



On the evolutionary stability of altruistic and spiteful preferences

Alex Possajennikov

Center for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Received 25 January 1999; received in revised form 25 January 1999; accepted 2 November 1999

Abstract

This note demonstrates that a result on evolutionary stability, presented by Bester and Güth [Bester, H., Güth, W., 1998. *Journal of Economic Behavior and Organization* 34, 193–209], applies under more general preference and payoff functions. ©2000 Elsevier Science B.V. All rights reserved.

JEL classification: C72; A13

Keywords: Indirect evolution; Endogeneous preferences

1. Introduction

The usual assumption in economics is that agents care only about their own profit and not about the profit of others. In classic papers, Alchian (1950) and Friedman (1953) supported the assumption with the argument that agents who fail to maximize their profit will be eliminated by economic evolution. Recently, however, Bester and Güth (1998) showed that, in a strategic interaction context with complete information, it is not necessarily true. They showed that altruism may be favored by evolution. The set of preferences, however, was restricted in an essential manner. In this note I will show that the Bester and Güth analysis extends to much broader preference parameter regions, and that the result extends to spiteful preferences as well as altruism. I also present an example when the sign of the strategic interaction (complements versus substitutes in the material payoff function) does not dictate the sign of preferences favored by evolution (altruism versus spite). Independently, Bolle (2000) also explains how the Bester and Güth results can be extended to allow spite.

2. The model

In the Bester and Güth model, two players play a symmetric game with ‘material payoffs’ given by

$$U_1(x, y) = x(ky + m - x), \quad U_2(x, y) = y(kx + m - y) \quad (1)$$

where $x \geq 0$ and $y \geq 0$ are the strategies of Player 1 and Player 2, respectively, and where k and $m > 0$ are parameters. These payoff functions would, for example, be profit functions in a differentiated products oligopoly game where strategies are quantities supplied, demand is linear, and costs are zero. The parameter k allows the products to be complements ($k > 0$) or substitutes ($k < 0$) in demand. Bester and Güth assume that the players do not necessarily maximize their material payoffs, but rather weighted sums of own and opponent’s payoffs:

$$V_1(x, y) = U_1(x, y) + \alpha U_2(x, y), \quad V_2(x, y) = U_2(x, y) + \beta U_1(x, y). \quad (2)$$

Here α and β are preference parameters, positive under altruism, negative under spite, and zero under classical own profit maximization.

Assuming α and β are subject to evolutionary pressure, what values would evolve? In particular, Bester and Güth asked what values for α and β are evolutionarily stable? Under the assumptions that $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $-1 < k < 1$ ($k \neq 0$), they found $\alpha = \beta = (k/(2 - k))$ to be the unique evolutionarily stable outcome when $k > 0$ and $\alpha = \beta = 0$ to be the unique evolutionarily stable outcome when $k < 0$. This result implies that the sign of the strategic interaction in the material payoff function (the sign of k) determines the sign of the evolutionarily stable preference parameters α and β .

Bolle showed that relaxing the restrictions on the preference parameters to $-\infty < \alpha \leq 1$ and $-\infty < \beta \leq 1$, thus allowing spite, leads to $\alpha = \beta = (k/(2 - k))$ to be the unique evolutionarily stable outcome, independently of the sign of k . Under his result, the sign of the strategic interaction in the material payoff function (the sign of k) still dictates the sign of the evolutionarily stable α and β .

In the proposition below, all of the restrictions on the preference parameters (α, β) will be relaxed. I will also remove the restrictions on k and (x, y) . It will be shown that $\alpha = \beta = (k/(2 - k))$ is still the evolutionarily stable outcome when one exists, and that one does exist so long as $-2 \leq k < 1$ ($k \neq 0$) or $k > 2$. But first I will argue why parameters might lie outside the bounds assumed by Bester and Güth.

Bolle has already explained why arbitrarily large negative values of α and β should be allowed. If spite is to be ruled out, it should be ruled out by the evolutionary logic, not by assumption. However, the same argument applies to large positive values of α and β . When α and β exceed 1, it merely means that agents care more about their opponents than about themselves, a coherent logical possibility.

Nevertheless, when $\alpha, \beta \in \mathbb{R}$ there are two reasons why the approach of Bester and Güth does not generalize automatically. The first reason is that if one keeps the restrictions $x \geq 0$, $y \geq 0$ then corner solutions often appear. This would require further assumptions on the cases with no or more than one equilibria than the ones below. Doing this would make the analysis more complicated without changing the qualitative results. Therefore, for the ease of the exposition, I consider as strategy space the whole real line \mathbb{R} . One of

the economic interpretations for such a strategy space can be an oligopoly model when a firm may buy back existing stocks from the market. Other interpretations are also possible. In a bidding context, an agent may demand a subsidy as well as offer a payment. In an effort context, a worker might be motivated to sabotage. In an international duopoly, the two governments might impose either tariffs or subsidies. In an investment game, an agent disinvest as well as invest.

The second problem with the extension of the preference parameter space is that there may be no equilibrium of the game with given preference parameters, or there may be a continuum of equilibria. In such cases a possible approach is to extend the fitness function $U_1^*(\alpha, \beta)$, defined later, to such preference parameters by continuity in the first argument. Continuity in the first argument is chosen because it avoids unnecessary technical difficulties since an evolutionarily stable preference parameter is found by maximizing the fitness function with respect to the first argument.

While the parameters α and β are determined endogenously, and therefore should not be restricted a priori, parameter k is exogenous. In most economic application it makes sense to have $-1 < k < 1$. In the differentiated product oligopoly interpretation above, $-1 < k < 1$ would mean that own price has more effect on own demand than the other price. Nonetheless, well behaved demand functions may lie outside the restriction. For example, in an international duopoly context, the output of the firm in the other country may well be more important to price than home output. As another example, a player in an effort game may perform a task that is more important to the other player. The case $k > 2$ is especially interesting because then, under the evolutionarily stable outcome $\alpha = \beta = (k/(2 - k))$, the sign of α and β are opposite to the sign of k .

The following proposition holds.

Proposition. *If $-2 \leq k < 1$ ($k \neq 0$) or $k > 2$, then $\alpha = \beta = (k/(2 - k))$ is the unique evolutionarily stable value for the preference parameters. If $k < -2$, $k = 0$ or $1 \leq k \leq 2$, there is no evolutionarily stable values for α and β .*

Proof. The players maximize their subjective utility functions (Eq. (2)). The unique equilibrium of the game is given by

$$x^*(\alpha, \beta) = \frac{m(k(\alpha + 1) + 2)}{4 - k^2(\alpha + 1)(\beta + 1)}, \quad y^*(\alpha, \beta) = \frac{m(k(\beta + 1) + 2)}{4 - k^2(\alpha + 1)(\beta + 1)} \quad (3)$$

if $4 - k^2(\alpha + 1)(\beta + 1) \neq 0$. The material payoff (*fitness*) of Player 1 as a function of preference parameters through the implied equilibrium strategies of both players is

$$U_1^*(\alpha, \beta) = -\frac{m^2(k(\alpha + 1) + 2)(k^2\alpha(\beta + 1) + k(\alpha - 1) - 2)}{(4 - k^2(\alpha + 1)(\beta + 1))^2} \quad (4)$$

while the material payoff function of Player 2 satisfies $U_2^*(\beta, \alpha) = U_1^*(\alpha, \beta)$. It follows from the equilibrium strategies in Eq. (3) that the game with preference parameters α', β' does not have a unique equilibrium if $4 - k^2(\alpha' + 1)(\beta' + 1) = 0$. As argued above, for such α', β' the fitness function is extended by continuity in the first argument as $U_1^*(\alpha', \beta') = \lim_{\alpha \rightarrow \alpha'} \lim_{\beta \rightarrow \beta'} U_1^*(\alpha, \beta)$. This limit always exists on the extended real line $\mathbb{R} \cup \{\pm\infty\}$. With this extension, the function $U_1^*(\alpha, \beta)$ is differentiable with respect to the first argument at $\beta = \alpha$.

An evolutionarily stable preference parameter in a symmetric context is a parameter α^* satisfying

$$(a) U_1^*(\alpha^*, \alpha^*) \geq U_1^*(\alpha, \alpha^*) \quad \forall \alpha,$$

$$(b) \text{ if } U_1^*(\alpha^*, \alpha^*) = U_1^*(\alpha, \alpha^*) \text{ for } \alpha \neq \alpha^*, \text{ then } U_1^*(\alpha^*, \alpha) > U_1^*(\alpha, \alpha).$$

To check condition (a), fix the second argument of $U_1^*(\alpha, \beta)$, find maxima of $U_1^*(\alpha, \beta)$ with respect to the first argument and equate the arguments. The first-order condition is

$$\alpha = -\frac{k(\beta + 1)(k + 2)}{\beta k(k - 2) + k^2 - 2k - 4}. \quad (5)$$

After equating β to α , possible candidates for evolutionarily stable parameters are

$$\alpha_1^* = -\frac{k + 2}{k}, \quad \alpha_2^* = \frac{k}{2 - k}. \quad (6)$$

Consider $\alpha_1^* = -(k + 2)/k$. With the above extension of the fitness function $U_1^*(\alpha, -(k + 2)/k) = m^2/4 \quad \forall \alpha$. Condition (a) for evolutionary stability is satisfied with equality for any α . However, $U_1^*(-(k + 2)/k, 0) = 0 < U_1^*(0, 0) = (m^2/(k - 2)^2)$, thus condition (b) is not satisfied and $\alpha_1^* = -(k + 2)/k$ is not evolutionarily stable.

To prove that α_2^* is the unique best preference parameter against itself, consider

$$\begin{aligned} & U_1^*\left(\frac{k}{2 - k}, \frac{k}{2 - k}\right) - U_1^*\left(\alpha, \frac{k}{2 - k}\right) \\ &= \frac{m^2(k + 2)(k - 2)}{16(k - 1)} - \frac{m^2(k + 2)(k - 2)(k^2(\alpha + 1)^2 - 4)}{4(k^2(\alpha + 1) + 2k - 4)^2} \\ &= \frac{m^2(k + 2)(k - 2)k^2(k(\alpha + 1) - 2\alpha)^2}{16(k - 1)(k^2(\alpha + 1) + 2k - 4)^2}. \end{aligned}$$

The last expression is strictly positive for $-2 < k < 1$ ($k \neq 0$) and for $k > 2$ unless $\alpha = \alpha_2^*$ and it is negative for $k < -2$ and for $1 < k < 2$. Thus, α_2^* is evolutionarily stable when $-2 < k < 1$ ($k \neq 0$) or $k > 2$ and it is not evolutionarily stable for $k < -2$ and for $1 < k < 2$. For boundaries one can check that α_2^* is evolutionarily stable for $k = -2$ but not for $k = 1$ or $k = 2$. When $k = 0$, there is no strategic interaction and players with all preferences parameters have the same fitness. Therefore, no preference parameter is evolutionarily stable when $k = 0$. \square

For $0 < k < 1$ the result is the same as in Bester and Güth: some degree of altruism is evolutionarily stable, $\alpha^* > 0$. For $-2 \leq k < 0$ the result is a natural extension of their result, indicating that the evolutionarily stable preference parameter has some degree of spite ($\alpha^* < 0$). An interesting new result appears when $k > 2$. Then evolutionarily stable α^* is negative and larger than one in absolute value. Thus, a large degree of spite, up to minimizing opponent's payoff ($\alpha^* \rightarrow -\infty$ when $k \rightarrow 2$ from above) is evolutionarily stable when the degree of interdependence between players' strategies is high, though it is a positive interdependence! This example shows that complementarity (positive strategic interaction) in the material payoff function may lead to negative evolutionarily stable preference parameter.

3. Conclusion

This note, together with the note of Bolle, demonstrates that the Bester and Güth result applies under a much broader range of conditions than they discussed. Both spite and altruism may be evolutionarily stable when there is strategic interaction.

I agree with Professor Bolle that, though the results look rather gloomy, predicting spiteful preferences to be evolutionarily stable rather often, they should be taken as they are. If one admits the existence of altruism, one should also admit the existence of spite or envy. However, it can be shown that with incomplete information, i.e., when the preferences of the opponent are not known, only egoistic preferences (or preferences equivalent to them) survive evolution. This result holds for quite general models of indirect evolution (see Ely and Yilankaya, 1997). Thus, uncertainty would eliminate spiteful preferences though it would eliminate altruistic ones too.

Acknowledgements

This paper is a part of my Ph.D. thesis. I am grateful to Maria Montero, Eric van Damme, Fred Palomino, Friedel Bolle, and John Conlisk for useful comments. A part of the work was done during my visit to University College London that was made possible by a grant from Dutch Organization for Scientific Research (NWO).

References

- Alchian, A., 1950. Uncertainty, evolution and economic theory. *Journal of Political Economy* 58, 211–221.
- Bester, H., Güth, W., 1998. Is altruism evolutionarily stable? *Journal of Economic Behavior and Organization* 34, 193–209.
- Bolle, F., 2000. Is altruism evolutionarily stable? And envy and malevolence? *Journal of Economic Behavior and Organization* (in press).
- Ely, J., Yilankaya, O., 1997. Nash Equilibrium and the Evolution of Preferences. Discussion Paper No. 1191, Northwestern University.
- Friedman, M., 1953. *The Methodology of Positive Economics, Essays in Positive Economics*. University of Chicago Press, Chicago.