Discussion Paper, Series A, No. 2005-157

A Game-Theoretical Approach to the Formation of Ethical Norms

Kazuo Machino

December, 2005

abstract

This essay shows the formation of the ethical-norms analytically by modeling them with a Bayesian game played by bounded-rational players. They are bounded-rational in the sense that they have limited memory. The players' limited memory makes them forget other choices they had and think their relatively successful choice a convention. Deviation from it causes payoff decrease, thus, creates an incentive to penalize the deviator. Finally, reinforcement mechanism of repeating penalty makes the socially beneficial but personally costly convention an ethical norm.

1. Introduction

One of the most important unresolved problems for social scientists is the social dilemma (or the free-riders problem). However, in reality, people in a community usually manage the problems well guided by their ethical-norms. Although the ethical-norms are not usually discussed in economics with an exception of the social choice theory, they are now studied by some non-standard economists, i.e., evolutionary-game theorists and experimental and psychological economists. After refining the equilibrium concepts to the extremely sophisticated level, game-theorists have to reconsider the reality of their players who are supposed to choose ultra-rational equilibrium strategies. Some hypotheses of the alternative not-so-rational players are also supported by many experiments performed by real human players. The real players in the various game-experiments are consistently showing that they are not *homo economicus*. They are *homo reciprocans*, *homo equals*, *homo parochious*, and etc. Psychological (or behavioral) economics provides several theories for explaining such anomalies.[1]

My main research interest is to find out how people developed such ethical norms. Unlike Social Choice theory, experimental economics, or game theory where players' objectives are given axiomatically, I would like to show the formation of the ethical-norms analytically by modeling them with a non-cooperative game played by bounded-rational actors. Actors are bounded-rational in the sense that they do not know what the best action is and that they have limited memory. Although the players gradually learn whether their strategy is good or bad, after finding a good strategy they forget that they had other choices.

The word 'ethic' means 'a set of moral principles or values' or 'a theory or system of moral values' (Webster dictionary). The word 'norm' means 'a principle of right action binding upon the members of a group and serving to guide, control or regulate proper and acceptable behavior' (Webster dictionary). Both words have a meaning of 'principle(s)' that are some kinds of rules of the society. It is obvious that we, human-beings, have ethical norms since we all live in some societies. If everyone lives alone, no rule is necessary. Therefore, we can assume that these rules make members of the society choose socially beneficial but personally costly behaviors, e.g. other-oriented, egalitarian, reciprocal or parochial behaviors voluntarily or

[1] For example see Gintis (2000) Ch. 11.

involuntarily.

There is also a hint in the definitions of both words about what kind of rules they are. Each definition includes an adjective like ‘moral’ or ‘right’ that needs value judgment. Since moral values or right actions are different among cultures or religions, ethical norms must be also different among societies. Therefore, finding common features in the various ethical norms seems one natural direction of the research. That is a type of the axiomatic approach. However, my approach is to find out how people get their ethical norms. Since we already have our ethical norms through education and socialization without remembering when we got them, even if we can explain what our ethical norms are, it seems difficult to know how to get them. Furthermore, since people who taught us our ethical norms were educated or socialized by the former-generation and the former-generation was taught their norms by their former-generation, and so on, it seems impossible to find out the origins of our education and socialization.

I propose to go back to the society of primates where the structure of society was much simpler than that of ours. Recent ethological studies show that primates, which also live in a community, show many signs that they also have their behavioral norms very similar to our ethical norms. The ethologists found that the primates showed some contradicting social behaviors. They fight over the leader's position with selfish reasons (e.g., food, reproductive opportunities), and, at the same time, even the leader seems to follow their social rules against his self-oriented interest. Furthermore, when a primate who deviated from a social rule, he behaves as if he feels guilty even if no witness is there (cf. de Waal, 1996 pp 108-111). Based on the ethological researches, we can think out the following simple scenario as a process of the formation of some “ethical” norms.

(i) One member becomes the boss.
(ii) The boss “persuades” other members to take a socially beneficial, but personally costly action. Because she gets most of the benefit, she has a motive and ability for persuading them. The boss may use his/her physical force to “persuade” them.
(iii) These things may become rules in the long-run since they are socially beneficial.
(iv) Members become to feel that deviating from the rules deserves punishment since the deviation prevents the realization of their expected benefit.
(v) Since all the members now share the above feeling, a deviation makes him/her feel guilty.

Using this scenario, I would like to build my model from the scratch in the next

section. Even for the modeling from the scratch, some minimum assumptions on players are necessary in order for them to act like humans. They must have abilities to empathize in the sense that they can imagine how others feel about their behaviors and consequently react against them.[2]

2. Model

Like in the primate's community, in our society many ambitious people want to become the leader. Their motives to become the leader may or may not be selfish. In our general framework, however, their non-selfish or socially beneficial motivations are acquired by education and socialization in the society. The leader of a primitive community was likely to be the boss of the most powerful family or the strongest boss of all the families in the community. In such a community, everyone who sees an opportunity to become the leader of the community tries to get the position from instinct and the one who has the best combination of the physical strength and intelligence will get it. Thus, we can assume that the original motivation to become a leader is selfish.

We suppose that people in a primitive community live most of their lives in the same way their predecessors lived. In a few occasions members have to decide whether or not they accept their leader's new proposal. That may happen more often than usual when the community is in some crisis such as food shortage, natural disaster, or threat from outside. In such occasions the leader of the community will persuade or force the members of the community to sacrifice some of their tangible or intangible resources for the society. In other words, the leader asks the members socially beneficial but personally costly actions. Since those who have most stakes in the community will benefit the most from such an action, a leader, usually one of the most affluent members of the community, would have an incentive and resource to persuade them.

No one, including the leader, knows for sure whether the leader's new proposal is actually good for the community ex ante. If the action the leader proposed is believed to be beneficial to the community after repeating it many times, it is a "good" action. If the result is not good for the community, it is not good even if the leader's intention was good. Then, the "good" action has been internalized in the members' preferences in the long-run, and consequently they choose that action voluntarily.

---

[2] Binmore (1994) stresses the importance of empathy for explaining moral phenomena naturalistically.

We would like to model this process, where the “good” action becomes an ethical norm, in the following four steps.

Step 1: The leader asks the members (potentially) socially beneficial but personally costly actions. Members of the community decide whether or not they accept their leader’s new proposal. After the members, who accepted the leader’s proposal, carry out their task, “nature” decides whether it brings success or failure.

Step 2: After repeating the stage game many times, the incomplete information game will become an almost complete information game.

In this step, the incomplete information game in Step 1 becomes a complete information game, since if the action the leader proposes is socially beneficial, people become aware the fact as they repeat it many times.

Step 3: After repeating the “good” action game many times with limited memory, it becomes a convention (forgetting alternatives and payoff structure).

We can use Young’s model (1993) for describing the second sub-step, which we will not explain in detail in this paper.

Step 4: If the convention is socially beneficial, it becomes a norm (or may become an ethical norm).

① Find that deviating from it decreases the other players’ payoffs.
② The community penalizes the deviation.
③ After repeating the stage game many times, the deviation makes the deviator feel guilty.

2-1. Step 1

Imagine a small community with one leader and non-leader members. A leader proposes the community members to take socially beneficial but personally costly action. The leader proposes it either maybe the community faces some crisis or maybe because she just wants to try it for some reason. Since no one, including the leader, knows for sure whether the leader’s new proposal is actually good for the community *ex ante*, this situation is formalized as a following Bayesian game.

■ Players: "nature", a leader, (non-leader) members ({1, 2,…, M})

The members are arranged in order of the subjective benefit from the leader's proposal, i.e., b_i. For simplification, for $i, j \in \{1, 2,\ldots, M\}$, we assume that $b_i \neq b_j$ if $i \neq j$.

- Strategies:

  "nature": {good proposal, bad proposal} × {success, failure}

  a leader: {persuade, not persuade}

  a member: {accept, reject}

- Timing of the game & payoff

  Timing the game is as follows. It is also shown in Figure 1.

  ① "Nature" decides whether the leader's proposal is good or bad.

  ② The leader decides whether she persuades the members or not. If the leader thinks her expected payoff from persuading them ($\mathbf{E}\pi^L_{pt}$) would be greater than her expected payoff from not persuading them ($\mathbf{E}\pi^L_{nt}$), she will try to persuade them.

  ③ Each member accepts or rejects their leader's persuasion based on the comparison between his expected payoff from taking the leader's proposed action ($\mathbf{E}\pi_{at}$) and his expected payoff from rejecting it ($\mathbf{E}\pi_{nt}$).

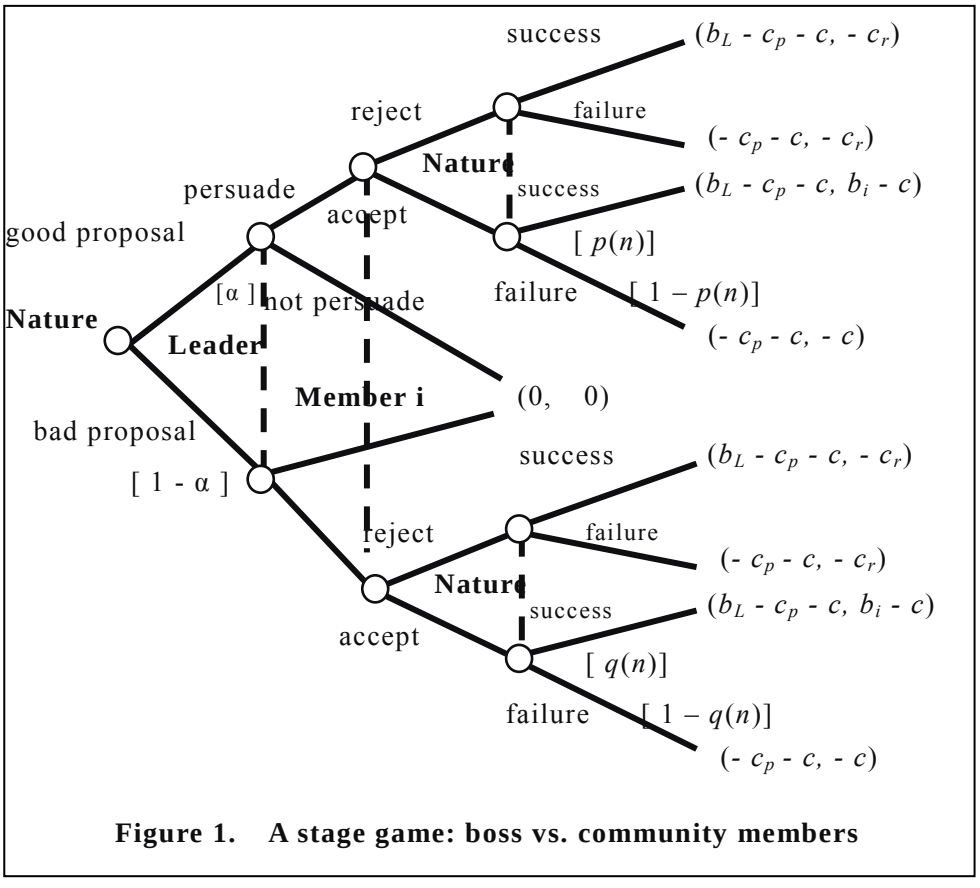

Figure 1. A stage game: boss vs. community members

④ "Nature" decides success or failure of the actions.

Denote the players' benefits and costs as follows.

b_L: leader's benefit when the action is successful

c_p: leader's persuasion cost

c: each player's cost of action

b_i: member i 's benefit when the action is successful, $c < b_i < b_L$

c_r: each member's cost of rejection

Let M and n_t be the number of (non-leader) members in the community and the number of members who accept the leader's persuasion at time t, respectively. Assume that $\mathbf{E}n_t = n_{t-1}$ and $\mathbf{E}n_1 = \alpha_1 M$, since it is difficult for a member to estimate n at time t.

Let $p(n_t)$ and $q(n_t)$ be the probability of the action's success when the proposal is good and when that is bad, respectively. Assume that they are both S-shape functions as shown in Figure 2 since an action like a potential ethical-norm could bring successful outcome only if the number of people who follow the norm is greater than a certain critical number. Assume also that $p(n_t) - p(n_t - 1) > 0$ and $q(n_t) - q(n_t - 1) > 0$. For simplification, assume $p(n_t) > q(n_t)\ \forall n_t$.

In this game, even if the proposal is successful, a member who rejects the proposal cannot get benefit. Since we suppose this game takes place in a small community, the leader or the community member can watch each other, thus, exclude the person who rejected. In this sense, the benefit from the action proposed by the leader is like a club good. For the same reason, we assume all the players including the leader have a common belief α_t about how good the proposal is at time t.

A member of this community will accept his leader's persuasion if his expectation of the benefit from the proposed action is greater than his cost of the action. His expected payoff from the proposed action at time t when he accepts the leader's persuasion is

$$\mathbf{E}\pi_{at} = \alpha_t\{p(n_{t-1})(b_i - c) - (1 - p(n_{t-1}))c\} + (1 - \alpha_t)\{q(n_{t-1})(b_i - c) - (1 - q(n_{t-1}))c\}$$
$$= \{\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1})\}b_i - c$$

where $p(0) = q(0) = 0$, $p(M) = p^*$, $q(M) = q^*$. Note that $p^* > q^*$ since $p(n) > q(n)\ \forall n$.

His expected payoff from rejecting the leader's persuasion at time t is

$$\pi_{rt} = -c_r$$

Thus, a member will accept the leader's persuasion if

$$\mathbf{E}\pi_{at} > \pi_{rt},$$

that is,

$$\{\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1})\}b_i > c - c_r \qquad (1)$$

The leader's expected payoff from persuading at time t is

$$\begin{aligned}\mathbf{E}\pi^{L}_{pt} &= \alpha_t\{p(n_{t-1})(b_L - c_p - c) - (1 - p(n_{t-1}))(c_p + c)\} \\ &\quad + (1 - \alpha_t)\{q(n_{t-1})(b_L - c_p - c) - (1 - q(n_{t-1}))(c_p + c)\} \\ &= \alpha_t(p(n_{t-1}) - q(n_{t-1}))b_L + q(n_{t-1})b_L - c_p - c,\end{aligned}$$

and her expected payoff from not persuading at time t is

$$\pi^{L}_{nt} = 0.$$

Thus, the leader will try to persuade her idea if

$$\mathbf{E}\pi^{L}_{pt} > \pi^{L}_{nt},$$

that is

$$\alpha_t(p(n_{t-1}) - q(n_{t-1}))b_L + q(n_{t-1})b_L - c_p - c > 0,$$

or

$$\{\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1})\}b_L > c + c_p \qquad (2)$$

Inequalities (1) and (2) are rewritten as follows.

$$\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1}) > (c - c_r)/b_i \qquad (1')$$

$$\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1}) > (c + c_p)/b_L \qquad (2')$$

Since $1 \geq \alpha_t \geq 0$ and $1 > p(n_{t-1}) > q(n_{t-1}) > 0$,

$$1 > \alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1}) > 0.$$

Then, if the cost of taking the action is less than the cost of rejecting the leader's persuasion, i.e., $c < c_r$, inequality (1') is always satisfied. This is the case where the leader's proposal is undoubtedly good or where the leader is so powerful that the cost for rejecting any proposal by the leader is very high. Since we would like to analyze more interesting case, we assume that $c \geq c_r$.

Define $n_d(\alpha_t)$ and $n_u(\alpha_t)$ as the minimum and the maximum numbers that satisfy inequality (1') given α_t, respectively. Also, define $n^L(\alpha_t)$ as the minimum number that satisfies inequality (2'). Since c, c_p, and b_L are given and positive, $(c + c_p)/b_L$ is a fixed positive number.

From figure 2, we can see that if $n_{t-1} \geq n_d(\alpha_t)$, the number of members who accept the leader's proposal exceeds the number of people who accepted the same proposal in the previous period, i.e., $n_t \geq n_{t-1}$. It is, however, necessary that $n_{t-1} \geq n^L(\alpha_t)$.

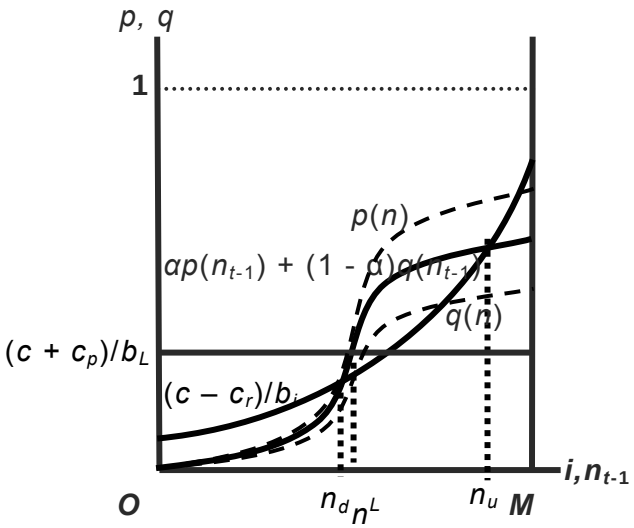

Figure 2 Inequalities (1') and (2')

Otherwise, the leader will not make the proposal.

If $n_{t-1} \geq n_u(\alpha_t)$, the number of members who accept the leader's proposal is less than the number of people who accepted the same proposal in the previous period i.e., $n_t < n_{t-1}$. However, n_t is large enough to get the support from $n_u(\alpha_t)$ members in the next period.

After the members, who accepted the leader's proposal, take the action Nature decides its success or failure.

2-2. Step 2

Next, we will take up the super-game in which the first-step game is repeated infinitely. Since we assume $p(n_{t-1})$ and $q(n_{t-1})$ are S-shape functions as shown in Figure 2, the relationship between a member's benefit when the action is successful, i.e., b_i, and his decision whether he accepts the leader's proposal is not so simple. For example, if $i > n_u(\alpha_t)$, even if most people supported it in the previous period, the member whose b_i is very small, will reject the leader's proposal. Also, if $i < n_u(\alpha_t)$, even if his b_i is very large, he will reject the proposal because the support for the project is so small that it is too risky to follow it and pay some cost.

Define $n_o(\alpha_t)$ as the maximum number that does not satisfy inequality (1') when $b_i = b_1$ given α_t. First, assume that $\alpha_t = \alpha \ \forall t$ and $n^L(\alpha) < n_o(\alpha)$, that is, players' common belief that the proposal is good is fixed for all times and that the leader always propose. Then, there are three dynamic equilibria as shown in Figure 3.

(i) If $\mathbf{E}n_\tau < n_o(\alpha)$, no one accepts the leader's persuasion at the period, i.e., $n_\tau = 0$ and consequently $n_t = 0 \ \forall t > \tau$.

(ii) If $n_o(\alpha) \leq \mathbf{E}n_\tau < n_d(\alpha)$, the number of members who accepts the leader's proposal is less than the number of people who accepted the same proposal in the previous period, i.e., $n_\tau < n_{\tau-1}$. Then, n_t will decrease every period until $n_t < n_o(\alpha)$, and $n_t = 0$ afterwards.

(iii) If $n_d(\alpha) \leq \mathbf{E}n_\tau < n_u(\alpha)$, the number of members who accepts the leader's proposal exceeds the number of people who accepted the same proposal in the previous period, i.e., $n_t \geq n_{t-1}$. Then, n_t will increase every period until $n_t \geq n_u(\alpha)$ and n_t will remain $n_u(\alpha)$ afterwards.

(iv) If $n_u(\alpha) \leq \mathbf{E}n_\tau \leq M$, the number of members who accepts the leader's proposal is $n_u(\alpha)$ that is less than the number of people who accepted the same proposal in the previous period. Then, n_t will remain $n_u(\alpha)$ afterwards.

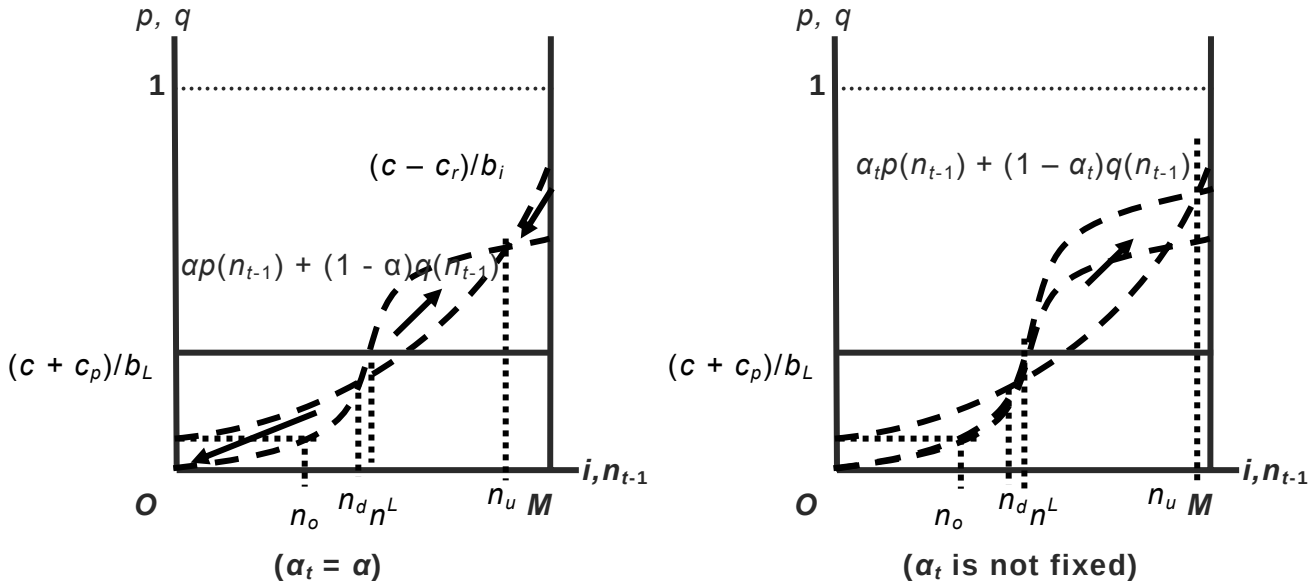

Figure 3 Dynamic Equilibria

This analysis suggests that if $\alpha_t = \alpha \ \forall t$ and if $n^L(\alpha) < n_o(\alpha)$, then, $n_t = 0$ and $n_t = n_u(\alpha)$ are the two stable equilibria. That is either no one accepts the leader's proposal or many of or all the members (if $p^* > (c - c_r)/b_M$) always accept it. If $n^L(\alpha) \geq n_o(\alpha)$, depending on the relative size of $n^L(\alpha)$ to $n_o(\alpha)$, $n_d(\alpha)$, and $n_u(\alpha)$, one or some of the above four results or the event (v) where the leader does not propose will occur.

These are the cases where players' common belief that the proposal is good at time t is constant, i.e., $\alpha_t = \alpha \ \forall t$. However, since $\alpha_{t+1} > \alpha_t$ if the social action is successful (and benefits the community) at period t and $\alpha_{t+1} < \alpha_t$ if it fails. More precisely, if the social action is successful,

$$\begin{aligned}\alpha_{t+1} &= \alpha_t p(n_{t-1})/\{\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1})\} \\ &= \alpha_t[p(n_{t-1})/\{\alpha_t p(n_{t-1}) + (1 - \alpha_t)q(n_{t-1})\}] \\ &> \alpha_t.\end{aligned}$$

If, however, it fails,

$$\begin{aligned}\alpha_{t+1} &= \alpha_t(1 - p(n_{t-1}))/\{\alpha_t(1 - p(n_{t-1})) + (1 - \alpha_t)(1 - q(n_{t-1}))\} \\ &= \alpha_t[(1 - p(n_{t-1}))/\{\alpha_t(1 - p(n_{t-1})) + (1 - \alpha_t)(1 - q(n_{t-1}))\}] \\ &< \alpha_t.\end{aligned}$$

If the social action is successful, α_{t+1} becomes greater than α_t. Thus, the left-hand side of inequalities (1') and (2') increase, and then, the number of members who accept the leader's proposal increases. As shown in Figure 3, if $\alpha_{t+1} > \alpha_t$, $n_o(\alpha_{t+1}) = n'_o < n_o(\alpha_t)$, $n_d(\alpha_{t+1}) = n'_d < n_d(\alpha_t)$, $n_u(\alpha_{t+1}) = n'_u > n_u(\alpha_t)$, and $n^L(\alpha_{t+1}) = n'^L < n^L(\alpha_t)$. Since $n_d(\alpha_{t+1}) < n_d(\alpha_t)$ and $n^L(\alpha_{t+1}) < n^L(\alpha_t)$, more people are likely to accept the

leader's proposal and more proposals are likely to be made by the leader. If α_{t+1} is less than α_t, the direction of the changes are the opposites.

Therefore, if $\mathbf{E}n_1 = \alpha_1 M \geq n_d(\alpha_1)$, $\alpha_1 M \geq n^L(\alpha_1)$, if the social action does not fail too often to keep the proposal alive, and if the proposal is "good," α_t will converge to one in the long run. In other words, people will be convinced that the leader's proposal is good for the community.

If the proposal is good and if the initial point is adequate, we can assume that $p \approx p^*$ and $q \approx q^*$ are almost constant after some period $t0$. Define s as the number of successes before period t. Then, if the proposal is good,

$$\alpha_t = \alpha_{t0}p^s(1 - p)^{t-s}/\{\alpha_{t0}p^s(1 - p)^{t-s} + (1 - \alpha_{t0})q^s(1 - q)^{t-s}\}$$
$$= 1/\{1 + (1 - \alpha_{t0})(q/p)^s\{(1 - q)/(1 - p)\}^{t-s}.$$

Then,

$$\lim_{t\to\infty} \alpha_t = \lim_{t\to\infty} 1/\{1 + (1 - \alpha_{t0})(q/p)^{pt}\{(1 - q)/(1 - p)\}^{t-pt}$$

$$= \lim_{t\to\infty} 1/\{1 + (1 - \alpha_{t0})[(q/p)^p\{(1 - q)/(1 - p)\}^{1-p}]^t.$$

Since the maximum value of

$$(q/p)^p\{(1 - q)/(1 - p)\}^{1-p}$$

is one when $p = q = 1/2$,

$$(q/p)^p\{(1 - q)/(1 - p)\}^{1-p} < 1$$

because $p > q$.

Thus,

$$\lim_{t\to\infty} [(q/p)^p\{(1 - q)/(1 - p)\}^{1-p}]^t = 0.$$

Therefore,

$$\lim_{t\to\infty} \alpha_t = 1.$$

Now, we have the following proposition.

Proposition 1: If the leader of the community has a proposal such that enough people for satisfying (1') and (2') accept, in the long-run, all the people will believe that the proposal is good for the community.

However, in this stage those who accept the leader's proposal do so because their expected benefits are greater than their costs. The action is not a norm or a convention.

2-3. Step 3

The process for the social action to become a convention is as follows. First, since α_t becomes one in the long run if the proposal is "good," inequality (1') becomes inequality (1").

$$p(n_{t-1}) > (c - c_r)/b_i \qquad (1'')$$

Since this is a realized proposal, $n_{t-1} > n^L(\alpha_t)$. Then, inequality (1") always holds. If we apply Young's model (1993) for the convention with limited memory, we can explain the process like the one shown in Figure 4. In this process players will forget the strategies they did not use more than certain periods. With this limited memory, a strategy that was not used for a long time becomes a convention. Then, since the leader does not have to "persuade" the community members, her cost becomes same as the others'.

2-4. Step 4

However, when the social action becomes convention, deviating from the convention is regarded as "bad" for the community. Since every player behaves assuming that others also behave like him, the deviator's unexpected behavior will change all the payoffs. Since this convention is socially beneficial, the payoff changes are likely to be downward. Thus, they have incentives to punish the deviator.[3]

Repeating this stage game (the upper game in Figure 4) many times is like a socialization process. By reinforcement mechanism, this stage game transforms.[4] For example, the new game looks like the one in Figure 4 (the lower one). Here, this action is the norm, i.e., 'a principle of right action binding upon the members of a group and serving to guide, control or regulate proper and acceptable behavior. (Webster)' The structure of this game tree is same as that of the previous game. However, the payoff structures in the two games are different. In the new game, deviation from the norm makes the player feel guilty. Moreover, every player is willing to pay the cost of penalizing the deviation. This new game is formalized as follows.

- Players: a leader, (non-leader) members
- Strategies:

  leader: {propose, not propose}

  a member: {(follow, punish), (follow, ignore), deviate}

[3] Knight (1992 Ch. 3) points out that deviation from an institutionalized convention, which is an equilibrium of the game, prevents other players from getting their expected payoff and that they have an incentive to penalize the deviator.

[4] Machino (2003) formalizes one type of socialization (or reinforcement) mechanism.

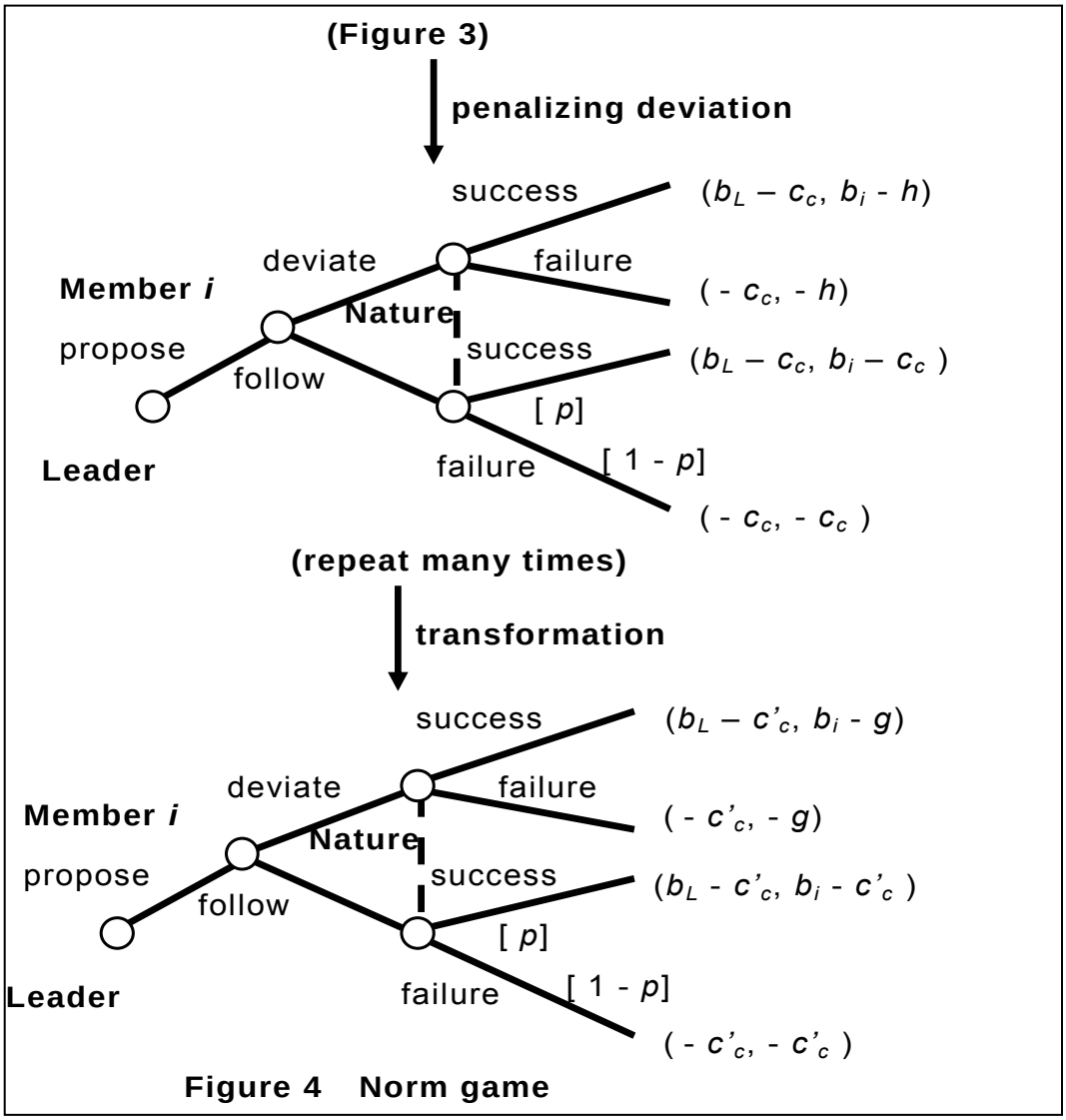

Figure 4 Norm game

■ Payoffs & game tree (Figure 4)

Definitions of all the variables are same as those in the first game except for the following three new variables.

h: penalty for the deviation

g: subjective cost (feeling guilty) for the deviation

c_c: "persuasion" costs including the cost for penalizing the deviator

c'_c: reduced "persuasion" costs (reduced because the deviator feels guilty)

3. Discussion

This paper tries to explain the formation of ethical norms as internalizing a sense of valuing socially beneficial but personally costly actions, i.e., ethical norms. I would like to show the process analytically because I believe that such an analysis is lacking

for this area of the study. Players' limited memory and reinforcement mechanism are the important points in modeling although the latter is not elaborated here. The leader's initiative is also important. The existence of the leader is explained by the heterogeneity of the community members.

As mentioned in Step 1, the community in our model is supposed to be a small primitive community. We are not sure whether we can apply our analysis to the formation of ethical norms in a big society like ours. However, many experimental and evolutionary-game studies mentioned at the begging of this paper shows that a small portion of non *homo economicus* players could proliferate in a big population. Therefore it is possible that even in a big society its small communities are the base for the formation of ethical norms.

The other important related research is how to maintain the ethical norms. Many developed societies seem to loose their work ethics along with many other ethical norms as they become affluent. I believe that our modeling framework is also useful for answering that question.

References

[1] Binmore, Ken (1994) *Game Theory and Social Contracts I*, MIT Press.

[2] Gintis, Herbert (2000) *Game Theory Evolving*, Princeton University Press

[3] de Waal, Frans (1996) *Good Natured -- The origins of Right and Wrong in Humans and Other Animals* --, Harvard University Press.

[4] Knight, Jack (1992) *Institutions and Social Conflict*, Cambridge University Press.

[5] Machino, Kazuo（2003）"*Rinriteki Kihan no Game Ron teki bunseki*" (Japanese) *Keiaigaku Kenkyu*（Hokkaido University）53 (3) 283－297.

[6] Young, Peyton, H. (1993) "The Evolution of Convention," *Econometrica*, Vol. 61, p57-84.