Co-evolution of punishment and cooperation: A computational exploration

Eser Sekercioglu

State University of New York at Stony Brook

Department of Political Science

Stony Brook, NY

Introduction

Recent literature on human cooperation shows that humans are usually willing to punish defectors in social dilemma type situations. (1–8). Particularly punishment of free riders is now established as an effective method of facilitating and sustaining cooperation in public good provision games. (9–11). And in fact such punitive strategies are shown to even sustain strategies that are pay off inferior (9)

Experimental studies spanning a wide variety of small scale societies in addition to university undergraduate samples show that all human societies are willing to punish defectors at a personal cost and level of punishment is found to be positively related to the level of altruistic contributions to collective efforts (12). This observation raises an evolutionary puzzle. In social dilemmas free riders who do not pay the cost of contributing but enjoy the benefits of the joint effort fare better than cooperating individuals. Punishment solves this initial dilemma by reducing the payoff of the free rider but since punishment is costly a second order social dilemma is created. Thus, punishment seems to be an altruistic act which itself needs an evolutionary explanation. This raises an evolutionary problem: In joint enterprises, free-riding individuals who do not contribute, but who exploit the efforts of others, fare better than those who pay the cost of contributing. If successful behavior spreads, for instance through imitation, these defectors will eventually take over. Punishment reduces the defectors' payoff, and thus may solve the social dilemma. However, because punishment is costly, it also reduces the punishers' payoff. This raises a "second-order social dilemma" As Colman (13) puts it, "We seem to have replaced the problem of explaining cooperation with that of explaining altruistic punishment" Thus, the emergence and maintenance of cooperation and altruism among non-kin remains a major puzzle. In fact in 2005 Science cited Evolution of Cooperation as one of the most important topics of research for the first quarter of the 21st century.

The most common formal methodological approach to tackle the problem of evolution of cooperation has been to model social dilemmas, often public good provision or common pool resource management and use evolutionary game theory to investigate population dynamics. The common practice is to first define "types" of players who always employ the same strategy and then analyze the replicator dynamics that emerge from this discrete strategy space. (14-20, 21-27) In addition many computational models, informed by the emerging replicator dynamics are used to analyze population dynamics. (19,20) Ordinarily these models (both game theoretic and computational) define types with easily defined behavioral properties. Defectors (or free riders) never contribute to a public good, cooperators altruistically contribute to the public good, punishers both contribute to the public good and punish defectors (some models also include second order punishment as well) and non-participants (or loners) do not contribute to the public good but they do not share the benefits either; they simply do not take part. Modeled as such, punishment, cooperation and defection are discrete behavioral traits: You either defect or not, you either cooperate or not, you either punish or not…

Most evolutionary game theoretic models assume a discrete (usually with a small number of pure strategies) strategy space. This choice is a tradeoff between realism and tractability. Replicator and population dynamics in a discrete strategy spaces are well defined and thoroughly studied (17). Of course there is no theoretical reason to justify that cooperation, punishment, defection or abstinence in humans are discrete traits. In fact experimental data suggests that there is substantial variation in levels of both cooperation and punishment (citation). Is there reason to believe that we might learn more about cooperation and punishment by modeling cooperativeness and punitiveness as continuous traits? In other words instead of imagining a gene that switches cooperation and punishment on and off, should we imagine a gene (or gene complex) that determine the level of cooperativeness and punitiveness?

There are several reasons which would justify such an endeavor. First of all, current studies implicitly assume that defectors have no way of responding to altruistic punishment since defectors always contribute nothing. However, literature on Machiavellian intelligence suggests that primates, but especially, humans have a unique ability to use what we call a “theory of mind”. Machiavellian intelligence is best defined by the ability of individuals to appraise the situation at hand and manipulate other individuals. Such tactics require successful detection mechanisms. In a constant influx of social dilemma like situations it is not difficult to imagine that Machiavellian intelligent individuals to adopt a level of cooperativeness that is barely enough to escape punishment. In other words defectors, who suffer punishment, may well start cooperating only enough to avoid punishment and still fare better than punishers. This is especially important in analyzing common pool resource management. Study of the tragedy of commons since being initially formulated by Hardin (34) has drawn a lot of scholarly interest. One common finding is that common pool resource management usually involves sanctions. How ultimately selfish agents adopt to sanctions is important to decipher. Selfish agents who would otherwise be free riders may learn to cooperate enough to avoid sanctions, thus limiting the effectiveness of resource management.

Second, current models fix cooperation and punishment together. Using discrete strategy spaces we are unable to analyze the co-evolution of punishment and cooperation. Modeling cooperation and punishment as continuous traits and allowing them to co-evolve in a dynamic model can be important. More specifically, we can analyze how high a level of cooperativeness can be maintained through punishment, how severe a punishment level is required to maintain cooperation, etc.

To this end I first develop an evolutionary game theoretical framework for continuous strategy spaces, define replicator dynamics and then using this framework develop a computational model to test whether modeling cooperation and punishment as continuous traits yields different results than current models.

Replicator dynamics: From discrete to continuous strategy space

We first assume an infinite population and a finite set of strategies $S=(s_1,s_2,...s_n)$. When a player A plays the strategy s_i against a player B who plays strategy s_2, then payoff to A is given by U_{ij}. When the fraction of the population that plays strategy s_i at time t is $p_i(t)$, then the average payoff to playing strategy s_i is given by:

$$\pi_i(p) = \sum_{j=1}^{N} Ui_j p_j \quad (1)$$

And the average payoff in the population is given by;

$$\bar{\pi}(p) = \sum_{i=1}^{n} \pi_i(p) p_i \text{ or equivalently , } \bar{\pi}(p) = \sum_{i,j=1}^{n} p_i U_{ij} p_j \text{ . } (2)$$

Change in the fraction p_i as the time progresses is proportional to the difference between payoff to p_i and the average payoffs. Thus, change in population fractions is given by;

$$\dot{p}_i = (\pi_i(p) - \bar{\pi}(p)) p_i \quad (3)$$

Solution set to the ordinary differential equation (3) gives the population dynamics. Now, we want to transform this discrete dynamics to continuous dynamics. In order to that we will replace the discrete strategy set S with a continuous variable s such that $s \in R^n$ and the population fractions p_i (t) by a probability distribution P(s,t). Then the payoff matrix U_{ij} is replaced by a payoff function U(s,s’) defining the payoff to s when playing against s’. Replicator dynamics then become;

$$\frac{\partial P(s,t)}{\partial t} = (\pi(s,P) - \bar{\pi}(P)) P(s,t) \quad (4)$$

Where $\pi(s,P)$ is the payoff to strategy s and $\bar{\pi}(P)$ is the total average payoff. These quantities in turn can be defined as follows:

$\pi(s,P) = \int \text{U(s,s')P(s',t)}d\text{s}'$ (5)

and

$\bar{\pi}(P) = \int \pi\text{(s,P)P(s,t)}d\text{s}$ (6)

We can show that a solution to (4) always exists as long as it is twice space differentiable and once time differentiable (see appendix).

What we have shown so far is a simple transformation of the strategy space from discrete to continuous. We can use the equation (4) to study stability conditions as well. See (citation) for similar results. But what we are interested is how we can take replicator dynamics defined in (4) and use it to formulate a computational model to investigate co-evolution of punishment and cooperation. In the next section we discuss a simple computational model that is derived from the continuous replicator dynamics defined above.

Co-evolving cooperation and punishment

We assume that there is a finite population of (N) agents who play a public good provision game is discrete time. Each agent has a cooperativeness characteristic (α)and a punishment characteristics (β). We assume that these characteristics are drawn from a bivariate probability distribution P(α,β). At each time period t, the population engages in a public good provision effort. The payoffs to each individual is given as follows:

if $\alpha_i < \Phi$

$$u_i = \sigma - \alpha_i + \Pi / N \sum \alpha_i - \beta_i - \theta * 1 / N \sum \beta_i$$

if $\alpha_i \geq \Phi$

$$u_i = \sigma - \alpha_i + \Pi / N \sum \alpha_i - \beta_i$$

Where, Φ is a set level of minimum acceptable cooperativeness, Π is the public good multiplier such that Π>1, σ is the initial endowment and θ is the punishment multiplier such that θ>0 (for efficient punishment θ>1).

This is a simplistic setting where the entire population plays the game simultaneously and together. We can readily model randomly matched smaller groups playing the game but the current configuration sets a baseline model to simulate replicator dynamics. We assume that P(α,β) is a single peaked bivariate distribution. In the following model we assume a bivariate normal distribution, mostly out of convenience.

After each round of play a new population is drawn from a distribution parameters of which is updated in the following manner:

Each agent i, influences the distribution of next generation proportional to its payoff. We calculate the proportion $u_i / \overline{u}$. The weighted α_i and β_i are calculated so that

$$\hat{\alpha}_i = u_i / \overline{u} * \alpha_i \quad \text{and} \quad \hat{\beta}_i = u_i / \overline{u} * \beta_i$$

The bivariate normal from which the next generation of agents are sampled from is then defined as having means $1/N\sum_{i=1}^{N}\hat{\alpha}_i$ and $1/N\sum_{i=1}^{N}\hat{\beta}_i$ and the population correlation of α and β is estimated from the sample correlation between α_i and β_i. Thus, after each round the new probability distribution is estimated in proportion to the payoffs. This updating mechanism approaches the replicator dynamic defined in (4) as the population size approaches infinity.

Simulations[1]

One interesting result that emerges is that even in this simplistic setting[2]; with no non-participants defined altruistic punishment is able to sustain a positive amount of cooperation. But more importantly, this sustained cooperation depends on the coevolvability of cooperation and punishment. When cooperation and punishment are allowed to co-evolve the population sustains a positive levels of cooperation, and a small punishment is sufficient to deter defectors to adopt a minimal level of cooperativeness. (Figure 1)

[Figure 1 about here]

While average payoffs stabilize in time the advantage of cooperation is not overwhelming. if there were non-participants in this setting, their average utility (which we assume is equal to the endowment of 100) would be higher than the average population utility.

[Figure 2 about here]

The issue of second order free riding is not handled here. While sustaining cooperation, albeit at relatively low levels, is interesting even higher levels of cooperation can be attained if second order punishment is included in the model.

---

[1] All simulations are coded in MATLAB. Code will be made available soon on the web.

[2] In all simulations Π =2, Φ = mean cooperativeness, θ =3. These parameters were selected to conform to existing formal and computational models as well as these parameters are often used in experimental studies.

Cooperation is sustained only if punishment and cooperation could co-evolve. When we break the link between cooperation and punishment[3] cooperation unravels eventually. This makes intuitive sense because potential irrational non-cooperative punishers would pay the double cost of punishing and being punished. Thus, punishment and cooperation cannot be independently evolving strategies. At the risk of stating the obvious, then, modeling the co-evolutionary relationship between the level of cooperation and punishment seems important.

[Figure 3]

We also observe that allowing cooperation and punishment as continuous traits also allows us to analyze the regression-to-median type of dynamic. Punishment can only sustain cooperation at the minimum acceptable levels. The simulations reported here used a threshold that is dependent on the population, i.e. only those who are less cooperative than the mean population are punished. This formulation forces the level of cooperation down. Setting fixed levels would alter the population dynamics.

And finally as a crude test of similarity to existing models, when we do not allow the punishment gene to have any phenotypic effects, i.e. we still allow the punishment gene to evolve but payoff function does not include punishment, cooperation unravels fast and the level of punishment fluctuates randomly around the initial levels. (see figure 4)

[Figure 4 about here]

Conclusion

This paper is still in its infancy, yet the initial findings are encouraging. While continuous replicator dynamics are usually cumbersome to work with, especially if we try to explicitly define the probability distributions. However, the insights from the formal models can greatly inform computational models and help evaluate results obtained from lab experiments. In this paper we have shown that continuous replicator dynamics can be derived and used to create computational models. More specifically we argued that such an endeavor is especially important for students of altruistic punishment in particular and evolution of cooperation in general. This study should be improved by additional simulations somewhat exploring the parameter space. Fixed versus floating punishment criteria will be compared in future versions of this paper as well.

---

[3] New population samples are not drawn from a bivariate distribution but from two independent normal distributions.

References

1. E. Fehr, S. Gächter, Nature 415, 137 (2002).

2. E. Fehr, U. Fischbacher, Nature 425, 785 (2003).

3. P. Hammerstein, Ed., Genetic and Cultural Evolution of Cooperation (MIT Press, Cambridge, MA, 2003).

4. M. E. Price, L. Cosmides, J. Tooby, Evol. Hum. Behav. 23, 203 (2002).

5. H. Gintis, S. Bowles, R. Boyd, E. Fehr, Eds., Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life (MIT Press, Cambridge, MA, 2005).

6. D. J.-F. de Quervain et al., Science 305, 1254 (2004).

7. C. F. Camerer, E. Fehr, Science 311, 47 (2006).

8. M. Nakamaru, Y. Iwasa, J. Theor. Biol. 240, 475 (2006).

9. R. Boyd, P. J. Richerson, Ethol. Sociobiol. 13, 171 (1992).

10. J. Henrich, R. Boyd, J. Theor. Biol. 208, 79 (2001).

11. R. Boyd, H. Gintis, S. Bowles, P. Richerson, Proc. Natl. Acad. Sci. U.S.A. 100, 3531 (2003).

12. J. Henrich et al., Science 312, 1767 (2006).

13. A. Colman, Nature 440, 744 (2006).

14. J. H. Fowler, Proc. Natl. Acad. Sci. U.S.A. 102, 7047 (2005).

15. H. Brandt, C. Hauert, K. Sigmund, Proc. Natl. Acad. Sci. U.S.A. 103, 495 (2006).

16. M. A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Nature 428, 646 (2004).

17. Weibull, J. W. (1995) Evolutionary Game Theory (MIT Press, Cambridge,MA).

18. D. Fudenberg, L. A. Imhof, J. Econ. Theory 131, 251 (2006).

19. C. Hauert, S. De Monte, J. Hofbauer, K. Sigmund, Science 296, 1129 (2002).

20. C. Hauert, S. De Monte, J. Hofbauer, K. Sigmund,J. Theor. Biol. 218, 187 (2002).

21. D. Semmann, H.-J. Krambeck, M. Milinski, Nature 425, 390 (2003).

22 (2002). 25. K. Panchanathan, R. Boyd, Nature 432,499 (2004).

26. J. H. Fowler, Nature 437, E8 (2005).

27. K. Panchanathan, R. Boyd, Nature 437, E8 (2005).

28. O. Gürerk, B. Irlenbush, B. Rockenbach, Science 312, 108 (2006).

29. E. Fehr, B. Rockenbach, Nature 422, 137 (2003).

30. C. Wedekind, M. Milinski, Science 288, 850 (2000).

31. B. Rockenbach, M. Milinski, Nature 444, 718 (2006).

32. M. Milinski, D. Semmann, H.-J. Krambeck, M. Marotzke,Proc. Natl. Acad. Sci. U.S.A. 103, 3994 (2006).

33. A. W. Johnson, T. Earle, The Evolution of Human Societies: from Foraging Group to Agrarian State (Stanford Univ. Press, Stanford, CA, 1987).

34. G. Hardin, Science 162, 1243 (1968).

35. Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. (2003) Proc. Natl. Acad. Sci. USA 100, 3531–3535.

36. Gintis, H. (2000) J. Theor. Biol. 206, 169–179.

37. Gintis, H. (2003) J. Theor. Biol. 220, 407–418.

Figure 1 Positve level of Cooperation is sustained when punishment and cooperation are allowed to be correlated. (N=1000, 3000 generations, mean α_i=50, mean β_i=15, Π=2, Φ= mean cooperativeness, θ=3)

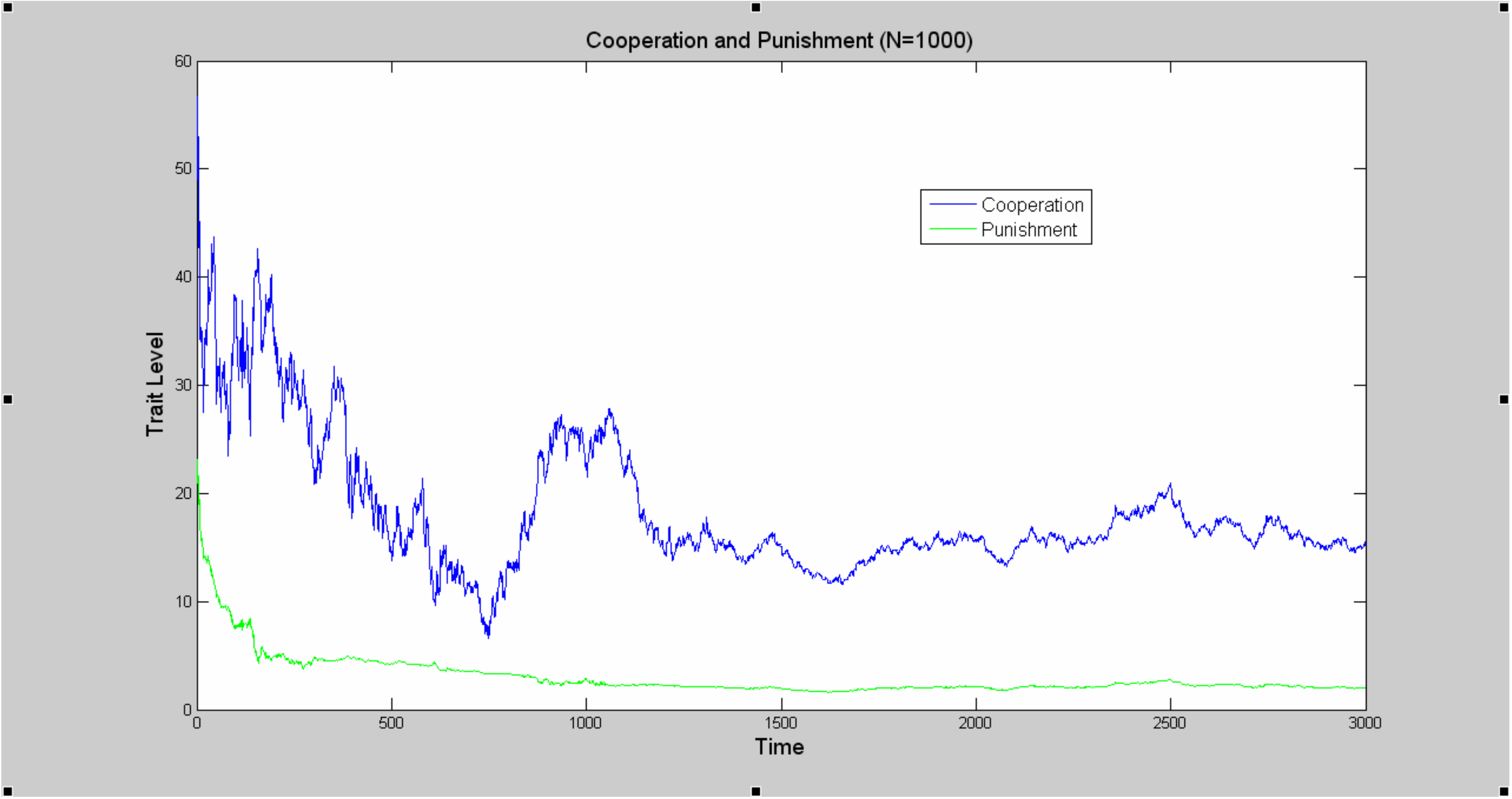

Figure 2 Average payoffs (N=1000, 3000 generations, mean α_i=50, mean β_i=15, Π=2, Φ = mean cooperativeness, θ=3)

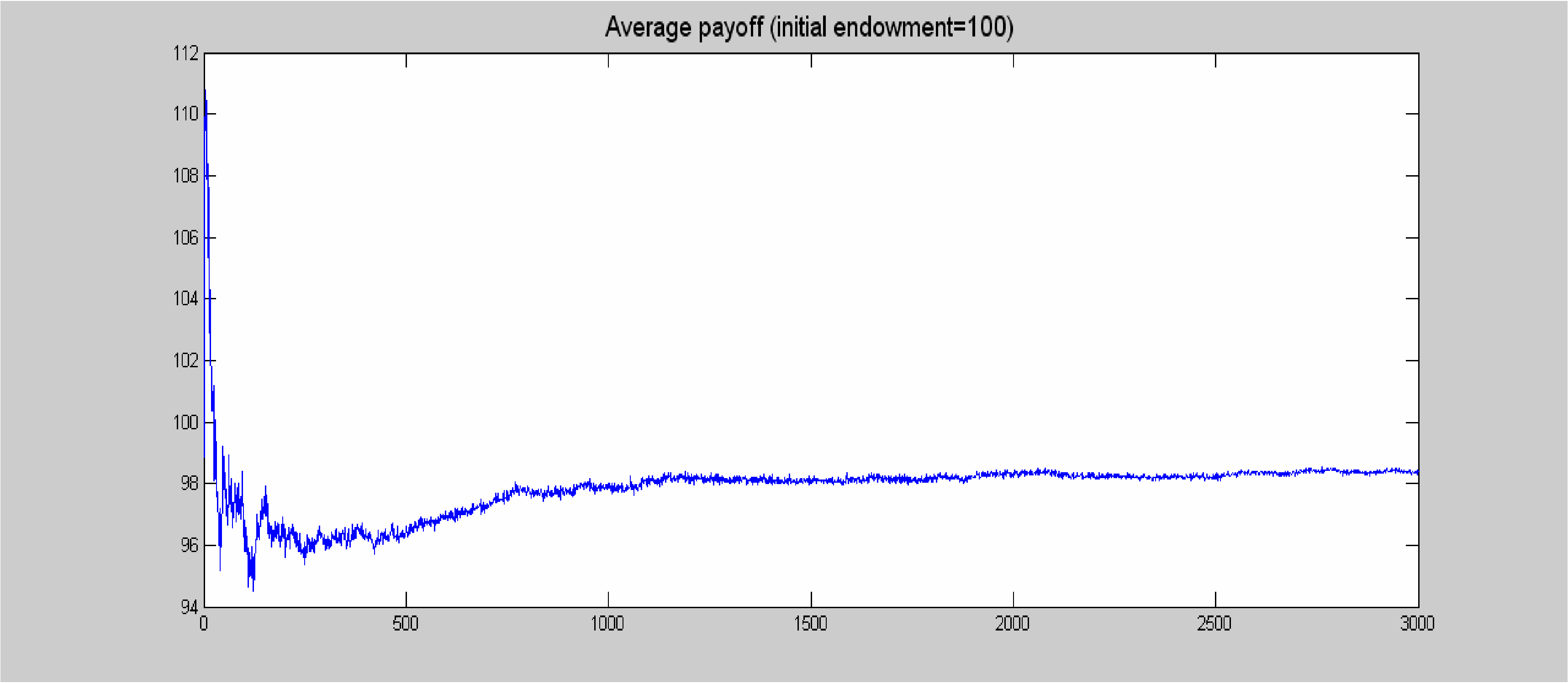

Figure 3: Cooperation unravels when cooperation and punishment is not allowed to co-evolve. (N=1000, 1000 generations, mean α_i=50, mean β_i=15, Π=2, Φ= mean cooperativeness, θ=3)

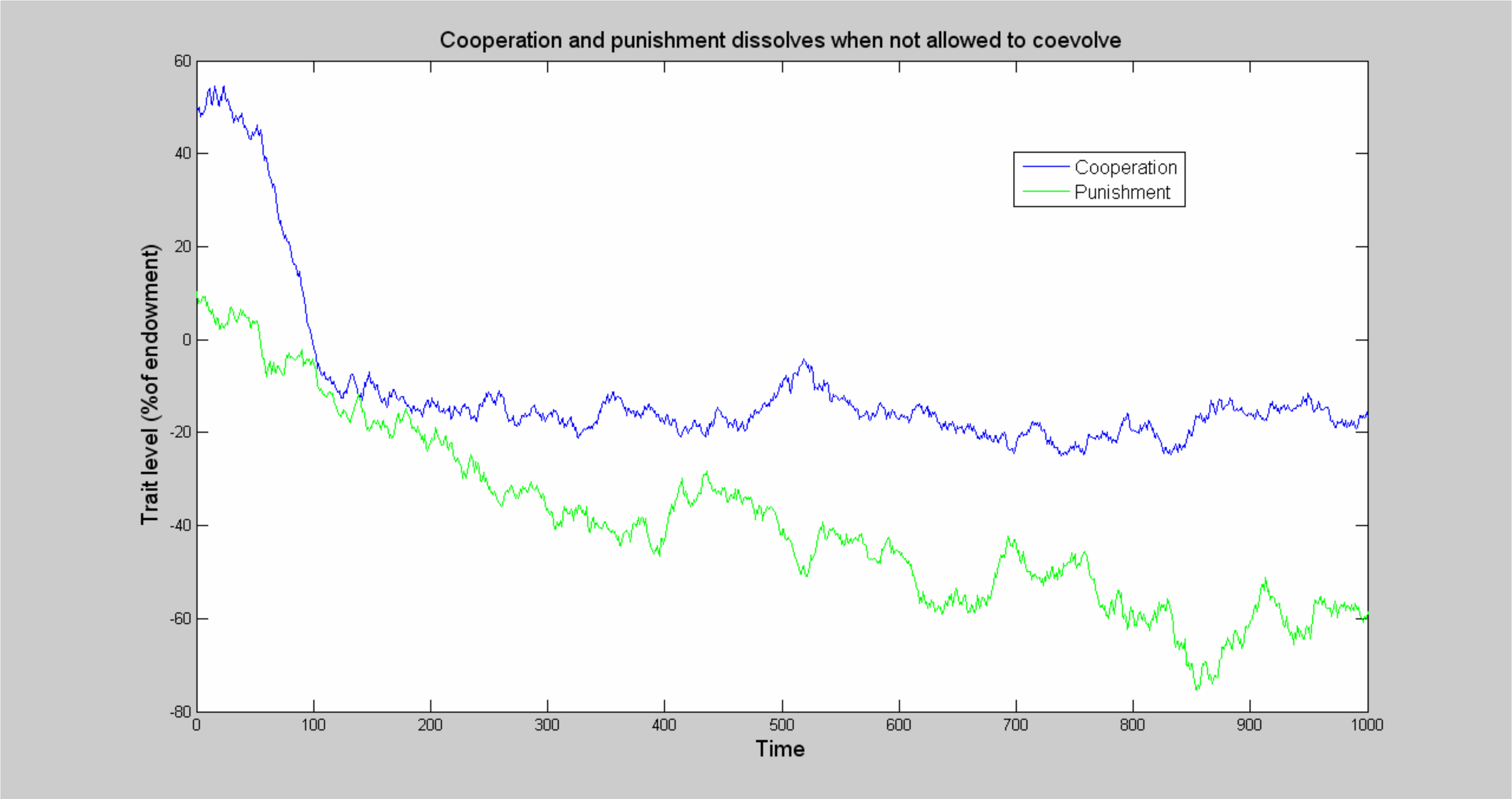

Figure 4 Cooperation unravels when punishment gene is not allowed to have phenotypic effects. (N=1000, 1000 generations, mean α_i =50, mean β_i =10, Π =2, Φ = mean cooperativeness, θ =3)