



Is altruism bad for cooperation?☆

Sung-Ha Hwang^a, Samuel Bowles^{b,c,*}

^a School of Economics, Sogang University, Republic of Korea

^b Santa Fe Institute, USA

^c University of Siena, Italy

ARTICLE INFO

Article history:

Received 11 July 2011

Received in revised form 29 May 2012

Accepted 20 June 2012

Available online 28 June 2012

JEL classification:

D64 (altruism)

H41 (public goods)

D03 (behavioral economics)

Keywords:

Public goods

Altruism

Spite

Reciprocity

Punishment

Cooperation

ABSTRACT

Some philosophers and social scientists have stressed the importance for good government of an altruistic citizenry that values the well being of fellow citizens. Economists, however, have emphasized the need for incentives that induce even the self-interested to contribute to the public good. Implicitly most have assumed that these two approaches are complementary or at worst additive. But this need not be the case. Behavioral experiments find that if reciprocity-minded subjects feel hostility towards free riders and enjoy inflicting harm on them, the incentives provided by the anticipated punishment support near efficient levels of contributions to a public good. Cooperation may also be supported if altruistic individuals internalize the group benefits that their contributions produce. But the effects of these two supports for high levels of cooperation may be less than additive. Using a utility function embodying both reciprocity and altruism we show that unconditional altruism attenuates the punishment motive and thus may reduce the level of punishment inflicted on defectors, resulting in lower levels of contribution. Increases in altruism may also reduce the level of benefits from the public project net of contribution costs and punishment costs. The range over which altruism inhibits cooperation and reduces material payoffs is greater, the stronger is the reciprocity motive among group members.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

When Adam Smith famously proposed that the actions of the self interested economic man might implement a socially desirable allocation of resources he added “Nor is it always worse for the society that . . . he intends only his own gain” (Smith, 1776). Smith was aware, of course, that in public goods settings and other social dilemmas, a concern for the well being of others may improve allocational efficiency. But economists have also elaborated the dark side of other-regarding preferences, including the way that inequity averse preferences may result in a smaller joint surplus, and hostility towards outsiders may restrict opportunities for mutually beneficial exchange. Simple, unconditional altruism, however, seems an unlikely candidate as the culprit in such deviations from efficient allocation. But we will show that altruism may reduce contributions to a public good, resulting in a smaller joint surplus than otherwise would be available.

Both altruism and reciprocity may motivate individuals to contribute to the provision of a public good. Altruism induces the individual to unconditionally value the payoff of other individuals, while reciprocity implies a valuation of the others' payoffs that is conditional on their contributions (or other indications of their type). Reciprocators may value the payoffs of

☆ We thank the Behavioral Science Program of the Santa Fe Institute, the U.S. National Science Foundation, the University of Siena and the European Science Foundation for support of this project and Roland Benabou, Theodore Bergstrom, Drew Fudenberg, Simon Gaechter, David Levine, Louis Putterman, Rajiv Sethi, Joel Sobel, Elisabeth Wood, and two anonymous referees for helpful comments.

* Corresponding author at: Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, USA.

E-mail addresses: sunghah@sogang.ac.kr (S.-H. Hwang), bowles@santafe.edu (S. Bowles).

low contributors negatively and be motivated to reduce the payoffs of defectors at a cost to themselves, when this option is available. The prospect of punishment of low contributions may induce individuals to contribute more than they otherwise would, thereby sustaining cooperation in groups where formal constraints and incentives are insufficient (Fehr and Gaechter, 2000; Anderson and Putterman, 2006; Boyd et al., 2010).

We explore the possibility that the two motives for contribution – a positive valuation of the payoffs of others and a desire to avoid the punishment induced by a negative valuation of one's payoffs by others – may work at cross purposes. Specifically we show that by attenuating the punishment motive, a general increase in the level of unconditional altruism may reduce rather than increase contributions.

Thus, while one often refers to individuals as being 'cooperative' or 'uncooperative', the motives supporting high levels of cooperation are heterogeneous, and they need not work synergistically. For example, experimental evidence indicates that unconditional altruists among American student subjects are significantly less likely to punish low contributors in a public goods game (Carpenter et al., 2009): a standard deviation increase in an individual's level of altruism reduced the amount he spent on punishment by 0.13 standard deviations. Also consistent with a possible conflict between altruistic motives for contributing to a public good and a willingness to discipline free riders is the fact that, among Russian urban and rural adults, those who contributed more to the public good punished low contributors significantly less, conditional on (a) the contribution level of others, (b) the amount by which the target of the punishment contributed less than the punisher, and (c) a large number of demographic and occupational controls (Gaechter and Hermann, 2011).

We model a public goods problem in which voluntary contributions are sustained both because altruistic citizens value the benefits conferred on others and reciprocal citizens punish free riders. We show that an increase in the level of altruism in a population may reduce the level of contributions and the benefits of the public project. This occurs because more altruistic individuals, while predisposed to contribute, are less willing to punish free riders. The idea behind the result – that seemingly good motives need not be synergistic – is as ancient as the contrast between the retribution-based morality typically attributed to the Old Testament and the unconditional generosity advocated in parts of the New Testament.

A key assumption in our model is that people have an intrinsic motivation to punish shirkers, not simply an instrumental desire to alter their behavior or to affect the distribution of payoffs so as to either reduce unfairness or to enhance the punisher's own relative payoffs. This is similar to what Boyd and Richerson (1992) call retribution punishment and the analogue of Andreoni's (1990) warm glow altruism. That subjects view punishment of shirkers also as retribution rather than simply as instrumental towards affecting behavior is consistent with the recent public goods with punishment experiment of Falk et al. (2005). The game was one shot, ruling out behavior modification as a motive for punishing low contributors, and the punishment technology was such that punishment could not alter the difference in payoff between the punisher and the target (the cost to the punisher was the same as that inflicted on the target). Nonetheless, sixty per cent of cooperators punished defectors.

Further evidence for our assumption that punishment is non-strategic comes from the public goods experiment of Fudenberg and Pathak (2010). As in the standard game, following each round of contributions subjects were given information on the contributions of fellow group members and had the opportunity to deduct some of their own payoffs in order to lower the payoffs of another in the group. But unlike the usual treatment in which the targets of punishment are informed of the level of punishment received after each round, in the Fudenberg and Pathak experiment the levels of punishment were not to be revealed until the experiment was over, and those who punished others knew this. Thus the experimental design ruled out modifying the behavior of shirkers as a motive for punishment. Consistent with what the authors term a "pure preference" motivation for punishment, subjects nonetheless punished shirkers, leading the authors to conclude that "agents enjoy punishment, where 'enjoyment' includes anger and a desire for retribution." There is considerable further evidence for our non-strategic modeling of punishment (de Quervain et al., 2004; Casari and Luini, 2012; Gaechter and Hermann, 2011; Anderson and Putterman, 2006).

In the next section we use the ideas of Levine (1998), Rabin (1993), and Falk and Fischbacher (2006) to explore the joint effects of altruism toward fellow group members and reciprocity-based hostility towards low contributors in a public goods game. In Section 3 we study the Nash equilibrium levels of punishment and contribution under varying levels of unconditional altruism of the members of a group. We show first that the relationship between the level of altruism and contributions is non-monotonic, and that there exists a range of levels of altruism over which increases in altruism reduce both equilibrium levels of contribution and the sum of benefits from the public project, net of the costs of contributing and the costs of punishing. Second, we show that the range for which altruism is bad for both cooperation and net benefits is larger the more reciprocal are the group members. For simplicity of exposition and clarity of the underlying causal mechanisms we initially assume a homogeneous population. In Section 4 we extend this model to a heterogeneous population and show that our main results and key insights still hold. In heterogeneous populations we can also show that the greater the frequency of altruistic reciprocators in the population the wider is the range for which increased levels of altruism in their functions will decrease average contributions. In the penultimate section we consider a number of caveats and possible extensions. In the conclusion we suggest some implications for how social preferences may support cooperation despite the sometimes counterproductive effects of increased altruism and the costly nature of punishment. In Appendix A we present the proofs of Propositions 1 and 2.

Similar in spirit to our first result is the finding of Bernheim and Stark (1988) that increased altruism among two family members in a repeated game setting may be welfare-reducing (see also Nakao, 2008; Alger and Weibull, 2010). However, our setting is a non-repeated public goods game rather than a repeated dyadic interaction; and rather than the simple

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات