November, 2012

Ethics and the Volunteers' Dilemma

Ted C Bergstrom, *University of California, Santa Barbara*

Ethics and the Volunteer's Dilemma

Ted Bergstrom, et al

November 5, 2012

Driving home on a well-travelled street, you encounter a broken traffic signal or perhaps a pile of traffic-obstructing debris. You wonder whether to take the trouble to phone traffic authorities about this condition. You realize that many other commuters face the same choice. If someone else calls, your effort will be wasted, but if everybody believes that someone else will call, the hazard will go unreported.

1 Volunteer's Dilemma with Identical Players

Andreas Diekmann modeled this situation as an n-player symmetric game which he called the *Volunteer's Dilemma* [7]. In the Volunteer's Dilemma game, each player can choose to take action or not. If at least one person takes action, then all of those who take action must pay a cost c and will receive net benefits of $b - c > 0$, while those who do not take action enjoy the benefits, but do not pay the cost, and thus receive net benefits of b. If no player takes action, then all players get a payoff of 0.

1.1 Symmetric Nash Equilibrium

In the Volunteer's Dilemma with two or more players, there cannot be a symmetric Nash equilibrium in which all take action, since if anyone else acts, one's own best response is not to act. Nor can there be a symmetric Nash equilibrium in which none take action, since if nobody else acts, one's own best response is to act. The only symmetric Nash equilibrium for this game is one in which each player uses a mixed strategy; taking action with a positive probability less than 1.

In a mixed strategy equilibrium, each player is indifferent between taking action and not doing so. Anyone who takes action is certain to have a net payoff of $b-c$. In equilibrium, all players must be indifferent between taking action and not taking action. Therefore, regardless of the number of players, the expected utility of each player in a symmetric Nash equilibrium must be $b - c$.

In a mixed strategy equilibrium for n players who each take action with independent probability p_n, a player who chooses the strategy "do not act" will not pay any cost and will enjoy the benefit b if at least one other player takes action. Let $q_n = 1 - p_n$. If all other players take action with probability p_n then the probability that at least one of the others takes action is $1 - q_n^{n-1}$. Therefore the expected payoff from the strategy "do not act" is $b\left(1 - q_n^{n-1}\right)$. It follows that in equilibrium,

$$b\left(1 - q_n^{n-1}\right) = b - c \tag{1}$$

and hence

$$q_n = \left(\frac{c}{b}\right)^{\frac{1}{n-1}}. \tag{2}$$

From Equation 2 it follows that in symmetric Nash equilibrium for an n player Volunteer's Dilemma, the probability that no player takes action is

$$q_n^n = \left(\frac{c}{b}\right)^{\frac{n}{n-1}} \tag{3}$$

which is an increasing function of n and which approaches c/b in the limit as n gets large.

This leaves us with a vexing conundrum. The technology of the Volunteer's Dilemma game allows the potential for significant benefits from the formation of larger groups; an action taken by a single person is sufficient to benefit the entire group, no matter how large the group. Yet, in symmetric Nash equilibrium for this game, as the number of players increases, the probability that nobody takes action increases and the expected payoff to each player remains constant at $b - c$.

1.2 Efficiency and Optimal Symmetric Mixed Strategy

Inefficiency of symmetric Nash equilibrium in the Volunteer's Dilemma game arises from two sources. One is the standard problem of neglected externalities. Individuals ignore the fact that an increase in their own probability of taking action exerts a positive externality on the expected payoffs of all other players. The second source of inefficiency is a coordination problem. In mixed strategy equilibrium, players do not know the actions that have been taken by others. Thus, in equilibrium, there is a positive probability that more than one player takes costly action, although the action of only one is needed to produce benefits for all.

Sometimes it is possible to coordinate the actions of players so that if there is more than one volunteer, only a single volunteer will be selected to perform the task. For example, potential donors of stem cells from bone marrow or blood aphoresis join a registry of persons who have declared their willingness to donate if their contributions are needed. When a patient is in need of a transplant, if one or more potential donors of this patient's immunity type have volunteered, the registry selects exactly one of these volunteers to make the donation. [3]. Jeroen Weesie [11] and Theodore Bergstrom [2] analyze the comparative statics of Nash equilibrium for versions of Volunteer's Dilemma in which at most one of the volunteers is required to pay.

Sometimes duplication of effort can be avoided because potential volunteers can see immediately whether someone else has "beat them to it." Bergstrom [1] studies the case of passers-by on a more or less crowded highway, who are presented sequentially with the opportunity to help a distressed traveler. Christopher Bliss and Barry Nalebuff [6], Marc Bilodeau and Al Slivinski [4] and Weesie [10] analyze a war-of-attrition game in which the first person to take action is observed by all and where benefits diminish as time passes. In deciding when to act, players face a trade-off between the costs of postponement and the possibility that if one waits a little longer, action will be unnecessary because someone else will have done it.

This paper studies situations where such coordination is technically infeasible. In the example discussed at the beginning of this paper, the cost of informing authorities would be minimized if only one commuter took action. But how can this be accomplished? It would not be cost-effective for the commuters who notice the problem to assemble and choose one of their number to contact the authorities.

1.3 An appeal to ethics

If players could be persuaded to abide by a self-enforced ethical rule, they would all be better off than they are in the symmetric Nash equilibrium. In the absence of a coordinating device, it is not possible to avoid duplication of effort when when the efforts of only one are needed. However, as we will see, even without coordination, there is a symmetric ethical rule that would lead players, using independent strategies, to improve on the symmetric Nash equilibrium. To find this optimal symmetric rule, we seek a strategy that satisfies the *Kantian* principle: "Use the strategy that you would wish that everyone would use".

Proposition 1. *In an n-player Volunteer's Dilemma, there is an optimal symmetric rule that requires each player to use a mixed strategy in which the the probability that a player takes action is greater than that in symmetric Nash equilibrium. If the probability that an individual does not take action is x_n under the optimal symmetric rule and q_n in Nash equilibrium, then*

$$x_n = n^{\frac{-1}{n-1}} q_n.$$

Proof. Where the mandated strategy is of the form: "take action with probability $1-x$", the probability that at least one player takes action is $1-x^n$, and the expected cost to each player of following the strategy is $c(1-x)$. The expected utility of every player is

$$b(1-x^n) - c(1-x). \tag{4}$$

Taking the derivative of expression 4, and arranging terms, we see that expected utility is maximized at $x = x_n$, when

$$x_n = n^{\frac{-1}{n-1}} \left(\frac{c}{b}\right)^{\frac{1}{n-1}}. \tag{5}$$

From Equations 5 and Equation 2, it follows that

$$x_n = n^{\frac{-1}{n-1}} q_n. \tag{6}$$

Since $n^{\frac{-1}{n-1}} < 1$ for all $n > 1$, it must be that $0 < x_n < q_n$ and hence the $1 > 1 - x_n > 1 - q_n$, which means that the probability of that an individual takes action under the optimal symmetric rule is less than one, but greater than the probability of taking action in Nash equilibrium. □

Proposition 2. *In the limit as the number of players approaches infinity, under the optimal symmetric rule, the probability that any single individual takes action approaches zero, but the probability that at least one player takes action approaches one.*

Proof. Equation 5 implies that

$$\begin{aligned}\lim_{n\to\infty} \ln x_n &= \lim_{n\to\infty}\left(\frac{-1}{n-1}\right)\ln n + \lim_{n\to\infty}\left(\frac{1}{n-1}\right)\frac{c}{b} \\ &= \lim_{n\to\infty}\left(\frac{-\ln n}{n-1}\right) \\ &= 0. \end{aligned} \tag{7}$$

It also follows form Equation 5 that

$$\begin{aligned}\lim_{n\to\infty} \ln x_n^n &= \lim_{n\to\infty}\left(\frac{-n}{n-1}\right)\ln n + \lim_{n\to\infty}\left(\frac{n}{n-1}\right)\frac{c}{b} \\ &= \lim_{n\to\infty}\left(\frac{-n\ln n}{n-1}\right) + \frac{c}{b} \\ &= -\infty, \end{aligned} \tag{8}$$

where the final equalities in Equations 7 and 8 are direct consequences of application of L'Hospital's rule. Since $\lim_{n\to\infty} \ln x_n = 0$, it must be that $\lim_{n\to\infty} x_n = 1$, and since $\lim_{n\to\infty} \ln x_n^n = -\infty$, it must be that $\lim_{n\to\infty} x_n^n = 0$. Therefore as $n \to \infty$, the limiting probability that any single individual acts is $1 - \lim_{n\to\infty} x_n = 0$ and the probability that at least one individual acts is $1 - \lim_{n\to\infty} x_n^n = 1$.

□

Table 1 compares symmetric Nash equilibria and symmetric optima as the number of players is varied, with the parameters b and c set at $b = 1$, $c = .9$. The first column of the table shows the number of players. The second and third columns show the probability of taking actions for an individual player in Nash equilibrium and the symmetric optimum, respectively. The fourth and fifth columns show the equilibrium probability that at least one player takes action. The last two columns show the utility achieved by each player in Nash equilibrium and the symmetric optimum.

Table 1: Symmetric Nash equilibria and Symmetric Optima ($c/b = .9$)

n	$p_N = 1 - q_n$	$p_O = 1 - x_n$	$P_N = 1 - q_n^n$	$P_O = 1 - x_n^n$	u_N	u_O
2	0.100	0.550	0.190	0.798	0.1	0.303
3	0.051	0.452	0.146	0.836	0.1	0428
4	0.035	0.392	0.131	0.863	0.1	0.511
5	0.026	0.349	0.123	0.883	0.1	0.569
25	0.004	0.129	0.104	0.969	0.1	0.852
50	0.002	0.079	0.102	0.983	0.1	0.913
100	0.001	0.047	0.101	0.991	0.1	0.950

Table 1 shows that when the number of players is small, in the optimal symmetric solution, players are much more likely to take action than in Nash equilibrium. But with small numbers of players, the probability that someone takes action is significantly smaller than one, even in the optimal solution. In both cases, as the number of players increases, the probability that an individual takes action falls toward zero As n increases, the Nash equilibrium probability that *at least one* player takes action falls and asymptotically approaches $1 - \frac{c}{b} = .10$. In contrast, in the symmetric optimum, as the number of players increases, the probability that at least one player takes action increases and asymptotically approaches unity. The expected utility of each player in Nash equilibrium remains at $b - c = 0.1$ for any number of players, while the expected utility of each player at the symmetric optimum rises and asymptotically approaches $b = 1$, which is the utility that a player who took no action would get if someone else were sure to take action.

2 Differing Costs and Incomplete Information

It is common practice to "simplify" game theoretic models like the Volunteer's Dilemma by assuming that all players have identical benefits and costs. While this simplification makes it easy to calculate a symmetric Nash equilibrium, the resulting mixed-strategy Nash equilibrium has an air of implausibility. In the symmetric mixed strategy equilibrium of the Volunteer's Dilemma game, all players are indifferent between the equilibrium mixed strategy and any other probability mix of the strategies "act" and "don't act." Given that this is the case, why should any player take the trouble to determine the equilibrium mixed strategy proportions and act accordingly?[1] If we allow the realistic possibility that different players have different costs of taking action, we avoid this conundrum and we can construct a manageable model in which players use pure strategies in a symmetric Nash equilibrium and where the optimal symmetric ethical rule recommends to each player a pure strategy that is determined by that player's realized cost of taking action.

Let us suppose that the costs c of taking action differ among players, while the benefits b received by each player are normalized to $b = 1$.[2] We assume that players are total strangers to each other and can not communicate before deciding whether to act. Individuals know their own costs, but do not know the costs of the other players in the game. We will suppose that

[1] Herbert Gintis [8] describes this quandry as "the mixing problem".

[2] Setting $b = 1$ for all players involves no real loss of generality for equilibrium analysis, since every player's choice of whether to take action is determined by the ratio of this player's costs to benefits.

players are chosen by independent draws from a population with a distribution of costs that is common knowledge.[3] We assume that the distribution from which players' costs are drawn has the following properties:

Assumption 1. *Players costs are drawn randomly from a population in which the cumulative distribution of costs is $F(\cdot)$, with support $[0,1]$. The corresponding density function $F'(\cdot)$ is continuous over the interval $c \in (0,1)$. This distribution has $F(0) < 1$ and $F(1) > 0$.*

The assumption that $F(0) < 1$ means that there is a positive probability that costs are positive and the assumption that $F(1) > 0$ that there is a positive probability that a player's costs are less than individual benefits, which have been normalized to 1.

2.1 Symmetric Nash Equilibrium

We can model this game as one that begins before individuals learn their costs. Then the game is a symmetric game in which a strategy for any player is a function of the costs that will be revealed to this player. There will be a symmetric Nash equilibrium in which every player uses a *threshold strategy* of the form: "Act if and only if your costs, c, turn out to be no larger than a common threshold level $\hat{c}$." The threshold strategy with threshold $\hat{c}$ will be a Nash equilibrium if and only if when all other players follow this rule, a player with realized cost c will have a higher payoff from acting if $c < \hat{c}$ and a higher payoff from not acting if $c > \hat{c}$. We have the following result.

Proposition 3. *For all $n \geq 2$, in an n-player Volunteer's Dilemma game with incomplete information, where the distribution of costs is common knowledge and satisfies Assumption 1, there is a unique threshold equilibrium with threshold $\hat{c}(n) \in (0,1)$. The equilibrium threshold level $\hat{c}(n)$ decreases as n increases.*

Proof. Let us define $G(c) = 1 - F(c)$. Where $b = 1$, if all other players use the threshold strategy with threshold $\hat{c}$, then for a player with cost c, the expected payoff to not acting is $1 - G(\hat{c})^{n-1}$ and the expected payoff to

[3] The assumption of incomplete information seems appropriate for games in which players are thrown together by chance for a single interaction. Situations where the same players are engaged in repeated encounters and know each other well might better be treated as games of complete information. Weesie [10] characterizes asymmetric equilibria for Volunteer's Dilemma games with differing payoffs, but complete information in which players know each other's payoffs.

acting is $1-c$. A player with cost c will be indifferent between acting and not acting if

$$1 - G(\hat{c})^{n-1} = 1 - c, \tag{9}$$

or equivalently,

$$G(\hat{c})^{n-1} = \hat{c}. \tag{10}$$

Assumption 1 implies that $G(\cdot) = 1 - F(\cdot)$ is a decreasing function and that $G(0) > 0$ and $G(1) < 1$. Let $H(c,n) = G(c)^{n-1} - c$. Then H is a continuous, strictly decreasing function such that $H(0,n) > 0$ and $H(1,n) < 0$. Therefore for any $n > 1$, there is exactly one solution $\hat{c}(n) \in (0,1)$ such that $H(\hat{c}(n), n) = 0$. This $\hat{c}(n)$ is a unique solution to Equation 10.

Taking the log of Equation 10 and differentiating with respect to n, we have

$$\ln G\left(\hat{c}(n)\right) + (n-1)\frac{\hat{c}(n)G'\left(\hat{c}(n)\right)}{G\left(\hat{c}(n)\right)}\left(\frac{\hat{c}'(n)}{\hat{c}(n)}\right) = \frac{\hat{c}'(n)}{\hat{c}(n)} \tag{11}$$

and hence

$$\frac{\hat{c}'(n)}{\hat{c}(n)} = \frac{\ln G\left(\hat{c}(n)\right)}{1 - (n-1)\frac{G'(\hat{c}(n))}{G(\hat{c}(n))}} \tag{12}$$

Since $0 \leq G\left(\hat{c}(n)\right) \leq 1$ and $G'\left(\hat{c}(n)\right) \leq 0$, the numerator of Equation 12 must be negative and the denominator must be positive. It follows that $\hat{c}'(n) < 0$ and hence $c(\cdot)$ is a decreasing function of n. □

Theorem 3 is reminiscent of Harsanyi's [9] purification theorem. Harsanyi shows that a mixed strategy equilibrium for a game with identical players is the limit point of of pure strategy equilibria for games in which the payoff functions of individuals have small perturbations around uniform payoffs for which the mixed strategy equilibrium is calculated. In the application treated here, there is no reason to believe that differences in costs are small deviations around some common value.

Proposition 4. *The probability that nobody takes action in a symmetric threshold equilibrium of n players increases with n if*

$$1 + \frac{\rho G'\left(\hat{c}(n)\right)}{G\left(\hat{c}(n)\right)}$$

is negative and decreases with n if this term is positive.

Proof. The equilibrium probability that nobody takes action if there are n players is $G\left(\hat{c}(n)\right)^n$. Since $\hat{c}(n)$ must satisfy Equation 10 it follows that

$$G\left(\hat{c}(n)\right)^n = \hat{c}(n)G\left(\hat{c}(n)\right) \tag{13}$$

Differentiating Equation 13, we have

$$\frac{\partial G\left(\hat{c}(n)\right)^n}{\partial n} = \hat{c}'(n)G\left(\hat{c}(n)\right)\left(1 + \frac{\hat{c}(n)G'\left(\hat{c}(n)\right)}{G\left(\hat{c}(n)\right)}\right) \tag{14}$$

According to Proposition 3, $\hat{c}'(n) < 0$. Therefore it must be that $G\left(\hat{c}(n)\right)^n$ decreases with n if

$$1 + \frac{\hat{c}(n)G'\left(\hat{c}(n)\right)}{G\left(\hat{c}(n)\right)}$$

is positive and increases with n if this term is negative.

□

Examples can be found where this is either positive or negative. For example if $G(c) = c^{-a}$, the expression in Equation 14 will be positive if $0 < a < 1$ and negative if $a > 1$. I need to work out some more examples and better yet have a useful general characterization of when the derivative is positive or negative. [5]

2.2 Optimal symmetric strategies

Let us evaluate possible symmetric mixed strategies from an initial position in which players do not yet know their own costs, but expect them to be drawn at random from the population distribution. From this standpoint, players have identical prospects and we can solve for an optimal symmetric strategy. We consider symmetric strategies that consist of a threshold cost level c^* and a mandate that any player should take action if and only if this player has costs $c \leq c^*$). If the threshold is set at c^*, the probability that any single player will not take action is $G(c^*)$ and the probability that at least one player will take action is $1 - G\left(c^*\right)^n$. Before individuals learn their own costs, the expected value of the costs each will have to pay is

$$\int_0^{c^*} xG'(x)dx.$$

Thus, if there are n players and if the threshold is set at c^*, then, before individual costs are revealed, the expected utility of every player l must be

$$(1 - G(c^*)^n) - \int_0^{c^*} xG'(x)dx. \tag{15}$$

This expression is maximized at $c^*(n)$ where

$$nG\left(c^*(n)\right)^{n-1} G'\left(c^*(n)\right) = c^*(n)G'\left(c^*(n)\right), \tag{16}$$

or equivalently when

$$\frac{G\left(c^*(n)\right)^{n-1}}{c^*(n)} = \frac{1}{n} \tag{17}$$

Where $\hat{c}(n)$ is the symmetric Nash equilibrium threshold, for n players, we see from Equations 10 and 17 that for any $n > 1$

$$\frac{G\left(c^*(n)\right)^{n-1}}{c^*(n)} = \frac{1}{n}\frac{G\left(\hat{c}(n)\right)^{n-1}}{\hat{c}(n)} \tag{18}$$

Since $G(c)$ is decreasing in c, it follows that

$$\frac{G(c)^{n-1}}{c}$$

is a strictly decreasing function of c. It therefore follows from Equation 18 that for all $n > 1$, $c^*(n) > \hat{c}(n)$.

References

[1] Theodore Bergstrom. The Good Samaritan and traffic on the Road to Jericho. Working paper (available at works.bepress.com/ted_bergstrom/112/), Department of Economics University of California Santa Barbara, 2012.

[2] Theodore Bergstrom. Single-payer volunteer's dilemma. Working paper, Department of Economics, University of California Santa Barbara, 2012.

[3] Theodore Bergstrom, Rodney Garratt, and Damien Sheehan-Connor. One chance in a million: Altruism and the bone marrow registry. *American Economic Review*, 99(4):1309–1334, 2009.

[4] Marc Bilodeau and Al Slivinski. Toilet cleaning and department chairing: Volunteering a public service. *Journal of Public Economics*, 59:299–308, 1996.

[5] Ken Binmore. *Playing for Real.* Oxford University Press, New York City, NY, 2007.

[6] Christopher Bliss and Barry Nalebuff. Dragon slaying and ballroom dancing: The private supply of a public good. *Journal of Public Economics*, 25:1–12, 1984.

[7] Andreas Diekmann. Volunteer's dilemma. *The Journal of Conflict Resolution*, 29(4):605–610, December 1985.

[8] Herbert Gintis. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, Princeton, N.J., 2009.

[9] John Harsanyi. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2:1–23, 1973.

[10] Jeroen Weesie. Asymmetry and timing in the volunteer's dilemma. *The Journal of Conflict Resolution*, 37(3):569–590, September 1993.

[11] Jeroen Weesie and Axel Franzen. Cost sharing in a volunteer's dilemma. *The Journal of Conflict Resolution*, 42(5):600–618–590, October 1998.