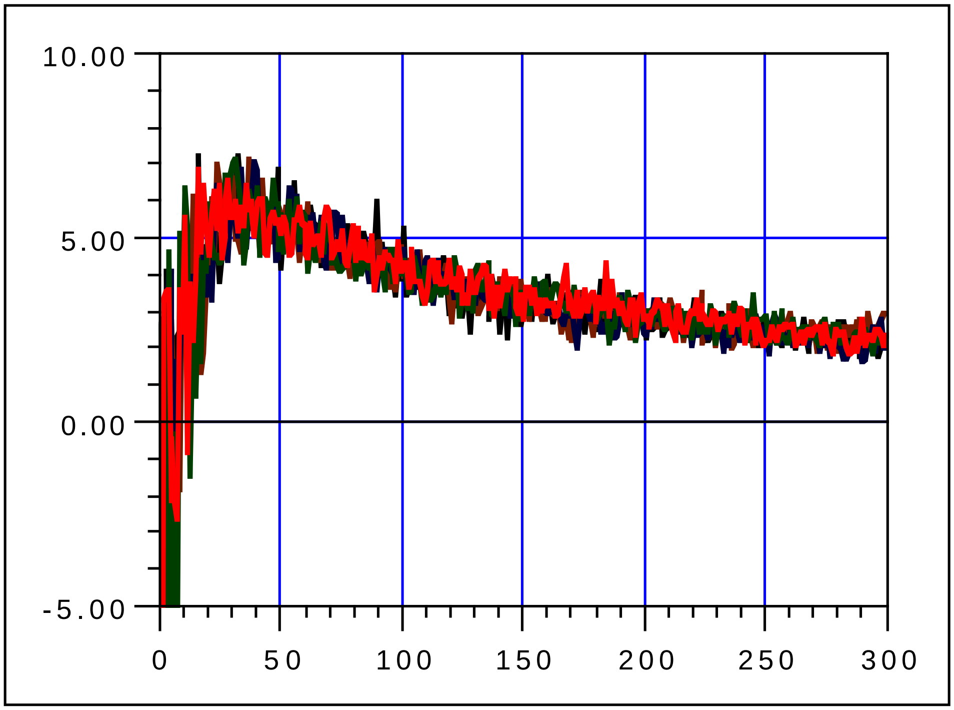


**Figure 8. The Average Net Payoff to Aggression**

Apparently the slow convergence to the subgame perfect equilibrium reflects the persistent effects of bad experiences with aggression in early stages when naive retaliators are quite common, and subsequent encounters with forgetful retaliators. If the agent has only a little experience, and a high resist value, the probability of aggression is significant even if past experience is against it -- "habits" have not yet been formed. But on the other hand, if the agent has enough experience so that the resist is quite small, the probability of a utility-decreasing choice is small, and to many potential aggressors, aggression appears on experience to be a utility-decreasing choice. In the words of my Pennsylvania-Dutch mother, "once burnt, twice shy."

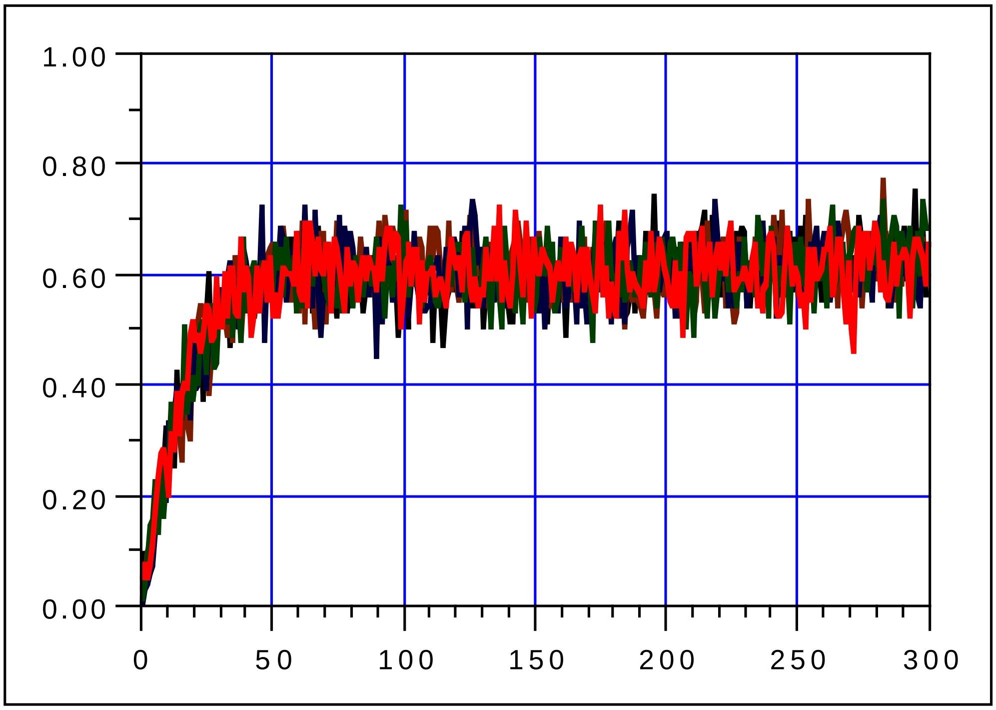


**Figure 9. The Proportion of Potential Aggressors with Bad Memories of Aggression**

It is, then, forgetting that causes retaliation but memory that limits aggression. It may be of some interest that these results deviate from rational self-interest (that is, subgame perfect equilibrium) in ways that suggest reciprocity. That is, retaliation is present when it makes the agent worse off, while aggression is limited beyond the extent justified by the real risk of retaliation. But, in fact, the agents simulated here have no reciprocity motives. Instead it is imperfect cognition, in the form of forgetting bad results on the one hand and remembering them and letting them inhibit exploration on the other, that produce the result.

## Arousal

In all the simulations reported thus far, learning is driven by increasing utility. In the impulse-filtering approach, "increasing utility" is identified with the "cognitive filter," but it is argued that the "cognitive filter" may be less effective when the individual is aroused. Arousal is readily modeled in the adapted simulated annealing framework used here. In the adapted simulated annealing model, the individual will choose the "better" alternative with a greater probability, but the distribution function depends on the resist ("temperature") and is flatter when the resist is greater, as shown by Figure 1. While the general tendency of the "annealing" process is to reduce the resist, arousal would increase it, corresponding to the general notion that the person who is aroused is "hot" and therefore unpredictable.

In applying this to the road rage game, we first situate the agents in a "line automaton," a cellular automaton of just one dimension. That is, each agent is situated at one point of a circular space of 1000 points. The agents in the four spaces on either side of a particular agent constitute that agent's "neighborhood." Agents become aroused when one of their neighbors is a victim of aggression. That is, when agent i is the second player in a random match and the first player chooses aggression, then the resist is increased by a certain proportion for players i-4, ..., i+4.

As usual, five series of simulations were run with the same five random number seeds as before. Arousal made some difference in the performance of the simulations. Figure 11 shows the average resist or "temperature" over all agents and all five families of simulations, in the simulations with as opposed to those without arousal, as they evolved over 500 steps of 100 games each. The average for simulations without arousal is shown in solid red, and the average for simulations with arousal is shown by the cross-hatched black curve. The resist declined more slowly and stabilized at a higher value with than without arousal. Since a higher resist or "hotter temperature" means less predictable

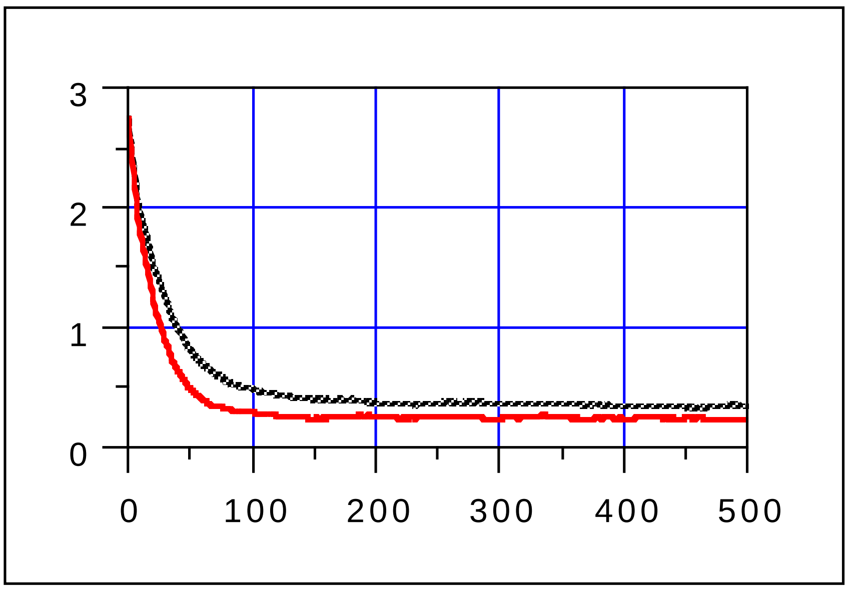


**Figure 10. The Resist with and without Arousal**

behavior, we might expect to see both more irrational behavior (retaliation and failure to aggress) and more learning in later stages of the simulation. These would tend to offset one another.

However, this is only partly realized. Figure 11 shows the average evolution of the proportion of incidents giving rise to retaliation for the simulations with and without arousal. As before, the simulations without arousal are shown in solid red and those with arousal in discontinuous black. On the whole, there is little difference. It may be that the learning problem for victims is just simple enough so that rationality is learned despite the arousal. Figure 12 shows the evolution of incidents of aggression with the same framework. We see that there is somewhat less aggression in simulations with than in those without arousal. Noting that aggression is rational in this game, less frequent aggression is consistent with less rational behavior in the presence of arousal. As with retaliation, there is little difference in the convergence of the two groups of simulations to the stable utility maximum of 7.5 for both agents. On the whole, then, arousal makes little difference in the results of the basic model, but does retard somewhat the tendency of potential aggressors to learn that it is rational to aggress.

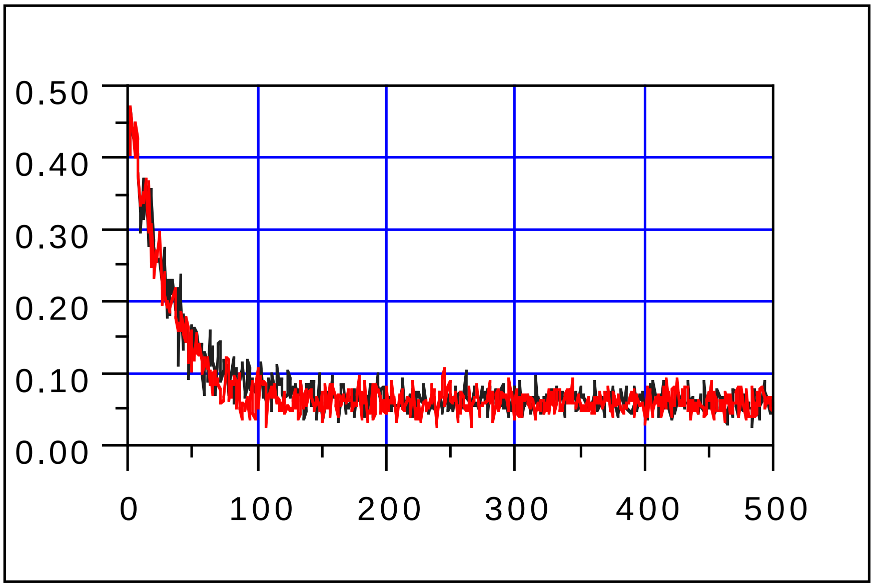


**Figure 11. Incidents of Retaliation With and Without Arousal**

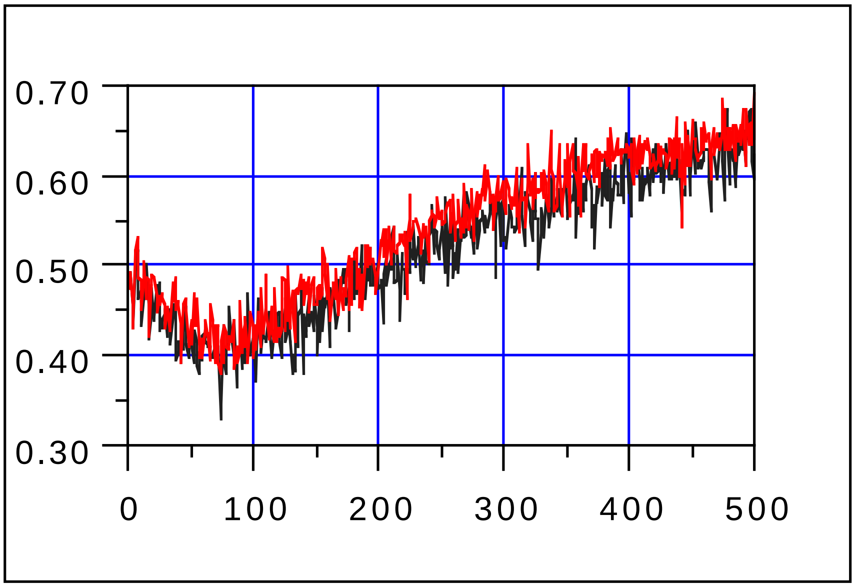


**Figure 12. Proportion of Incidents of Aggression**

## Enforcement

We now consider simulations in which there is law enforcement along the lines suggested by the policy of the Washington State Police. Agents who choose the strategy of aggression are, in some cases, stopped and fined 100. Of course, it is not possible to penalize every aggressor, and we expect that the effect of the policy will depend on the frequency of enforcement. We compare four sets of simulations in which aggressors are penalized with a probability of 0.005, 0.01, and 0.05, and in which there is no law enforcement. In these simulations, we do not have arousal.

We would expect to see enforcement have an effect on the rate of aggression. This is shown in Figure 13. This plot shows the average rate of aggression in four series of simulations, each series comprising five simulations with five random seeds, as it evolves over 500 periods of 100 plays each of the game. The average for simulations with no enforcement is shown in black; for simulations with enforcement at probability 0.005 in red, for simulations with enforcement at probability 0.01 in green, and with enforcement at probability 0.05 in blue. We see that enforcement at relatively low rates has some but relatively little effect relative to no enforcement, while enforcement at 0.05 results in a qualitatively different behavior, an overall decline in aggression. This reflects the subgame perfect equilibrium: with the probability of retaliation at about 0.06,

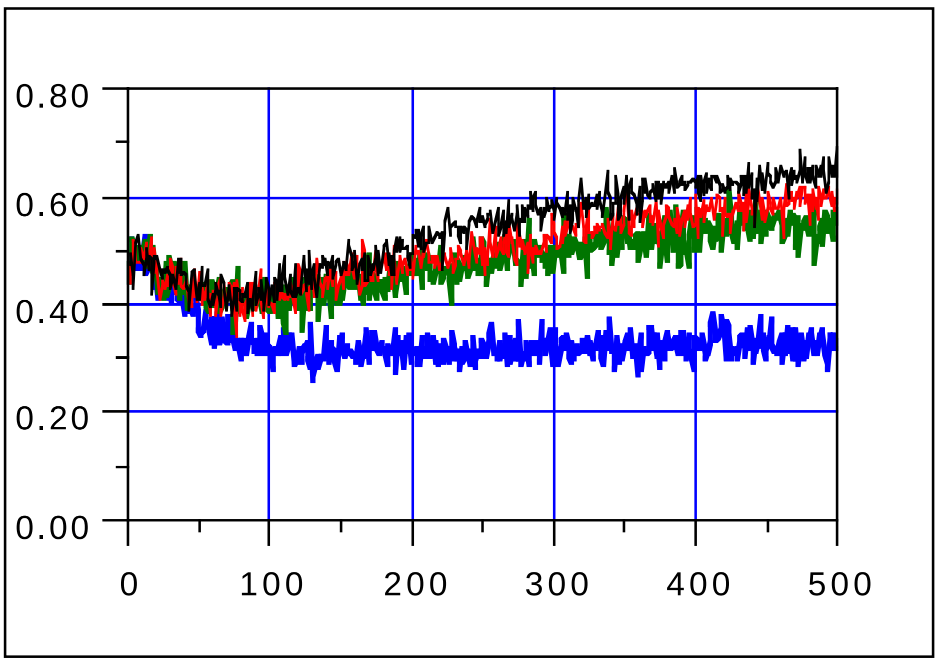


**Figure 13: The Probability of Aggression in Twenty Simulations -- Averages by Probability of Penalization**

aggression continues to be subgame perfect at a probability of punishment under about 0.035. Thus, at lower rates of punishment, law enforcement retards the rate of learning of subgame perfect strategies, which is in any case slow and incomplete; at higher rates it switches the subgame perfect equilibrium, and crossing that "tipping point" leads to qualitatively different behavior.

Since retaliatory behavior is not punished by law under this policy, we would expect little impact on the probability that victims retaliate in cases of aggression. This is borne out by the observations reported in Figure 14, with the four series of simulations coded as before.

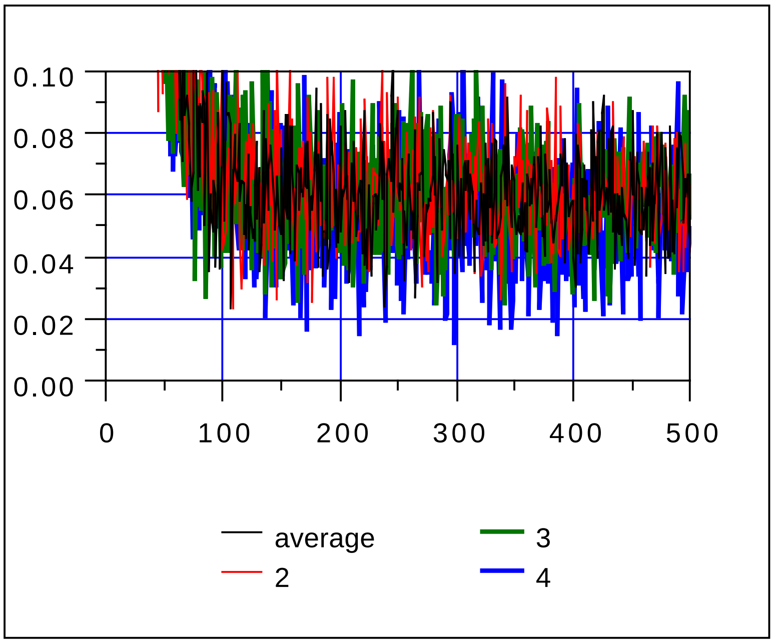


**Figure 14: The Conditional Probability of Retaliation in Twenty Simulations -- Averages by Probability of Penalization**

The objective of the policy is to affect the number of incidents of retaliation. This depends both on the probability of aggression and the probability of retaliation given that aggression does occur. Since we know that one component of the product is affected by the occurrence and frequency of enforcement and the other is not, we may anticipate the result shown in Figure 15, which is coded as before. We see a visible difference between enforcement at .05 and the other cases. In fact, there is a slight difference among the no enforcement case and the cases of enforcement at .005 and .01. The average unconditional rates of retaliation, during the last 10,000 plays of the game, for the four cases were as shown in Table 2. If we were to perform conventional statistical tests on these nonrandom data, we would judge the differences among the first three nonsignifigant.

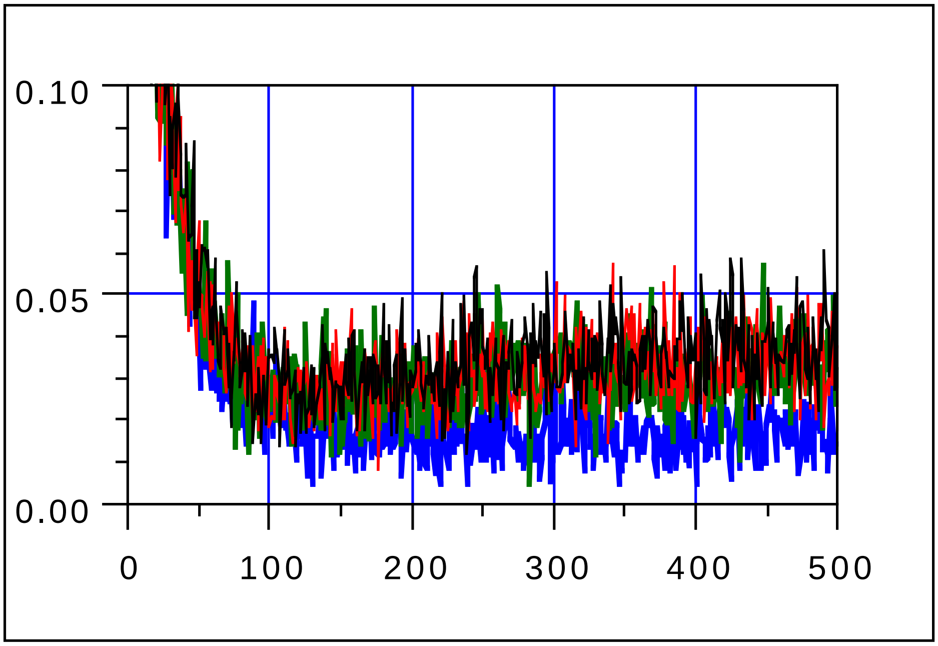


**Figure 15: The Unconditional Probability of Retaliation in Twenty Simulations -- Averages by Probability of Penalization**

**Table 2. Unconditional Probabilities of Retaliation by Enforcement Regime**

| case | no enforcement | 0.005 | 0.01 | 0.05 |
|---|---|---|---|---|
| rate of retaliation | 0.0375 | 0.0347 | 0.0323 | 0.0170 |

As we would expect, there is little or no difference in the rates of learning among the four series of simulations, and this series will not be shown.

The convergence of the average utility for the four series of simulations is shown in Figure 16. Here we see lower stable utility levels for the cases with law enforcement, and the utility is lower as the frequency of punishment rises. In part, this is unsurprising. The utilities here are those of the players in the game, and so are net of the fines imposed on those who attempt aggression. If the fines replace some taxes in financing public goods, then there would be an offsetting benefit from the public revenue through fines. Moreover, recall that the subgame perfect behavior is Pareto-optimal, so that offsetting efficiency gains are limited to the reduction in retaliations, while retaliation in later (near equilibrium) plays of the game is in any case very uncommon. Further yet, the cost of retaliatory behavior in this version of the game is slight by comparison to both the gains from aggression and the penalty. Nevertheless, this difference points up a fundamental social dilemma, the problem of any benevolent authority, which is illustrated by the well-known dilemma of monetary policy against inflation. If the penalties imposed were not

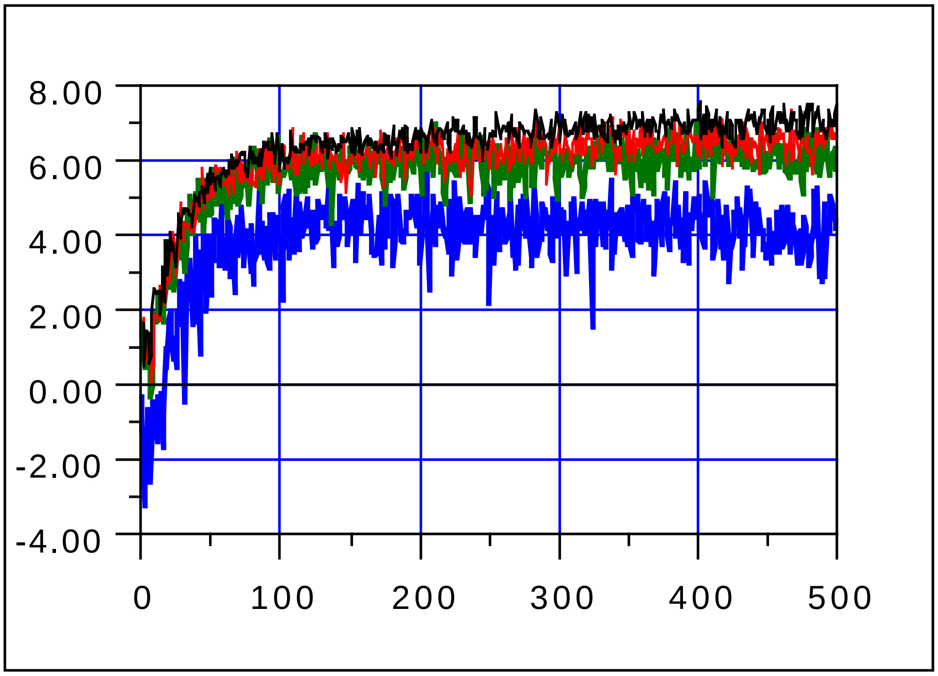


**Figure 16: The Convergence of Utility in Twenty Simulations -- Averages by Probability of Penalization**

fines but some non-transferable disutility -- for example, the if the emotional discomfort of the traffic stop were substantial relative to the fine -- so that the penalties were a net loss, then the efficiency of penalties for aggression would be subject to serious question.

As the next stage, a parallel series of simulations was conducted with arousal. Arousal was treated as before, with one qualification. As before, an act of aggression leads to arousal on the part of all the neighbors in an eight-person neighborhood in the line automaton; however, in addition, a legal penalty also creates arousal in the neighborhood of the person penalized.

As before, the model with arousal behaves in a way that is qualitatively similar to its behavior without arousal. Accordingly, we will limit our attention to the differences in the last 100 sessions, that is, the last 10000 matches. Table 3 shows the rate of aggression in eight series of simulations -- four with, and four without aggression, and with no enforcement or enforcement at three different rates.

**Table 3. Rate of Aggression by Enforcement and Arousal**

| | no enforcement | 0.005 | 0.01 | 0.05 |
|---|---|---|---|---|
| without arousal | 0.63 | 0.58 | 0.54 | 0.33 |
| with arousal | 0.61 | 0.57 | 0.53 | 0.31 |

Table 3 supports two immediate conclusions. First, arousal makes very little difference to the results. Second, that slight difference is in the direction of less aggression. It is striking that this is true even in the case of enforcement at the rate of 0.05, for which nonaggression is subgame perfect. Table 4 compares the rate of retaliation conditional on an act of aggression being chosen.

**Table 4. Conditional Rate of Retaliation by Enforcement and Arousal**

| | no enforcement | 0.005 | 0.01 | 0.05 |
|---|---|---|---|---|
| without arousal | 0.0595 | 0.0595 | 0.0594 | 0.0522 |
| with arousal | 0.0581 | 0.0610 | 0.0625 | 0.0639 |

We see that, in the absence of arousal, there is essentially no relationship between the rate of enforcement and the conditional rate of retaliation; but with arousal the rate of retaliation increases with the rate of enforcement! This may reflect two changes. Recall that retaliation is never subgame perfect, and consequently all incidents of retaliation are attributable to bounded rationality, i.e. either to forgetting or to arousal. Recall also that law enforcement activity increases arousal. Thus law enforcement can increase retaliatory activity. Second, victims learn only in cases in which aggression actually takes place. To the extent that the enforcement is successful in reducing incidents of aggression, learning non-retaliation is retarded. Since arousal also retards learning on the whole, these effects may reinforce one another.

Since the policy is directed toward reducing the total number of incidents of retaliation, Table 5 shows the proportion of all matches that led to retaliation in the eight series of simulations.

**Table 5. Unconditional Rate of Retaliation by Enforcement and Arousal**

| | no enforcement | 0.005 | 0.01 | 0.05 |
|---|---|---|---|---|
| without arousal | 0.0375 | 0.0347 | 0.0323 | 0.0170 |
| with arousal | 0.0355 | 0.0345 | 0.0330 | 0.0195 |

Once again, the first conclusion is that arousal makes little difference. The second conclusion is a bit more complex. With arousal, the increase in the conditional rate of retaliation at higher enforcement levels somewhat offsets the reduction in the rate of aggression with higher enforcement, so that there is a bit more retaliation with arousal than without at the highest rate of enforcement. Nevertheless, enforcement at 0.05 reduces the number of retaliatory incidents by 45% in the presence of arousal and by 55% in the absence of arousal, relative to no enforcement at all.

## Concluding Summary

These simulations have explored the learning of equilibrium in a game in which aggression is subgame perfect but retaliation is not, and used these explorations for a tentative assessment of law enforcement strategy with respect to road rate. The results suggest a mixed conclusion with respect to the application of subgame perfection. First, there is some tendency to learn the subgame perfect strategies, but the rates of learning differ for the two roles of aggressor and victim. With forgetfulness, victims' strategies stabilize at an equilibrium level within about 10,000 to 15,000 matches, while aggressor's strategies continue to converge through 50,000 matches and convergence still seems incomplete at that point. This slow convergence seems to reflect a long-lasting influence of rare events of retaliation on the learning of potential aggressors: "once burnt, twice shy." It seems that the subgame perfection of aggression is a more difficult learning problem than that of nonretaliation. The resulting relative scarcity of acts of aggression might, if it were observed in a real human population, be interpreted as a result of altruism or reciprocity, but there is no altruism nor reciprocity in these simulations, so the result is attributable to boundedly rational learning. When arousal is introduced, these tendencies are only slightly retarded.

Thus, in application to law enforcement, the subgame perfection criterion needs to be used with some care. Nevertheless, we find that it does predict the long-term trends as they vary with rates of enforcement of the penalty for aggression, in comparable sequences of simulation with varying rates of enforcement. In particular, only if the rate of enforcement is high enough so that nonaggression becomes subgame perfect will the tendency toward aggression and retaliation be qualitatively effected. In other cases, penalization of aggressive behavior only retards the tendency of potential aggressors to learn the subgame perfection of aggression. Thus, subgame perfection should indeed be used in assessing law enforcement policies, but with an understanding that it describes the long-term tendencies of experienced subjects and short-term responses may be relatively slight.

## References


Axtell, Robert (1998), "Varieties of Agent Implementations," (Paper presented at the fourth international conference of the Society for Computational Economics, Cambridge, England, June).

Bala, Venkatesh and Sanjeev Goyal (1998), "Learning from Neighbors," The Review of Economic Studies v. 65, pp. 595-621.

Berg, Joyce and John Dickhaut and Kevin McCabe (1995), "Trust, Reciprocity, and Social History," Games and Economic Behavior v. 10, no. 1 pp. 122-142.

Bikhchandani, Susil and David Hirschliefer and Ivo Welsh (1998), "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," Journal of Economic Perspectives (Summer).

Chen, Hsiao-Chi and James W. Friedman and Jacques-Franciose Thisse (1997), "Boundedly Rational Nash Equilibrium: A Probablistic Choice Approach," Games and Economic Behavior v. 18, pp. 32-54.

Dawid, Herbert (1996), "Genetic Algorithms as a Model of Adaptive Learning in Economic Systems," Central European Journal for Operations Research and Economics v. 4, no. 1 pp. 7-23.

Erev, Ido and Alvin E. Roth (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," American Economic Review v. 88, no. 4 (Sept.) pp. 848-881.

Fehr, E. and G. Kirchsteiger and G. Riedl (1993), "Does Fairness Prevent Market Clearing? an Empirical Investigation," Quarterly Journal of Economics v. 108, pp. 437-459.

Goffe, William L. (1996), "SIMANN: A Global Optimization Algorithm Using Simulated Annealing," Studies in Nonlinear Dynamics and Econometrics v. 1, no. 2 (Oct) pp. 169-176.

van der Heijden, E. C. M. and J. H. M. Nelissen, J. J. M. Potters and H. A. A. Verbon (1998), "The Poverty Game and the Pension Game: The Role of Reciprocity," Journal of Economic Psychology v. 19, no. 1 pp. 5-41.

Hoffman, Elizabeth and Kevin McCabe and Vernon L. Smith (1998), "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology," Economic Inquiry v. 36, no. 3 (July) pp. 335-352.

Levine, David (1998), "Modeling Altruism and Spitefulness in Experiments," Review of Economic Dynamics v. 1, no. 3 pp. 593-622.

Mailath, G. J. (1998), “Do people Play Nash Equilibrium? Lessons from Evolutionary Game Theory,” Journal of Economic Literature v. 36, no. 3 pp. 1347-1374.

McCabe, Kevin and Stephen Rassenti and Vernon Smith (1996), “Game Theory and Reciprocity in Some Extensive Form Games,” Proceedings of the National Academy of Science v. 13428, pp. -13421.

McCain, Roger A. (1992), A Framework for Cognitive Economics (New York, Praeger Publishers).

McCain, Roger A. (2000), Agent-Based Computer Simulation of Dichotomous Economic Growth (Boston: Kluwer).

Rabin, Matthew (1993), “Incorporating Fairness into Game Theory and Economics,” American Economic Review v. 83, no. 5 (Dec.) pp. 1281-1302.

Shaffer, Ronald E. and Gary W. Small (1997), “Learning Optimization from Nature,” Analytical Chemistry v. 69, (April) pp. 236-242.

Shefrin, Steven M. and Robert K. Triest (1992), “Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance,” Why People Pay Taxes edited by Joel Slemrod () pp. 192-218.

Sugden, R. (1984), “Reciprocity: The Supply of Public Goods Through Voluntary Contributions,” Economic Journal v. 94, pp. 72-87.

Watson, Tom (1998), “Aggressive Driver Apprehension Team,” radio news report, KPLU (National Public Radio).

Wu, Lihua and Wang Yuyun (1998), “An Introduction to Simulated Annealing Algorithms for the Computation of Economic Equilibrium,” Computational Economics v. 12, no. 2 (Oct) pp. 151-169.

## Notes

[1] For a witty survey of the issues, see *The Economist*.

[2] A survey of some recent literature is available from the author on request.

[3] What the aggressor gains is no less than what the victim loses, so either allocation is Pareto Optimal -- but even if that is not true, e.g. if the payoff to a successful aggressor is 9 rather than 10, the losses due to retaliation can be of a greater order of magnitude than those from aggression per se, e.g. manslaughter as against a few seconds' delay in traffic. Thus the game analysis focuses on losses due to retaliation.