Private Goods, Public Goods and Common Pools with *Homo Reciprocans*

By James C. Cox[1]

1. Introduction

Social dilemmas characterize environments in which individuals' exclusive pursuit of their own material self-interest can produce inefficient outcomes such as those resulting from under-provision of public goods and over-appropriation of common pool resources. Under-provision is manifested in foregone opportunities to create surplus through provision of a public good with greater value than the sum of its private costs. In a conceivable extreme outcome, free-riding on voluntary provision can lead to outright failure to provide a public good. Over-appropriation is manifested in realized destruction of surplus through appropriation of private goods with a sum of values less than the value of the common pool that has been exploited. In a conceivable extreme outcome, competition under the "rule of capture" for appropriation can lead to destruction of the common pool, a tragedy of the commons.

In practice, some public goods are voluntarily provided in positive (not necessarily optimal) amounts. And some common pools are maintained (not necessarily optimally) through voluntary cooperation. In this way the most pessimistic predictions coming from models of competition between narrowly selfish agents are often not manifested in outcomes. The question is: Why? Or, more descriptively, what are some answers to many specific questions about voluntary cooperation in environments characterized by social dilemmas?

This paper addresses the following questions. Does reciprocity support cooperative behavior that ameliorates under-provision? Is over-appropriation of a common-pool less of a problem when agents' behavior is characterized by reciprocity? Is reciprocity more or less effective in promoting efficiency in an environment with public good provision or one with common pool appropriation? Does asymmetric power exacerbate or ameliorate under-provision and over-appropriation? How can experiments be used to address these questions? Does

[1] This paper is based on the Presidential Address delivered at the Southern Economic Association's 81st Annual Meeting, November 21, 2011. Helpful comments and suggestions were provided by Vjollca Sadiraj and Kurt Schnier and by the students in my PhD Workshop in Experimental Economics. Financial support was provided by the National Science Foundation (grant number SES-0849590).

behavior in the experiments support or fail to support alternative theories that have testable implications for agent behavior in environments with public goods and common pools?

Dilemmas can also arise in environments with only private goods when opportunities for cooperatively increasing payoffs can be realized or squandered. Examples of such environments are provided by the (textbook) Stackelberg duopoly game and the investment game (Berg, Dickhaut, and McCabe, 1995). I ask whether experiments reveal cooperative behavior in these private property games and whether the same theory that organizes data from environments with public goods or common pools can also organize data from duopoly and investment games.

There are large literatures spanning several disciplines devoted to exploring questions about public good provision, common pool appropriation, and private property allocation. Surveying those literatures is not feasible for the space available here, and would in any case detract from the sharp focus I want to give to one approach. That approach is based on construction of pairs of (money) payoff equivalent games and design of experiments to get empirical insights from play of the games. Two types of game pairs are constructed: games with simultaneous moves by agents with the same feasible choices and games with sequential moves in which one of the movers has asymmetric power to affect final allocations.

The two games in each pair with either simultaneous or sequential moves are constructed to be (money) payoff equivalent. The two games in each pair are also strategy equivalent for the *homo economicus* model of narrowly selfish preferences and for models of social preferences in which preferences over distributions of material outcomes are a fixed characteristic of an agent (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Cox and Sadiraj, 2007). In contrast, the sequential move games are not strategy equivalent for *homo reciprocans* preferences (Cox, Friedman and Gjerstad, 2007; Cox, Friedman and Sadiraj, 2008) in which one agent's other-regarding preferences can depend on the prior actions of another agent. The simultaneous move games are strategy equivalent for *homo reciprocans* preferences as well as fixed (non-reciprocal) preferences. Theory for *homo reciprocans* yields insight into empirical regularities observed in public good and common pool games. For example *homo reciprocans* theory predicts that, in sequential move games, over-appropriation from a common pool will be a more serious problem than under-provision of a public good. This is observed in experiments.

2. Revealed Altruism Theory

Neoclassical theory, as represented by the classic works by Hicks (1946) and Samuelson (1947), develops a formal explanation of how observable choices by *homo economicus* are jointly determined by preferences and observable opportunities. Revealed altruism theory (Cox, Friedman and Sadiraj, 2008) applies that approach to develop a theory of reciprocity, that is, a model of *homo reciprocans*. The theory focuses on willingness to pay (WTP) amounts of one's own income to change another's income that is *revealed* by choices made in response to opportunities with an observable need to pay to make such changes, as defined by an opportunity set. Revealed altruism theory differs from neoclassical theory through its formal representation of how WTP can depend on *observed* actions by another that change one's own opportunities. Revealed altruism theory introduces to the literature a partial ordering of preferences, a partial ordering of opportunity sets, and two axioms about reciprocity that link the partial orderings.

The partial ordering of preferences is based on WTP, as follows. So long as marginal utilities are well-defined, my WTP an amount of my income *m* to change your income *y* is given by the ratio of my marginal utilities: $WTP(m,y) = u_y(m,y)/u_m(m,y)$. Two different preference orderings, A and B, over own and other's income can represent the preferences of two different agents or the preferences of the same agent in two different situations. If $WTP_A(m,y) \geq WTP_B(m,y)$ for all (m,y) in the domain of interest then preference ordering A is "more altruistic than" preference ordering B on that domain. In that case, one writes A MAT B.

The partial ordering of opportunity sets is stated here. Suppose that your actions affect my opportunities. If your actions increase my maximum possible payoff then I will regard them as generous unless it is the case that you also increase your own maximum possible payoff even more, in which case I might suspect that your real intention is just to benefit yourself. Consider two opportunity sets G and F. Let m_G^* and m_F^*, respectively, denote my maximum possible payoffs in G and F. Let y_G^* and y_F^* denote your maximum payoffs in the two sets.[2] Opportunity set G is "more generous than" opportunity set F if: (a) $m_G^* - m_F^* \geq 0$ and (b) $m_G^* - m_F^* \geq y_G^* - y_F^*$. In that case, one writes G MGT F. Part (a) states that your choice of G rather than F (weakly) increased my maximum possible payoff. Part (b) states that your choice of G rather than F did *not* increase your maximum payoff more than it did mine, thus clearly revealing your generosity towards me.

The essential property of reciprocal preferences is that my WTP depends on your prior actions that affect me. Revealed altruism theory specifies that an individual's preferences can become more or less altruistic depending on the actions of another agent, as follows. Axiom R states that if you choose my opportunity set G, when you could have chosen set F, and it is the case that G MGT F, then my preferences will become more altruistic towards you. More formally, let A_G denote my preferences when you choose (my) opportunity set G and A_F denote my preferences when you choose set F. Axiom R states: if G MGT F then A_G MAT A_F.

Revealed altruism theory contains a second axiom, Axiom S, that distinguishes between the effects on my preferences of your acts of commission or acts of omission or no opportunity to act. An informal description of Axiom S is that it says that the effect of Axiom R is stronger when a generous act (of commission) overturns the status quo than when an otherwise same act (of omission) merely upholds the status quo and yet stronger still than when there was no opportunity to act. See Cox, Friedman, and Sadiraj (2008, p. 41) for a formal definition of Axiom S.

3. Applications to Private Property Games

The same idiosyncratic features of reciprocal preferences that can rationalize play in public good and common pool games can also rationalize play in some private property games

[2] More formally, for any opportunity set H one has $m_H^* = \sup\{m : \exists y \geq 0 s.t. (m, y) \in H\}$ and $y_H^* = \sup\{y : \exists m \geq 0 s.t. (m, y) \in H\}$.

that deviate from predictions of *homo economicus* theory, most especially when those games involve sequential moves by players in the game. Seeing how *homo reciprocans* theory rationalizes play in private property games that involve fairness considerations yields insight into the theory, so I begin with this type of application.

I present two examples derived, respectively, from extension of the investment game (Berg, Dickhaut and McCabe, 1995) and truncation of the (familiar, textbook) Stackelberg duopoly game. These examples are rich enough to include most central features of economic activities including: competition (Stackelberg), efficiency (investment game), opportunity for collusion (Stackelberg), opportunity for cooperation (investment game), and exposure to defection (both games). They also offer two types of asymmetries: second mover advantage in the investment game versus first mover advantage in the Stackelberg game.

3.1 Investment Game and Dictator Control

My first example is derived from the investment game. In this game, subjects are divided into two groups. Each individual subject is given a $10 endowment. Each (first mover) subject in group A is randomly and anonymously paired with a (second mover) subject in group B. The group B subjects are instructed to keep their $10 endowments. The subjects in group A are informed that each of them, individually, can keep all of their $10 endowment or transfer any integer amount of it to the paired person in group B. Any amount transferred is tripled by the experimenter. The first mover's choice of zero or positive amount to send to the second mover determines the second mover's opportunity set.[3] Subsequently, each subject in group B can keep all of any amount received or return part or all of it to the paired first mover. Each individual makes only one choice (a one-shot game).

It is clear that for this game *homo economicus* theory predicts: (a) second movers will keep all of any tripled amounts sent by first movers; and (b) knowing this, first movers will not send any positive amount. Results from investment game experiments reported in Berg,

---

[3] Let the integer s, such that $0 \le s \le 10$, denote the amount sent by the first mover. Then the second mover's opportunity set contains ordered pairs of integers (m, y) that satisfy the budget constraint $\{(m, y) \in R_+^2 : m + y = 20 + 2s, m \in [10, 10 + 3s]\}$.

Dickhaut, and McCabe (1995) and numerous subsequent papers are that many first movers send positive amounts and many second movers return positive amounts.

A second mover may return a positive amount because she has positively reciprocal preferences or out of unconditional altruism. In order to find out whether altruism or positive reciprocity or both motivations can account for the positive returns, Cox (2004) introduced a dictator control game in which decision makers are given (by the experimenter) the same opportunity sets chosen by first movers in the investment game.[4] In other words, for each amount observed to have been transferred by a first mover in the investment game, a paired dictator game is constructed in which the dictator has the same opportunity set as did the second mover in the investment game. The *homo economicus* model, and models of social preferences in which preferences over distributions of material outcomes are a fixed characteristic of an agent (hereafter "fixed social preferences models") such as Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Cox and Sadiraj (2007), predict that an agent will make the same choice in the dictator control game as in the paired investment game because the opportunity set is identical in the two games. This provides a null hypothesis.

The investment game and second mover dictator control game provide data for a diagnostic test of Axioms R and S of revealed altruism theory. The larger the amount received by the second mover, the more generous is his opportunity set.[5] Hence Axiom R tells us that a larger amount sent by the first mover makes the second mover more altruistic than she is when a smaller amount is sent. In the dictator control game the passive subject has no feasible action whereas a first mover in the investment game can send a positive amount to the second mover by an act of commission. Therefore Axioms R and S, together, predict that second movers in the investment game will return larger amounts than dictators in the control game. This provides an alternative hypothesis. Data from an experiment reject the null hypothesis in favor of the

---

[4] The experimental design also includes another dictator control game to control for *first* mover altruism in order to identify trusting actions.

[5] Consider two amounts that could be sent by the first mover, s_G and s_F such that $10 \geq s_G \geq s_F \geq 0$ that, respectively, determine the second mover's opportunity sets G and F. Note that $m_G^* = 10 + 3s_G$ and $m_F^* = 10 + 3s_F$. Also, $y_G^* = 10 + 2s_G$ and $y_F^* = 10 + 2s_F$. Hence, G MGT F according to the above definition of the partial ordering, MGT.

alternative hypothesis implied by revealed altruism theory (Cox, 2004; Cox, Friedman and Sadiraj, 2008).

3.2 Stackelberg Mini-Game

A second example of a diagnostic test is provided by the Stackelberg mini-game (Cox, Friedman and Sadiraj, 2008) that is derived from the standard (textbook) Stackelberg duopoly game. If, in the standard game, the Stackelberg leader chooses a lower quantity then the Stackelberg follower gets a higher residual demand, which provides the follower with a more generous opportunity set. The *homo economicus* model predicts that the Stackelberg follower will choose the own-profit-maximizing quantity conditional on any residual demand. An experiment with the standard game produced data that systematically deviate from the predictions of the *homo economicus* model in the direction that is consistent with agents' altruism (Cox, Friedman and Sadiraj, 2008). The data are consistent with predictions of both fixed social preferences models and with the model of *homo reciprocans*.

Construction of two Stackelberg mini-games allows one to clearly discriminate between the observable implications of (a) reciprocal preferences and (b) fixed social preferences. Let the leader's opportunity set be $\{x, s\}$ in situation A and $\{s, z\}$ in situation B, where x, s and z are three integers such that $x < s < z$. Choice of quantity s in either situation A or B gives the follower the same residual demand and, hence, the same opportunity set of (m, y) money payoffs for follower and leader. The *homo economicus* model and fixed social preferences models predict that the follower will choose the same quantity of output (and resulting (m, y) money payoffs for follower and leader) when the leader chooses output s in either situation A or situation B. This provides a null hypothesis. Revealed altruism theory has different implications, as follows. Quantity s is the less generous (higher quantity, lower residual demand) choice in situation A but the more generous (lower quantity, higher residual demand) choice in situation B. Axioms R and S of revealed altruism theory predict that the follower will be more altruistic if the leader picks quantity s in situation B than he will if the leader picks the same quantity s in situation A; hence that theory predicts that the follower will choose a lower quantity (which gives the leader a higher price and higher money payoff) in situation B than in situation A in the event that the leader picks quantity s in both situations. This provides an alternative hypothesis.

Data from an experiment reject the null hypothesis in favor of the alternative hypothesis implied by revealed altruism theory (Cox, Friedman and Sadiraj, 2008).

4. Application to Paired Private Property and Common Property Trust Games

The design of the Stackelberg mini-game, which discriminates between the observable implications of reciprocal preferences and of fixed (*homo economicus* or social) distributional preferences, provides intuition for how to design experimental games that provide insight into the question of how private property endowments and common property endowments affect play in trust games. The example of this type of application (Cox, Ostrom, Walker, et al., 2009; Cox and Hall, 2010) is provided by the private property trust (PPT) game and the common property trust (CPT) game, both of which are derived from the investment game.

In the two-person PPT game, the first mover and second mover each have a *private property* endowment of $10, as in the original investment game (see above). The PPT game differs from the investment game in only one way: the second mover can return part or all of her own $10 endowment in addition to the tripled amount received (if any) from the first mover. This change is necessary to make comparisons with the two-person CPT game possible because the second mover is not required to extract any of the common property in the CPT game. The CPT game is the "inverse" game to the PPT game. In the CPT game, each *pair* of subjects is given an endowment of $40 (the maximum amount that can be generated for subject pairs in the investment game and 2-person PPT game) in a *common pool* which they can both appropriate according to the rules of the game. The first mover can withdraw up to $10, in whole dollar amounts, from the common pool and place the amount withdrawn in his private fund. However, each dollar placed in the private fund reduces the value of the common pool by $3. The second mover's decision is how to divide the value of the common pool, that remains after the first mover's extraction, between her private fund and the first mover's private fund.

If in the CPT game the first mover exploits the common pool and places z_{FM} in her private fund (i.e. reduces the value of the common pool by $3z_{FM}$) then the amount left for the second mover to allocate is $\pi_{CPT} = 40 - 3z_{FM}$. If in the PPT game the first mover *leaves* the amount x_{FM} in his private fund (i.e. sends $10 - x_{FM}$ to the second mover) then the amount that becomes available for the second mover to allocate is $\pi_{PPT} = 10 + 3(10 - x_{FM})$. The second mover has the same opportunity set in the CPT and PPT games in the event: ($*$) $z_{FM} = x_{FM}$. Given that condition ($*$) holds, for any feasible amount in the two games, all fixed preferences theories predict that the second mover will choose the same allocation (m^o, y^o) in the CPT and PPT games. In case of the *homo economicus* model, the second mover will, of course, allocate the whole residual fund (remaining after the first mover's extraction) to himself and nothing to the first mover. In case of fixed social preferences models, the second mover may allocate a positive amount to the first mover but the allocation will be the same in the CPT and PPT games because distributional preferences are a fixed characteristic of the agent and, given ($*$), the second mover's opportunity set is the same in the two games.

The *homo reciprocans* preferences in revealed altruism theory have quite different implications for these two games. It is straightforward to use the (above) definition of the MGT partial ordering of opportunity sets to show the following: (a) the second mover's opportunity set in the CPT game is the most generous possible if the first mover does not change the common endowment (i.e. withdraws nothing); (b) each additional dollar that the first mover withdraws in the CPT game makes the second mover's opportunity set incrementally less generous than it was, which according to Axioms R and S makes the second mover less altruistic than he was; (c) the second mover's opportunity set in the PPT game is the least generous possible if the first mover does not change the private endowment (i.e. sends nothing); and (d) each additional dollar that the first mover sends in the PPT game makes the second mover's opportunity set incrementally more generous than it was, which according to Axioms R and S makes the second mover more altruistic than she was. In this way, the theory predicts that the second mover's reciprocal preferences will *not* be the same in the CPT and PPT games in event ($*$) because the (same) opportunity set resulted from an ungenerous change from the endowed opportunity set in CPT and a generous change from the endowed opportunity set in PPT.

The *homo economicus* model and fixed social preferences models provide a null hypothesis: if ($*$) $z_{Fm} = x_{Fm}$ then the second mover will make the same allocations in the CPT and PPT games. The *homo reciprocans* model in revealed altruism theory provides an alternative hypothesis: if (*) $z_{Fm} = x_{Fm}$ then the second mover will allocate more money to the first mover in the PPT game than in the CPT game. Data from the experiment reported by Cox and Hall (2010) reject the null in favor of the alternative hypothesis.[6] In this specific way, reciprocity is observed more with private property endowment than with common property endowment, as predicted by the theory for *homo reciprocans*.

5. Application to Paired Public Good and Common Pool Games

Reciprocal preferences also provide insight into why over-appropriation from a common pool may be more problematic than under-provision of a public good. The example of this type of application (Cox, Ostrom, Sadiraj and Walker, 2012) is provided by provision and appropriation games, both of which are derived from the voluntary contributions public good game.

Typically, in both field and laboratory studies, differences exist in strategy and outcome spaces that make direct comparisons between provision and appropriation social dilemmas infeasible. A precise comparison of the two types of social dilemmas can be made by constructing pairs of payoff equivalent games following the design principle illustrated above by the PPT and CPT games.

In field environments, provision and appropriation social dilemmas often exist within larger economic and political contexts that involve asymmetries in power. This motivates construction of heterogeneous pairs of games that differ (across pairs) in power symmetry or asymmetry. For ease of exposition, we consider games with four players.

[6] Data from the Cox, et al. (2009) experiment do not reject the null hypothesis. The difference between the two experiments is that subjects had to earn their endowments in a laborious task in Cox and Hall (2010) whereas the endowments were a gift from the experimenters in Cox, et al. (2009).

5.1 Provision Games

A provision game is a surplus creation game (as are the investment and PPT games, above). Each of four agents is endowed with E "franks" (experimental currency units) of private property. Each frank contributed to the public good yields 3 franks. Let x_j denote the contribution to the public good by agent j. The frank payoff to agent j equals the amount $E - x_j$ of her endowment that is *not* contributed to the public good plus an equal share (in this case one-fourth) of the tripled amount contributed to the public good by all agents. In the symmetric provision game (Symmetric PG), all agents move simultaneously. Each of the four agents chooses the number of franks to contribute x_j, $j = 1, 2, 3, 4$, from the opportunity set $X = \{0, 1, 2, \cdots, E\}$.

In the boss provision game (Boss PG), three agents simultaneously move first. Subsequently, the boss observes their choices and then decides how much to contribute. Each first mover and the boss choose the number of franks to contribute x_j, $j = 1, 2, 3, 4$, from the same opportunity set X as in the Symmetric PG. In the king provision game (King PG), three agents simultaneously move first. Subsequently, the king observes their choices and then decides how much to contribute to the public good or how much to appropriate from the other three subjects' contributions to the public good. Each of the three first movers chooses the number of franks to contribute x_j, $j = 1, 2, 3$, from the same opportunity set X as in the Symmetric PG. The king can choose to contribute any non-negative number of franks up to his endowment E. Alternatively, the king can choose to take (in integer amounts) any part of the first movers' contributions.

5.2 Appropriation Games

An appropriation game is a surplus destruction game (as is the CPT game above). The common pool is endowed with 12E franks. Each frank extracted from the common pool increases the value of the private fund of the extracting agent by 1 frank while reducing the value of the common pool by 3 franks. The payoff to agent j equals the end value of his private fund plus an equal share (in this case one-fourth) of the remaining value of the common pool after the extractions by all agents. In the symmetric appropriation game (Symmetric AG), all agents move simultaneously. Each of the four agents chooses the number of franks to extract from the opportunity set $Z = \{0,1,2,\cdots,E\}$.

In the boss appropriation game (Boss AG), three agents simultaneously move first. Subsequently, the boss observes their choices and then decides how much to extract. Each of the four agents chooses an amount to extract from the same opportunity set as in the Symmetric AG. In the king appropriation game (King AG), three agents simultaneously move first. Subsequently, the king observes their choices and then decides how much of the remaining common pool to extract. Each of the three first movers chooses an amount to extract z_j, $j = 1,2,3$, from the same opportunity set as in the Symmetric AG. The king can choose to extract any amount up to the total remaining balance of the common pool after the first movers' extractions.

5.3 Implications of Fixed Preferences Models

The feasible allocations and associated payoffs differ across the symmetric, boss, and king versions of the provision game, and they differ across the three versions of the appropriation game. In contrast, the feasible allocations and associated payoffs for all agents are the same within each of the three pairs of games: (1) Symmetric PG and Symmetric AG, (2) Boss PG and Boss AG, and (3) King PG and King AG. If the number of franks *not* contributed to the public good (equal to individual endowment less contribution) equals the number of franks removed from the common pool by each of the four agents then all agents receive the same payoffs in each of the two games in any one of the three pairs of provision and appropriation games.[7] In this way, the two games within each pair are payoff equivalent. The two games within each pair are also strategy equivalent for the *homo economicus* model and fixed social preferences models; hence play is predicted to be the same within each pair of provision and appropriation games by those models. This provides a null hypothesis.

[7] Note that this statement does not preclude heterogeneous play across agents *within* a game; it is a statement about comparison of play *across* paired games.

5.4 Implications of Revealed Altruism Theory

In contrast, the two games in either a boss or king pair of games are not strategy equivalent for the *homo reciprocans* model in revealed altruism theory. The theory is developed in Cox, Freidman and Sadiraj (2008) for two-player extensive form games. Straightforward reinterpretation of the two-player reciprocal preferences model allows us to apply the theory to the second mover's preferences for her own payoffs and payoff of the three first movers in our boss and king treatments by defining the relevant "other player's" payoff as the average payoff of first movers. Define "my income" m as the second mover's money payoff in a game and define "your income" $\bar{y}$ as the average payoff of first movers in the game. Suppose the second mover's preferences for m and $\bar{y}$ satisfy Axiom R and Axiom S stated in section 2. In that case, the implications for second mover play can be derived by reasoning similar to that for the PPG and CPG games in section 4.

Consider the king games. The decisions by the first movers determine the opportunity set of the second mover ("king"). If the first movers don't contribute anything in the king provision game (i.e. don't change their private property endowments), this provides the second mover with the least generous opportunity set possible. Each additional frank that first movers contribute provides the second mover with an incrementally more generous opportunity set. If the first movers don't extract anything in the king appropriation game (i.e. don't change the common pool endowment), this provides the second mover with the most generous opportunity set possible. Each additional frank that first movers extract provides the second mover with an incrementally less generous opportunity set. If on average the number of franks *not* contributed in the provision game (equal to private good endowment less contribution) equals the average number of franks removed from the common pool by the three first movers then the second mover's preferences will be less altruistic in the appropriation game than in the provision game. Therefore, in the event that first movers on average leave the same amount in their private accounts in the king provision game that they place in their private accounts in the king appropriation game, the theory predicts that the second mover will take more money for himself in the appropriation game than in the provision game. This provides an alternative hypothesis.

Data from an experiment (Cox, Ostrom, Sadiraj and Walker, 2012) reject the null hypothesis in favor of the alternative hypothesis implied by revealed altruism theory. The first movers on average leave about the same amounts in their private accounts in the king provision game as first movers place in their private accounts (from extraction) in the king appropriation game. In response, as predicted by the *homo reciprocans* model in revealed altruism theory, second movers allocate significantly more to themselves, and less to the first movers, in the king appropriation game than in the king provision game. This behavior by second movers produces significantly lower efficiency (i.e. realized *total* payoff) in the king appropriation game than in the king provision game. In this specific way, asymmetric power makes over-appropriation from a common pool a more serious problem than under-provision of a public good, as predicted by the theory for *homo reciprocans*.

6. Social Class Differences and Homo Reciprocans: The Indian Caste System

Previous sections describe theory-testing experiments that focus on competition, cooperation, reciprocity, and defection with private goods, public goods and common pools. The subjects were drawn from the student population of American universities. No information was given to subjects about the individual characteristics of others in his/her own experiment group such as their race, sex, age, religion or social class.[8] The types of games used in these theory-testing experiments can also be used for other purposes, for example to study the implications of racial, ethnic, or social class differences for behavior in environments characterized by social dilemmas.

Reciprocal preferences reflect norms for behavior that develop over time in a society. In India, a central feature of the society is the ancient caste system of social stratification and social restriction.[9] The caste system continues to this day, and is especially strong in marital matching (Dugar, Bhattacharya and Reiley, 2011). One question of importance to the economic development of India is how the caste system affects behavior in social dilemma situations. Does

[8] Subjects interacted with 1 or 3 other people. They could look around the laboratory and see that there were 27 – 39 other individuals who looked like students who would be participating in the experiment groups (of size 2 or 4) but they could not know the racial, gender, etc. characteristics of the specific other individuals they would interact with.

[9] The caste system was documented in the travelogues of Fa Hien, a Buddhist monk who visited India circa 400 AD (Keay, 2000). It survived 200 years of British rule and subsequently had political and economic ramifications for independent India (Srinivas, 1957, 1976; Thorat and Newman, 2007).

intra-caste composition of a group ameliorate under-provision of public goods or over-appropriation from common pools? Does *inter*-caste composition of a group exacerbate these efficiency problems?

Cox and Sen (2012) apply the king versions of the provision and appropriation games to research behavioral properties of the caste system in villages in West Bengal, India. The central question is how differences in social caste interact with power asymmetry in social dilemma situations to affect efficiency of allocations with public goods and common pools.

The baseline (king) appropriation and provision treatments are implemented with mixed caste groups of subjects. The subjects are not informed of the caste status of others in these treatments. The caste treatments, implemented with the king appropriation and king provision games, include: (a) one high-caste second mover matched with three low-caste first movers; (b) one low-caste second mover matched with three high-caste first movers; (c) one low-caste second mover matched with three low-caste first movers; and (d) one high-caste second mover matched with three high-caste first movers.[10] The subjects are informed of the caste status of others in these treatments (but not other personal identifiers).

This application is intended to provide new insight into behavior in the social dilemmas embodied in the provision and appropriation games and on ways in which power asymmetry interacts with caste to affect outcomes. For example, how will behaviors change when a low caste person, who is traditionally at the lower end of the stratum, is given greater power while a high caste individual is given less power in the game? How does this compare with outcomes when roles are reversed? How do the high-low, low-high, low-low, and high-high caste treatments compare to the no-caste-information, symmetric treatment?

Data from this experiment show that the effects of caste differences overwhelm the effects of provision vs. appropriation game form on outcomes in the asymmetric power (kings) games. The (high-high) treatments with high caste first movers and a high caste second mover attain high efficiency in both the provision game and the appropriation game. The (high-low)

[10] Hindu society is divided into five distinct castes (and into sub-castes). For purposes of participation in the experiment, "higher" and "lower" castes were combined into two aggregated castes. Subjects were informed that other individuals were members of one or more of the specific castes contained in one of our aggregate categories. The labels "high" and "low" were not used in subject instructions nor in response forms.

treatment with high caste first movers and a low caste second mover, as well as the (low-high) treatment with low caste first movers and a high caste second mover yield lower efficiencies than the high-high treatment. The lowest efficiency is realized in the (low-low) treatment in which both first movers and the second mover are low caste. In the king provision game treatments, the average payoff efficiency in the low-low caste sessions is less than 21% of the payoff efficiency in the high-high sessions. In the king appropriation game treatments, the average payoff efficiency in the low-low caste sessions is less than 26% of the payoff efficiency in the high-high sessions.

This experiment yields insight into effects of strong social hierarchy inequalities on the behavior of *homo reciprocans* in social dilemmas involving provision and appropriation. The data from both asymmetric power (kings) games reveal an absence payoff-enhancing cooperation in both provision and appropriation games when low caste subjects participate. In this way, the legacy of 2,000 years of discrimination show up in this experiment, and the resulting patterns of cooperation failure are manifested in severe under-provision of a public good and over-appropriation from a common pool. These results suggest that the caste system may still be a deterrent to the economic development of India.

7. Summary

Neoclassical theory, as represented by the classic works by Hicks (1946) and Samuelson (1947), has a many-decades-long fruitful application to broad topic areas in economics. By suitable identification of “commodities” (such as your material payoff and mine) this theory provides a model of other-regarding preferences as well as the *homo economicus* model of narrow material self-interest. One defining characteristic of the theory is that preferences are a fixed characteristic of an agent that is independent of other agents’ actions. In this way neoclassical theory does not provide an explanation of reciprocity, by which I mean a pattern of behavior which reflects preferences for which my willingness to pay to increase or decrease your material payoff depends on your previous actions that affect me. By this definition, reciprocity is inconsistent with neoclassical (fixed) preferences over “commodities.” Reciprocity is also inconsistent with models of social preferences such as Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Cox and Sadiraj (2007) in which other-regarding preferences are a fixed characteristic of an agent that is invariant with actions by others.

This paper is concerned with several types of economic games in which behavior observed in experiments systematically differs from predictions of the *homo economicus* model and from predictions of fixed social preferences models. The unifying theme in discussion of all games is the way in which the *homo reciprocans* model in revealed altruism theory (Cox, Friedman, and Sadiraj, 2008) can successfully organize (or "explain") the data. Revealed altruism theory is a generalization of neoclassical preference theory that incorporates reciprocity with a formal explanation of how choices are determined by opportunities and by preferences over payoffs that vary with others' previous actions. The resulting theory includes neoclassical (fixed) preferences as the special case in which my willingness to pay to increase or decrease your payoff is invariant with your previous actions that affect me.

Revealed altruism theory has been applied to a large number of different types of economic games.[11] I here discuss a selection of applications that is broad enough to include private goods, public goods, and common pools and central features of economic activities including competition and reciprocity as well as efficiency, opportunity for collusion, and exposure to defection. Diagnostic tests are reported that discriminate between the implications of models of fixed (*homo economicus* and social) preferences and the implications of the *homo reciprocans* model in revealed altruism theory. These tests are provided by data from experiments with: (a) paired investment and dictator games; (b) paired Stackelberg mini-games; (c) paired private property and common property trust games; and (d) paired appropriation and provision games with sequential moves. The data are generally consistent with the *homo reciprocans* model and inconsistent with fixed preferences models. I also discuss an application of the appropriation and provision games to research on the implications of the Indian caste system for social dilemmas.

[11] See, for examples, the applications to more than 20 games in Cox, Friedman, and Gjerstad (2007), Cox, Friedman, and Sadiraj (2008), Cox, Ostrom, Walker, et al. (2009), Cox and Hall (2010), Cox, Ostrom, Sadiraj, and Walker (2012), and Cox, Servátka, and Vadovič (2012).

References

Berg, Joyce, John Dickhaut, and Kevin McCabe (1995). Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10(1): 122-42.

Bolton, G. E. and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90(1):166–93.

Charness, Gary and Matthew Rabin (2002). "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117: 817–69.

Cox, James C. (2004). "How to Identify Trust and Reciprocity," *Games and Economic Behavior*, 46 (2): 260-281.

Cox, James C., Daniel Friedman, and Steven Gjerstad (2007). "A Tractable Model of Reciprocity and Fairness." *Games and Economic Behavior*, 59(1): 17–45.

Cox, James C., Daniel Friedman, and Vjollca Sadiraj (2008). "Revealed Altruism." *Econometrica*, 76(1): 31–69.

Cox, James C. and Daniel Hall (2010). "Trust with Private and Common Property: Effects of Stronger Property Right Entitlements." *Games* 1(4): 527-550.

Cox, James C., Elinor Ostrom, Vjollca Sadiraj, and James M. Walker (2012). "Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas," working paper, Experimental Economics Center, Georgia State University.

Cox, James C. and Vjollca Sadiraj (2007). "On Modeling Voluntary Contributions to Public Goods." *Public Finance Review*, 35(2): 311–32.

Cox, James C. and Urmimala Sen (2012). "Caste, Efficiency and Fairness with Public Goods and Common Pool Resources," working paper, Experimental Economics Center, Georgia State University.

Cox, James C., Maroš Servátka, and Radovan Vadovič (2012). "Status Quo Effects in Fairness Games: Reciprocal Responses to Acts of Commission vs. Acts of Omission," Experimental Economics Center Working Paper 2012-03, Georgia State University.

Cox, James C., Elinor Ostrom, James M. Walker, Jaime A. Castillo, Eric Coleman, Robert Holahan, Michael Schoon, and Brian Steed (2009). "Trust in Private and Common Property Experiments," *Southern Economic Journal*, 75(4): 957-975.

Dugar, Subhasish**,** Haimanti Bhattacharya and David Reiley (2011). **"**Can't Buy Me Love? A Field Experiment Exploring the Trade-off Between Income and Caste-Status in an Indian Matrimonial Market," forthcoming in *Economic Inquiry*.

Fehr, Ernst and Klaus Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3): 817–68.

Hicks, John R. (1946). *Value and Capital*. Oxford: Clarendon Press.

Keay, John (2000). *India: A History*. HarperCollins: London.

Samuelson, Paul A. (1947). *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

Srinivas, M. N. (1957). "Caste in Modern India,"*Journal of Asian Studies*, 16 (4): 529-548.

Srinivas, M. N. (1976). *The Remembered Village*. University of California Press: Berkeley.

Thorat, S. and Newman, K. (2007). "Caste and Economic Discrimination: Causes, Consequences and Remedies," *Economic and Political Weekly*, XLII (41): 4121-4124.