

Are we capable of being altruistic?

Third Prize – Postgraduate Category

JACK WHYBROW* AND JIWEI ZHENG**

Abstract

This short essay will seek to address whether we are capable of being altruistic. The essay will start with a discussion of the definition of altruism both from a modern language and theoretical perspective before moving on to discuss the ways in which economists have attempted to capture altruism and concluding with a general discussion. The results suggest behavioural economists have as yet been unable to explain acts of kindness purely in altruistic terms and instead have elements of other motivations. That does not however suggest that we are not capable which incidentally may or may not be true, but given current experimental methods we are yet to be successful in isolating strong altruism.

Altruism

The Oxford English Dictionary (Paperback) defines altruism as "unselfish concern for other people" (Oxford English Dictionary, 2006, pp.20). It originates from the Italian 'altrui' meaning somebody else. To the authors of this work it appears that the definition of selfish and how this is defined is key to answering this question. The same source defines selfish as "concerned mainly with your own needs and wishes" (OED, 2006, pp.683). Needs and wishes are likewise defined as "want something because it is essential or important" (OED, 2006, pp.501) and "to

* PhD Applied Econometrics and Policy

**PhD Experimental and Behavioural Economics

feel a need or desire to have or do" (OED, 2006, pp.863). Similarly the body of literature on altruism defines this as: "Not abject self-sacrifice, but merely a willingness to act in the consideration of the interests of other persons, without the need of ulterior motives" (Thomas Nagel, 1970, p.79).

Traditionally economists like to model the psychological processes of agents in terms of expected utility¹ where utility (or utils) is a quantified measure of gain/satisfaction from consumption of material goods. Doing a good deed therefore does not involve any material consumption and in this sense satisfies the definition of altruism. All-in utility (Zizzo, 2000) however is not constrained by the same materiality condition therefore if an individual gains some utils from the act (referred to as warm glow (Andreoni, 1990)) then this is not pure altruism. Conversely the same solution is arrived at if there is a penalty (dis-utils) in their utility function² such as guilt associated with turning a blind eye to a fellow human in need -- inequity aversion, (Fehr and Schmidt, 1999). Therefore if an individual were to rationally stop help (assuming the cost of doing so is less than the disutility associated with doing nothing) then the individual is motivated by self-interest and therefore cannot be altruistic.

Experimental Economics – Measuring Altruism

Behavioural economists have attempted to use a variety of games to capture altruism, the most widely-used of which are the Ultimatum, Dictator and Trust Game. Each of these and their variations are conducted in experimental labs with test subjects. The Ultimatum Game (Guth et al., 1982) involves two participants who are randomly paired without complete information of who the other is paired with in a group of experimental subjects. The dictator (proposer) is given an endowment of experimental credits and then decides how much (if any) to allocate to the recipient (responder). These are subsequently exchanged at the end of the experiment into real monetary values. With knowledge of the dictator's proposal the recipient now decides whether or not to accept or reject. If the recipient chooses to accept then both parties receive what was proposed by the dictator. If however the recipient chooses to reject the dictator's allocation both participants receive 0. Assuming both players are rational and are only

¹ There is a substantial literature on Expected Utility, its alternatives and rationality although this has been omitted for the purposes of clarity.

² A deterministic way of expressing the happiness an individual feels for a certain event.

interested in their associated payoffs the dictator will reason that as long as his offer is strictly greater than 0 then the recipient should agree, maximizing their respective earnings. The dictator should therefore offer the smallest denomination of experimental credits and the recipient should accept.

What we find is that although this result is observed in some instances the majority offer between 40-60% of their endowment. Offers below 40% are often accompanied with a high rate of rejection. (e.g. Cameron, 1999; Croson 1996; Eckel and Grossman, 2001;) Fairness and inequity aversion are often used to justify why recipients reject as the negative utility from accepting offers below 40% are greater than the utility associated with the monetary payoff to the recipient, leading to the recipients to reject the allocation. It is however curious as to why dictators propose offers of more than 40%. One reason may be due to inequity aversion (Fehr and Schmidt, 1999) in that individuals are adverse to offers that are less than 'fair'. Other academics have also attempted to explain this by extending the number of periods in which participants are paired. The results suggest that through learning effects we observe some convergence to 40% although some dictators still offer consistently more than this.

Several reasons in addition to altruism have been proposed to explain this effect these are namely: inequity aversion, warm glow and experimenter demand effects (Zizzo, 2010). Considering inequity aversion, the dictator may have only shared the material payoff to her co-player because an unfair proposal would generate substantial negative utilities above those associated with a greater material payoff.³ Alternatively if we consider warm glow, in sharing the allocation more equally the individual may gain positive utilities above that which would have been generated from receiving a higher material payoff. Finally, experimenter demand effects may influence individual decisions and result in a higher incidence of sharing either because of experimenter scrutiny or because the participants act in a way the experimenters would hope to see.

The Dictator Game (Kahneman et al., 1986) is a variant of the Ultimatum Game but with no avenue of recourse for the recipient when faced with the dictators' proposal. The dictator is

³ $\psi = -\alpha (v-u)$ if $v \geq u$ and $\psi = -\beta (u-v)$ if $v < u$, $0 < \alpha < \beta$, or the difference between the dictator and his/her co-player's material payoffs will decrease the dictators all-in utilities, the dictator will choose to share some material payoffs to his/her co-player to avoid this decrease.

given an endowment of experimental credits and then decides how much (if any) to allocate to the recipient. Rational self-interested individuals would deduce that the recipient has no recourse to object to the dictator's decision and therefore should allocate 0 to the recipient. Although this result is witnessed in test conditions not all agents allocate 0. Of those that do not allocate 0 the dictator gives on average 40% of their endowment to the recipient. Some have tried to verify this result by turning the game from a one-shot game to multiple periods, keeping the allotted dictators and recipients constant throughout. Regardless of adding additional periods the mean does not converge (fully) to 0. This result suggests something systematic in the way dictators are choosing to allocate their endowment. As we saw in the ultimatum game similarly inequity aversion, warm glow and experimenter demand effects have also been proposed to explain this result in addition to altruism.

The trust game proposed by Berg, Dickhaut and McCabe (1995) like the previous two involves two participants who are randomly paired in a group of participants without knowledge of who the other is paired with. One participant is then allocated as the investor (truster) and the other the recipient (trustee). The investor is then given an endowment of experimental credits and then decides how much (if any) to invest. Usually the amount invested is then multiplied by some factor and transferred to the recipient. The recipient then decides how much if any to return to the investor. Fully rational investors would deduce that self-interested recipients will return 0 maximizing their payoff over a one period game, although both could do jointly better if they invested. Extending the game to multiple periods rational agents should reason the same by backwards induction. Suppose the trust game runs for 10 periods then the most rational thing for a self-interested profit-maximizing recipient to do is to return 0. An investor will reason this and therefore invest 0 on the ultimate round. The recipient reasoning this would therefore return 0 on the penultimate round and so forth.

What we find is that although this result is observed, many participants establish trust in the initial rounds only defaulting in the penultimate rounds. In some cases trust and fulfilling is also maintained throughout the whole experiment. Jacobsen and Sadrieh (1996) found participants invested 60% and were repaid 110%. Koford (1998) found the participants invested 70% and got 150% back. As with the previous two cases the results found in these experiments cannot be explained purely with altruism. Aside from inequity aversion, warm glow and experimenter

demand effects there are at least two other explanations. The first is 'kindness reciprocity' and the second is 'trust responsiveness'. Kindness reciprocity (Rabin, 1993; Falk and Fischbacher, 2001) occurs when a recipient reciprocates the trust of the trustee in returning some of the investment. 'Trust responsiveness' (Bacharach and Zizzo, 2007) occurs when the recipient fulfils this trust because they believe the investor trusts them. The more she believes so, the more she is going to fulfil the investor.

In summary we cannot explain the results purely with altruism instead we must use this in conjunction with other reasons. This may be because we do not yet possess the methodology to isolate pure altruism from individual players or simply it does not exist. It is however likely that some form of altruism may exist as in any experiment as in real life there will be differences between individuals and their underlying motivations.

Philosophical Incites and General Discussion

Experimental results aside it can however be argued that it is unreasonable to expect an individual to commit a purely altruistic act/action without some form of gain whether this be material or otherwise. Even if the individual commits an act that appears purely altruistic due to human, computational or forecasting errors (and it is inconceivable that an individual can gain from it), they are likely to have taken this action due to self-interested reasoning. Critiques of this would likely bring forth notable examples such as Nelson Mandela to name one but a few as an example of altruism. Or can it? It would be hard to refute that his fame has not allowed him a much more extravagant lifestyle than that which can be obtained through work. Notable examples aside it is important to remember altruism may come in many different forms such as someone who holds down three jobs and does his/her best for their family and still has the time to help a stranger in need. Although even here it may be argued that this individual has a vested interest in their children (sunk costs) and may help a stranger due to inequity aversion.

In the Wealth of Nations Adam Smith (attributed as the father of economics and by some of behavioural economics) commented about the idea that an individual may sympathize with another individual/being when they are experiencing a bad turn of events, due partly to we imagining ourselves in their shoes. Again this can be explained at least in part through inequity aversion. There may also even be something biological in why we respond this way which may

have allowed us to feel compassion for our own (as long as the risks to the individual are not too great) and in that way bolstered the species survival as a whole.

This essay has been primarily focused on interpreting the question as whether or not we are capable of committing an act without some form of gain. If the question is purely concerned with whether or not we are 'physically' capable of committing an altruistic act by error or otherwise then we must surely believe this is possible. Key to the argument presented above is whether or not we can indeed separate the 'act of altruism' from the 'belief system'. In essence separating the act from whether an individual believes they are doing a good deed⁴.

Conclusion

This short essay aimed to address whether or not we are capable of being altruistic. The question itself was broad and could be interpreted in a number of different ways. For example if we are to purely consider whether we are 'physically' capable of committing an altruistic act then we firmly agree with this view. If we are to consider the deeper issue of whether an individual would purposely commit an altruistic act we find no justification. The evidence presented within the main body of this essay suggests we cannot dispute altruism as possible justification for what we observe but the results can also be explained by other motivations. We do not yet have the methodology to isolate altruism. Indeed this may even be unattainable. In the meantime we must look to other disciplines for insights into this problem.

⁴ One would be justified in looking at relativism and determinism here, although these were omitted as the authors felt this was steering too far into the realm of philosophy in which neither author is well equipped to navigate.

Bibliography

1. Andreoni, James (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *Economic Journal* 100 (401): 464–477.
2. Bacharach, M., G. Guerra & D.J. Zizzo (2007). The Self-Fulfilling Property of Trust: An Experimental Study. *Theory and Decision*, 63, 349-388.
3. Berg Joyce, Dickhaut John and McCabe Kevin (1995). Trust, Reciprocity and Social History, *Games and Economic Behavior*, 10, 122-142.
4. Cameron, Lisa A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, *Economic Inquiry*,27,47-59.
5. Croson, Rachel T.A. (1996). Information in ultimate games: An experimental study, *Journal of Economic behavior and Organization*,30,197-212.
6. Eckel, Catherine C., and Philip Grossman (2001). Chivalry and solidarity in ultimatum games, *Economic Inquiry*,39,171-88.
7. Falk, A. and Fischbacher, U. (2001). Distributional consequences and intentions in a model of reciprocity, *Annales d’Economie et de Statistique*, 63–64, 111–129.
8. Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition and cooperation, *Quarterly Journal of Economics* 114, 817–868.
9. Guth, W., Schmittberger, R and Hernd S. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*,3,367-88.
10. Jacobson,E. and Abdolkarim, S. (1996). Experimental proof for the motivational importance of reciprocity. University of Bonn.
11. Kahneman,D. Jack,L.K. and Richard,T. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American economic review*, 76, 728-41.
12. Koford, K. (1998). Trust and reciprocity in Bulgaria: A replication of Berg, Dickhaut and McCabe (1995). University of Delaware Department of Economics working paper.
13. Nagel, Thomas (1970). *The Possibility of Altruism*. Oxford: Clarendon Press.
14. Oxford English Dictionary (2006) "Paperback Oxford English Dictionary" Soans, C., Hawker, S. & Elliott, J. (Eds) Oxford University Press
15. Rabin, M. (1993), Incorporating fairness into game theory and economics, *American Economic Review* 83, 1281–1302.
16. Zizzo, D.J. (2000). Relativity-sensitive behaviour in economics, Doctoral thesis, University of Oxford
17. Zizzo, D.J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13, 75-98.