Economics Letters 135 (2015) 133–136

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Interactive preferences

Heinrich H. Nax*, Ryan O. Murphy, Kurt A. Ackermann

Department of Social Sciences, ETH Zürich, Clausiusstr. 37, 8092 Zurich, Switzerland

HIGHLIGHTS

- A game-theoretic model of socially interactive preferences is proposed.
- Individuals' concerns for one another dynamically and reciprocally depend on each other.
- The interactions of preferences are tested experimentally.
- On average, pro-sociality is shown to diminish over time and across experiments.

ARTICLE INFO

Article history:
Received 15 June 2015
Received in revised form
27 July 2015
Accepted 7 August 2015
Available online 14 August 2015

Keywords:
Game theory
Social preferences
Preference evolution

ABSTRACT

Game theory presumes that agents have unique preference orderings over outcomes that prescribe unique preference orderings over actions in response to other players' actions, independent of other players' preferences. This independence assumption is necessary to permit game-theoretic best response reasoning, but at odds with introspection, because preferences toward one another often dynamically depend on each other. In this note, we propose a model of interactive preferences. The model is validated with data from a laboratory experiment. The main finding of our study is that pro-sociality diminishes over the course of the interactions.

1. Introduction

Mother Teresa does not defect in prisoners' dilemmas, because she cares for her opponents in ways that transform the games' mixed motives into other games where her and common motives are aligned (e.g., harmony). Cooperation thus emerges as a dominant strategy. The experimental economics literature is concerned with 'subjective expected utility corrections' (Gigerenzer and Selten, 2001) that modify players' utility representations to account for such other-regarding concerns. Numerous corrections have been proposed (e.g., Rabin, 1993; Levine, 1998; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) in light of laboratory evidence that manifests systematic deviations from narrow self-interest predictions (see Ledyard, 1995 and Chaudhuri, 2011 for reviews).[1] This route of enquiry is bothersome for many theoretical game theorists who question how these findings generalize beyond the laboratory.[2]

Missing from most alternative utility formulations are interactive components that meaningfully alter the game-theoretic analysis. Standard theory (von Neumann and Morgenstern, 1944) equips players with preferences that prescribe actions vis-à-vis others' actions, independent of others' preferences. Here, inspired by Rabin (1993) and Levine (1998), we propose a model of interactive preferences among players that depend on each other and investigate their dynamic interdependence. The model is validated with laboratory studies involving repeated voluntary contributions games (VCM; Isaac et al., 1985) sandwiched by two sets of dictator games (DG; Kahneman et al., 1986) used to evaluate individuals' social value orientation (SVO; Murphy et al., 2011). Our results show that, independent of unintended behavioral deviations, the proportion of behavior associated with pro-sociality diminishes over the course of the interactions and is replaced by individualism. These patterns carry over between VCMs and DGs. Our model fares predictively well.

2. Methods

2.1. Experimental setup

Experiments were run at ETH's Decision Science Laboratory during February 2013 involving 128 subjects in 6 sessions ($4 * 20$

* Corresponding author.
E-mail address: hnax@ethz.ch (H.H. Nax).

1 With some exceptions (e.g., Saijo and Nakamura, 1995; Saijo, 2008), many analytical set-ups have been biased as discussed in, for example, Burton-Chellew and West (2013) and Burton-Chellew et al. (2015).

2 See controversies in JEBO 73, 2010.

http://dx.doi.org/10.1016/j.econlet.2015.08.008
0165-1765/

+ 2 ∗ 24). Subjects were informed in detail and in advance of each stage of the experiment using standard instructions.[3] Every decision was monetarily incentivized, and subjects earned over 40 CHF (over 40 US$) on average. The experiment lasted roughly 90 min.

The experiment had the following three stages:

Stage 1: ***Initial SVO***. Subjects played 6 DGs choosing allocations in different ranges representing different himself-versus-other tradeoffs; for example, between 100 for himself and 50 for the other, (100, 50), and 50 for himself and 100 for the other, (50, 100).[4] The 6 decisions are represented as angles in the classical SVO ring (Griesinger and Livingston, 1973), and an individual's initial SVO is taken as the average angle, representing a compact indicator of his *ex ante* SVO.[5]

Stage 2: ***VCM***. Subjects played 10 VCMs in groups of 4 that were randomly formed in round 1 and then remained fixed for the remainder of rounds. In each round subjects made contributions and guessed others' average contributions (with incentives for accuracy). Before each round, players were informed of the previous-period contributions. (More details will be provided shortly.)

Stage 3: ***Final SVO***. Stage 1 is repeated, thus measuring individuals' *ex post* SVOs.

Our analysis focuses on 22 data points per person, namely his 2 – initial and final – SVOs, plus his 10 contributions and 10 guesses about others' contributions from the VCM, yielding a total of 2,816 data points.

2.2. The model

2.2.1. Static model

Population $N = \{1, 2, 3, 4\}$ plays a VCM with marginal per capita rate of return $r = 0.4$ and budget $B = 20$. Each $i \in N$ sets a private contribution $c_i \in B$ which, jointly with the others' average contribution, c_{-i}, results in payoff

$$\phi_i = 20 - c_i + 0.4(c_i + 3c_{-i}).$$

We assume i's utility depends on payoffs in Cobb–Douglas form

$$u_i(c) = \phi_i^{1-\alpha_i} * \phi_{-i}^{\alpha_i}, \tag{1}$$

where $\alpha_i \in [0, 1]$ measures player i's concern for others. The nonlinearity of expression (1) distinguishes it from most representations, including (Levine, 1998), thus rationalizing intermediate contributions in terms of intermediate concerns. We obtain the following expression for α_i by assuming c_i is chosen optimally given his guess about c_{-i} (expressed as $\widehat{c}_{-i}$):

$$\alpha_i = \frac{0.6\phi_{-i}(c_i, \widehat{c}_{-i})}{0.4\phi_i(c_i, \widehat{c}_{-i}) + 0.6\phi_{-i}(c_i, \widehat{c}_{-i})}. \tag{2}$$

Note that $\partial\alpha_i/\partial c_i > 0$ and $\partial\alpha_i/\partial\widehat{c}_{-i} < 0$, that is, higher own contributions (holding beliefs about others constant) indicate more concern for others, and higher beliefs regarding others' contributions (keeping own contributions fixed) indicate less concern for others.

The interdependence of preferences results from imposing that, in static equilibrium, $\alpha_i = \widehat{\alpha}_{-i}$, where $\widehat{\alpha}_{-i}$ is i's belief about α_{-i}.[6] The resulting set of equilibria, the general structure of which is under investigation in an ongoing study, contains the standard case (when $\alpha_i = \alpha_{-i} = 0$) but also new ones when $\alpha_i = \alpha_{-i} > 0$ as in fairness equilibria (Rabin, 1993).

2.2.2. Dynamic components

The above game repeats with revelation of past outcomes. Each period t, suppose i contributes to maximize expression (1) so that expression (2) implies α_i^t given $(c_i^t, \widehat{c}_{-i}^t)$. We assume α_i^t is updated in light of evidence by

$$\alpha_i^t = (1 - \beta_i^t)\alpha_i^{t-1} + \beta_i\widetilde{\alpha}_{-i}^{t-1}, \tag{3}$$

where $\widetilde{\alpha}_{-i}^{t-1}$ is i's deduction of α_{-i}^{t-1} from previous-period evidence, and $\beta_i^t \in [0, 1]$ measures i's period-t degree of belief responsiveness.

2.3. Estimation strategy

2.3.1. Classification

Initial SVOs are used to classify individuals as 'individualistic' and 'pro-social'. An individual is pro-social (individualistic) according to the SVO measure if his SVO-angle is ≥ 22.5 (< 22.5) degree.[7] The initial SVO classifications are used to predict initial VCM contributions

'Responsive' and 'unresponsive' types are classified based on the VCM data. Individual i is said to be responsive (unresponsive) if the estimation of expression (3) in light of his VCM decisions from rounds 2–10 yields an average coefficient for β_i^t which is positive (not positive).

2.3.2. Prediction

We use our estimated 2×2 typology (from initial SVO and VCM) to make predictions regarding final SVO classifications, which we shall assess in light of the recorded final SVOs. We shall use the following terminology: an individual is associated with a VCM group matching that is said to be 'individualistic' ('pro-social') if those players he is matched with, on average, contribute less (more) than himself.

We predict unresponsive types (pro-social and individualistic alike) not to change their preferences. We predict responsive types to change their types in the direction of their interaction partners as matched with during the VCM group matching. Hence, a responsive pro-social (individualist) in a VCM group matching that is pro-social (individualistic) will remain pro-social (individualistic). A responsive pro-social (individualist) matched with individualistic (pro-social) others, however, may become individualistic (pro-social), dependent on the action/payoff difference between himself and his opponents. In particular, whichever payoff difference is larger we shall assume will be associated with a preference-change flow of probability one, and the lesser payoff-difference to be proportional to that flow depending on the relative payoff difference.

[3] See Murphy and Ackermann (2013) for details.

[4] The remaining 5 choices are amongst linear combinations in the ranges [(100, 50), (85, 85)], [(50, 100), (85, 15)], [(50, 100), (85, 85)], [(85, 85), (85, 15)], and [(85, 15), (100, 50)].

[5] Angles close to 0° represent individualistic preferences in the sense of material self-interest, angles ≥22.5° indicate pro-sociality.

[6] A weaker assumption in the same spirit would be to weigh this dependence by some parameter as in Levine (1998), something we shall introduce via 'responsiveness' instead.

[7] See Murphy and Ackermann (2013) for a more fine-grained categorization.