Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas*

James C. Cox,[1] Elinor Ostrom,[2] Vjollca Sadiraj,[1] and James M. Walker [3]

Abstract:

Social dilemmas characterize decision environments in which individuals' exclusive pursuit of their own material self-interest can produce inefficient allocations. Social dilemmas are most commonly studied in provision games, such as public goods games and trust games, in which the social dilemma can be manifested in foregone opportunities to create surplus. Appropriation games are sometimes used to study social dilemmas which can be manifested in destruction of surplus, as is typical in common-pool resource extraction games. A central question is whether social dilemmas are more serious for inhibiting creation of surplus or in promoting its destruction. This question is addressed in this study with an experiment involving three pairs of payoff-equivalent provision and appropriation games. Some game pairs are symmetric while others involve asymmetric power relationships. We find that play of symmetric provision and appropriation games produces comparable efficiency. In contrast, power asymmetry leads to significantly lower efficiency in an appropriation game than in a payoff-equivalent provision game. This outcome can be rationalized by reciprocal preference theory but not by models of unconditional social preferences.

---

[1]Experimental Economics Center and Department of Economics, Georgia State University
[2] Deceased founder of the Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis, Indiana University and the Center for the Study of Institutional Diversity, Arizona State University
[3] Department of Economics and Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis, Indiana University

* Helpful comments and suggestions were provided by an anonymous referee. Financial support was provided by the National Science Foundation (grant numbers SES–0849590 and SES–0849551).

Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas

1. Introduction

Social dilemmas characterize settings where a divergence exists between expected outcomes from individuals pursuing strategies based on narrow self-interest versus groups pursuing strategies based on the interests of the group as a whole. A large literature in several disciplines studies specific manifestations of social dilemma situations (Axelrod, 1981; Gautschi, 2000; Heibing, et al., 2011; Marshall, 2004). Two prominent areas in the economics literature are public goods games and trust games. These are typically *surplus creation* games in which the central question is whether free riding or absence of trust leads to an opportunity cost that a potential surplus is not created nor provided for a group. For example, in the one period voluntary contributions public good game reported by Walker and Halloran (2004), decision makers on average failed to create 47 percent of the feasible surplus.[1] In the investment (or trust) game reported by Berg, Dickhaut, and McCabe (1995), decision makers on average failed to create 48 percent of the feasible surplus.[2]

The ultimatum game is a well-known *surplus destruction* game; in a typical game, the entire surplus available to the two players is destroyed if the responder rejects a proposed split. In the seminal ultimatum game study reported by Guth, et al. (1982), 10 percent of the feasible surplus was destroyed by “inexperienced” subjects. This figure increased to 29 percent with experienced subjects.[3] Another, well known example of a surplus destruction game is appropriation from a common-pool resource. Walker, et al. (1990), report data for a multiple-decision-round setting where players, on average, over-appropriated to the point of destroying the entire available surplus from the common pool, consistent with the outcome referred to as “the tragedy of the commons.”[4]

An open empirical question is whether social dilemmas are more serious when related to under-provision or over-appropriation in comparable environments. In the field, and most prior laboratory studies, critical differences exist in the opportunity sets that make direct comparisons between provision and appropriation social dilemmas infeasible. We address the question by constructing three pairs of provision and appropriation games. The two games within each pair are payoff equivalent.

In the appropriation game in a payoff-equivalent pair, the value of the total endowment is: (a) strictly greater than the value of the total endowment in the provision game; but (b) equal to the maximum attainable total payoff in the provision game. The endowment in an appropriation game is a Group Fund from which surplus-destroying extractions can be made by participants. The endowments in a provision game are Individual Funds, from which surplus-creating contributions to a Group Fund can be made by participants. One experimental question is whether the theoretical equivalence of these appropriation and provision games fails empirically and, if so, whether efficiency (or realized economic surplus) is lower in the provision game or the appropriation game in a theoretically equivalent pair.

In field environments, institutions for provision and appropriation often exist within larger economic and political contexts that involve asymmetries in power. This motivates the treatments reported herein that focus on the implications of symmetric versus asymmetric power in paired provision and appropriation games.

We examine strategies and outcomes in three pairs of games. Each pair consists of a provision game and an appropriation game that have the same set of feasible allocations and payoffs. The only difference between the two games within a pair is whether the agents' initial endowments are private property or common property. In contrast, pairs of games differ in their types of asymmetry. In the baseline games all N agents move simultaneously. In contrast, in the boss game N-1 "workers" simultaneously move first and the "boss" subsequently determines everyone's payoff after observing the workers' play of the game. The "king" (being sovereign) has even more power: the king not only moves last, after observing the (simultaneous) first moves of the "peasants," he can also appropriate all surplus created in the provision game or not previously destroyed in the appropriation game. The design of the experiment "crosses" the (provision or appropriation) game form treatments with the (baseline or boss or king) power treatments in a 2 X 3 design.

The experimental design provides new insights into the ways in which (a) provision versus appropriation and (b) power symmetry versus power asymmetry affect behavior in environments characterized by social dilemmas. The two games within each pair are payoff equivalent. Self-regarding (or *homo economicus*) preference theory and unconditional models of other-regarding preferences, including the social preference theories of Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Cox and Sadiraj (2007), predict that agents will make choices that yield the same payoffs in the baseline provision and appropriation games. These theories also predict that the boss (resp. king) provision game is equivalent to the boss (resp. king) appropriation game. Unconditional preference theories do not predict that agents will make the same choices in a boss or king game as they do in the baseline game. But the unconditional preference theories do predict that the appropriation and provision games in each (baseline or boss or king) pair of games are equivalent. Reciprocal preference theory has quite different implications. Provision and appropriation games in either of the asymmetric power (boss or king) pairs are not equivalent according to revealed altruism theory (Cox, Friedman, and Sadiraj, 2008). That theory makes specific predictions about how play will differ in an asymmetric (boss or king) *provision* game from the paired asymmetric *appropriation* game.

2. Provision and Appropriation Games with Symmetric and Asymmetric Power

A pair of games consists of a provision game and an appropriation game. The games can be played by any number of agents *N* larger than two. In our baseline games with symmetric power, all *N* players move simultaneously. In the asymmetric power games, *N-1* players simultaneously move first and one player moves second.

2.a. Simultaneous-Move Provision Game

The simultaneous-move provision game (PG) is a contributions game in which *N* agents (simultaneously) choose amounts they will contribute from their endowed Individual Funds to a Group Fund that yields a surplus to be shared equally among all group members. Each agent is endowed with e "tokens" in an Individual Fund and can choose an amount x_j from the feasible set $\{0,1,2,\cdots,e\}$ to contribute to the Group Fund. Contributions to the Group Fund create surplus; each "token" added to the Group Fund decreases the value of the Individual Fund of the contributor by \$1 and increases the value of the Group Fund by \$*M*, where $N > M > 1$. From the perspective of the literature in experimental economics, it is most natural to view the provision game as a linear "voluntary contributions mechanism," where contributions create a non-rival public good. In this case, contributions create a symmetric positive externality to each group member. However, note that one can also view provision as creating a common pool that is shared equally among group members. Of course, the nature of the good (and the interpretation of the decision setting) would be critical if one were to examine the effect of a change in group size. By definition, an increase in group size would not alter the individual externality created by contributions in a pure public good setting, while increasing group size would diminish the individual share of the Group Fund allocated in a common-pool setting.

In summary, each agent is endowed with e tokens worth \$1 each. Each token contributed to the Group Fund yields $\$M$. Let x_j denote the contribution to the Group Fund by agent j. Each of the N agents chooses the number of tokens to contribute x_j, $j = 1, 2, \cdots N$, from the feasible set $\{0, 1, 2, \cdots, e\}$. The dollar payoff to agent i equals the amount of her endowment that is retained in her Individual Fund (i.e. *not* contributed to the Group Fund) plus an equal ($1/N$ share) of M times the total amount contributed to the Group Fund by all agents. There is a social dilemma because $N > M > 1$. The money payoff to a representative agent i can be written as:

$$(1) \quad \pi_i^p = e - x_i + M \sum_{j=1}^{N} x_j / N$$

2.b. Simultaneous-Move Appropriation Game

In the simultaneous-move appropriation game (AG), N agents (simultaneously) decide how much to extract from a Group Fund. The N agents are jointly endowed with $E = Ne$ tokens in a Group Fund that have value $\$ME$. Each agent can choose an amount z_j from the set $\{0, 1, 2, \cdots, e\}$ to extract from the Group Fund. Extractions from the Group Fund destroy surplus; each token removed from the Group Fund increases the Individual Fund of the extractor by \$1 but reduces the value of the Group Fund by $\$M$ where, as above, $N > M > 1$. Agents share equally in the remaining value of the Group Fund after all extractions. Similar to the point made above regarding the provision game, it is most natural to view the appropriation game as a common-pool resource game where, through extraction, agents destroy surplus. Note, however, one could view an appropriation game as one where agents appropriate resources that would have been available to provide a public good. Taking resources prior to public good production destroys the surplus that would have been created.

In summary, the Group Fund is endowed with $E = Ne$ tokens worth $\$M$ each, for a starting total value $\$ME$. Each token extracted from the Group Fund increases the value of the Individual Fund of the extracting agent by \$1 while reducing the value of the Group Fund by $\$M$.

Each of the *N* agents chooses the number of tokens to extract z_j, $j = 1, 2, \cdots, N$, from the feasible set $\{0, 1, 2, \cdots, e\}$. The dollar payoff to agent *i* equals the end value of his Individual Fund plus an equal ($1/N$) share of the remaining value of the Group Fund after the extractions by all agents. As above, there is a social dilemma because $N > M > 1$. The payoff to a representative agent can be written as:

$$(2) \qquad \pi_i^a = z_i + M(E - \sum_{j=1}^{N} z_j)/N$$

2.c. Boss Provision Game (BPG) and Boss Appropriation Game (BAG)

In the BPG, *N*-1 agents ("workers") simultaneously move first to choose the number of tokens they will contribute to the Group Fund. Subsequently, the boss (agent $j = N$) observes their choices and then decides how much, if anything, to contribute. The boss's decision determines all players' final payoffs. Each of the *N* agents chooses the number of tokens to contribute x_j, $j = 1, 2, \cdots N$, from the same feasible set as in the (baseline) PG game.

In the BAG, *N*-1 agents simultaneously move first to choose the number of tokens they will extract from the Group Fund. Subsequently, the boss observes their choices and then decides how much to extract, which determines all players' final payoffs. Each of the *N* agents chooses the amount to extract z_j, $j = 1, 2, \cdots, N$ from the same feasible set as in the (baseline) AG game.

2.d. King Provision Game (KPG) and King Appropriation Game (KAG)

In KPG, *N*-1 agents ("peasants") simultaneously move first. Subsequently, the king (agent $j = N$) observes their choices and then decides how much to contribute or how much to appropriate from the other agents' contributions. Each of the first movers chooses the number of tokens to contribute x_j, $j = 1, 2, \cdots, N-1$ from the same feasible set as in the PG and BPG games. The king can choose to contribute any non-negative number of tokens up to his endowment e to the Group Fund. Alternatively, the king can choose to take (in integer amounts) any part of the tokens contributed by the *N*-1 peasants if it is strictly positive. Define $S_{-N} = -\sum_{j=1}^{N-1} x_j$. The king can choose an amount x_N (to take or contribute) from the feasible set $K_{PG} = \{S_{-N}, S_{-N}+1, S_{-N}+2, \cdots, e\}$.

In KAG, N-1 agents simultaneously move first. Subsequently, the king observes their choices and then decides how many of the remaining tokens (if any) to extract. Each of the N-1 first movers chooses an amount to extract z_j, $j = 1, 2, \cdots, N-1$ from the same feasible set as in the AG and BAG games. The king chooses an amount z_N to extract from the feasible set of integers $\mathrm{K}_{AG} = \{0, 1, 2, \cdots, E - \sum_{j=1}^{N-1} z_j\}$.

3. Theory of Provision and Appropriation Games

Each pair of provision and appropriation games is constructed to be payoff equivalent, as follows. If the amount z_j added to the Individual Fund in the appropriation game equals the amount $e - x_j$ retained in the Individual Fund (i.e. *not* contributed to the Group Fund) in the provision game, for each player $j = 1, 2, \cdots, N$, then the payoff to any agent i is the same in both games.[5] This follows immediately from statements (1) and (2) by noting that they imply $\pi_i^p(\boldsymbol{x})$ = $\pi_i^a(\boldsymbol{z})$, when $\boldsymbol{z} = \boldsymbol{e} - \boldsymbol{x}$, where boldface letters denote vectors representing choices by N agents.

Several testable hypotheses will be derived in this section. Each hypothesis could be stated in terms of total payoffs or average payoff or ending value of the Group Fund because, for our games, each of these measures is a monotonic function of the others, as follows. Define: T = total payoff to all subjects; A = average payoff across subjects; and G = ending *value* of the Group Fund. For our appropriation and provision games, one has $T = Ne + [(M-1)/M]G$ and $A = T/N$. Since $M - 1 > 0$, each measure of outcome from an experimental treatment is monotonically increasing in the other two measures; hence equivalent comparisons of data across experimental treatments can be stated in terms of T, A, or G. In the following, we will phrase hypotheses variously in terms of any one of these three measures which seems most natural in context.

First consider the straightforward implications of *homo economicus* preferences for play in our one-shot games. Given any expectations about play by others, an agent with *homo economicus* preferences will contribute zero to the Group Fund in a provision game and will extract the maximum possible amount from the Group Fund in an appropriation game. This can be formalized as follows.

Proposition 1. Assume *homo economicus* preferences. Agents have a dominant strategy to contribute zero to the Group Fund in a provision game or extract the maximum amount possible from the Group Fund in an appropriation game.

Proof: See appendix 1.

Proposition 1 implies the following testable hypothesis.

> *Hypothesis 1*: Average earnings of players in a provision or appropriation game will be the minimum possible amount $\$e$.

3.a. Simultaneous-Move Provision and Appropriation Games

We now consider implications of other-regarding preferences. As noted above, if players retain the same amounts in their Individual Funds in the provision game as they add to their Individual Funds in the appropriation game they will receive the same payoffs in the two games. Let $\succeq_i^{\Gamma}$ denote player i's preference in game Γ. Payoff equivalence between the games implies that $\boldsymbol{e} - \boldsymbol{g} \succeq_i^{PG} \boldsymbol{e} - \boldsymbol{h}$ for contributions to the Group Fund in the provision game if and only if $\boldsymbol{g} \succeq_i^{AG} \boldsymbol{h}$ for appropriations to the Individual Funds in the appropriation game. This in turn implies that a vector of amounts $\boldsymbol{g}^*$ retained in Individual Funds (i.e. *not* contributed to the Group Fund) is a Nash equilibrium in the provision game if and only if the vector of amounts $\boldsymbol{g^*}$ (appropriated into Individual Funds from the Group Fund) is Nash equilibrium in the appropriation game. An immediate implication of the last statement is that the payoffs attained from a Nash equilibrium is the same for both games. Formally, one has the following proposition.

Proposition 2. Assume either social preferences or *homo economicus* preferences. In the simultaneous-move games, a vector of appropriations $\boldsymbol{g^*}$ into Individual Funds in the appropriation game

a. is a Nash equilibrium if and only if the vector of amounts $\boldsymbol{g^*}$ retained in Individual Funds is a Nash equilibrium in the provision game
b. allocates to player *i* the same payoff in the appropriation game as does the vector of amounts $\boldsymbol{g^*}$ retained in Individual Funds in the provision game.

Proof: See appendix 1.

Note that this proposition states an equivalence between equilibrium payoff vectors in the simultaneous-move provision and appropriation games. It does *not* state that all players make the same choice in either one of the games. Proposition 2 implies the following testable hypothesis.

> *Hypothesis 2*: Average earnings of players are the same in the simultaneous provision and appropriation games.

3.b. Sequential-Move Provision and Appropriation Games with Fixed Preferences

Theoretical properties of sequential-move (boss and king) provision and appropriation games depend on the distinction between fixed preferences and reciprocal preferences. We here consider the implications of fixed preferences models. The most familiar type of fixed preferences are those for the *homo economicus* model in which an agent's preferences vary with her own material payoffs, but are invariant with the material payoffs of others. Other fixed preferences models relax the material selfishness property. Examples are given by models of social preferences such as Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Cox and Sadiraj (2007). In these *homo economicus* and social preferences models, the preferences of an agent are a fixed characteristic of the agent that is independent of the actions of other agents. Fixed preferences models are distinct from models of reciprocal preferences (Cox, Friedman and Gjerstad, 2007; Cox, Friedman, and Sadiraj, 2008) in which an agent's other-regarding preferences can be conditioned on prior actions of others. For fixed (social or *homo economicus*) preferences models, the boss (resp. king) provision game is theoretically equivalent to the boss (resp. king) appropriation game, as stated in Proposition 3.

The second mover (boss or king) is, without loss of generality, denoted as agent N. Let $\boldsymbol{g}_{-N}$ denote the vector of appropriations into Individual Funds by the N-1 first movers in the sequential appropriation game. Let g_N denote the second mover's choice according to the best reply function, $\boldsymbol{br(x_{-N}})$ at information set determined by $\boldsymbol{g}_{-N}$. Let $\boldsymbol{g}_{-N}$ denote the vector of amounts retained in Individual Funds by the N-1 first movers in the sequential provision game, which implies that first movers vector of contributions to the Group Fund is $\boldsymbol{e} - \boldsymbol{g}_{-N}$. Let $e - g_N = \boldsymbol{br(e - g_{-N}})$ the second mover's best response contribution in the Group Fund in the provision game.

Proposition 3. For fixed (*homo economicus* and social) preferences, in the sequential-move games a vector of appropriations $\boldsymbol{g}^*$ into Individual Funds in the appropriation game

a. is an outcome of a Nash equilibrium if and only if the vector of amounts retained $\boldsymbol{g}^*$ in Individual Funds is an outcome of a Nash equilibrium in the provision game
b. allocates to player i the same payoff as the vector of amounts retained $\boldsymbol{g}^*$ in Individual Funds in the provision game.

Proof: See appendix 1.

Proposition 3 implies the following testable hypothesis.

Hypothesis 3: Average earnings of players are the same in the sequential provision and appropriation games.

We have compared the provision game with the appropriation game in each pair of baseline, boss, and king games. Next, we compare the boss provision (resp. appropriation) game with the king provision (resp. appropriation) game. Recall that, for a given (provision or appropriation) game form the only difference between the (sequential) boss game and the (sequential) king game is that at any given information set the boss's opportunity set is a strict subset of the king's opportunity set. In the provision game, the boss can contribute non-negative amounts to the Group Fund whereas the king can either contribute amounts to the Group Fund or remove amounts from the Group Fund that were contributed by first movers. Any amount that the king removes from the Group Fund reduces the sum total payoffs to all players compared to a choice of zero.

For any given choices by the first movers, replacing the choice of a negative amount by a king with the closest amount that meets the non-negativity constraint in a boss game (which is zero) increases the total group payoff and reduces inequality of payoffs. For any given vector of contributions of the first movers, the total group payoff in a boss provision (resp. appropriation) game is not lower than the group payoff in a king provision (resp. appropriation) game. This result is formally presented in the following proposition.

Proposition 4. For sequential-move finite (provision or appropriation) games, for any vector of first movers' choices $\boldsymbol{x}$, the total group payoff from $\boldsymbol{(x,br^{B}(x))}$ in the boss game is (weakly) higher than the total group payoff from $\boldsymbol{(x,br^{K}(x))}$ in the king game.[6]

Proof: See appendix 1.

Proposition 4 implies the following testable hypothesis.

Hypothesis 4: For any given contributions of the first movers, the total group earnings in a sequential king game are not larger than total group earnings in a sequential boss game.

3.c. Sequential-Move Provision and Appropriation Games with Reciprocal Preferences

We have compared provision vs. appropriation game forms and boss vs. king power treatments with fixed (*homo economicus* or social) preferences models. We now consider the implications of reciprocity. Fixed preferences are fundamentally different from reciprocal preferences in which an agent's other-regarding preferences can be dependent on the prior actions of other agents. Cox, Friedman, and Sadiraj (2008) presents a model of reciprocity based on two partial orderings, of opportunity sets and of preferences, and two axioms that link the partial orderings. The theory focuses on how a second mover's willingness to pay (WTP) amounts of her own material payoff to change a first mover's payoff can be affected by the generosity of the first mover's previous actions.

If a first mover's previous action increases the second mover's maximum possible payoff then the second mover will regard him as generous unless it is the case that the first mover increases his own maximum possible payoff even more, in which case the real intention of the first mover may be just to benefit himself. For any given opportunity sets G and F, let m_G^* and m_F^*, respectively, denote a second mover's ("my") maximum possible payoffs in G and F. Let y_G^* and y_F^* denote a first mover's ("your") maximum payoffs in the two sets. Opportunity set G is "more generous than" opportunity set F for me (the second mover) if: (a) $m_G^* - m_F^* \geq 0$ and (b) $m_G^* - m_F^* \geq y_G^* - y_F^*$. In that case, one writes $G \text{ MGT } F$. Part (a) states that a first mover's choice of G rather than F (weakly) increases the second mover's maximum possible payoff. Part (b) states that the first mover's choice of G rather than F did *not* increase his own maximum payoff more than it did the second mover's maximum payoff, thus clearly revealing generosity.

The essential property of reciprocal preferences is that a second mover's WTP can depend on a first mover's prior actions, as represented by Axiom R in Cox, Friedman and Sadiraj (2008). Axiom R states that if you choose my opportunity set G, when you could have chosen set F, and it is the case that $G \text{ MGT } F$, then my preferences will become more altruistic towards you.

Reciprocal preferences can also depend on the distinction between acts of commission and acts of omission or no opportunity to act as in Axiom S of Cox, Friedman, and Sadiraj (2008, p. 41). 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.

An extension to $N > 2$ players of the model in Cox, Friedman, and Sadiraj (2008) shows that play by second movers is predicted to be different in sequential provision games than in sequential appropriation 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 king (resp. boss) appropriation game is the most generous possible if the first movers do *not* change the Group Fund (i.e. they appropriate nothing for their Individual Funds); (b) each additional token that any first mover appropriates in the king (resp. boss) appropriation 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 king (resp. boss) provision game is the least generous possible if the first movers do not change their private endowments (i.e. provide nothing to the Group Fund); and (d) each additional dollar that any first mover provides to the Group Fund 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.

Let $\boldsymbol{g}_{-N}$ be the vector of amounts retained in their Individual Funds (i.e. *not* provided to the Group Fund) by first movers in a sequential provision game and let $\boldsymbol{g}_{-N}$ be the vector of appropriations from the Group Fund in a sequential appropriation game. In that case, Axioms R and S imply that the second mover's reciprocal preferences will *not* be the same in the king (resp. boss) appropriation game as in the king (resp. boss) provision game because the (same) second mover opportunity set resulted from an *ungenerous* change from the endowed opportunity set in the sequential (boss or king) appropriation game and a *generous* change from the endowed opportunity set in the sequential (boss or king) provision game. This intuition is formalized in the following proposition about the second mover's best responses in the sequential provision and appropriation games.

Proposition 5. Let first movers retain $\boldsymbol{g}_{-N}$ in their Individual Funds in the provision game and add $\boldsymbol{g}_{-N}$ to their Individual Funds in the appropriation game. A second mover with reciprocal preferences characterized by Axioms R and S will add more to his Individual Fund in the appropriation game than he retains in his Individual Fund in the provision game.

Proof: See appendix 1.

Proposition 5 implies the following testable hypothesis.

> *Hypothesis 5*: Bosses' (resp. kings') contributions to the Group Fund in the provision game are larger than the amounts they leave in the Group Fund in the appropriation game.

4. Experiment Results

Experiment sessions were conducted at both Georgia State University and Indiana University.[7] In each session, subjects were recruited from subject data bases that included undergraduates from a wide range of disciplines. Via the computer, the subjects were privately and anonymously assigned to four-person groups. No subject could identify which of the others in the room was assigned to their group. Since no information passed across groups, each session involved numerous independent groups. At the beginning of each session, subjects privately read a set of instructions that explained the decision setting. In addition, an experimenter reviewed the instructions publicly. The games described above were operationalized in a one-shot decision setting with a double-blind payoff protocol. The game settings and incentives were induced in the following manner.

In the PG treatment, each individual is endowed with 10 tokens worth \$1 each in what was referred to in the experiments as his/her Individual Fund. The decision task of each individual is whether to move tokens to a Group Fund. Any tokens moved to the Group Fund are tripled in value. Individual earnings equal the end value of the Individual Fund plus one-fourth of the end value of the Group Fund. Second movers in the BPG and KPG treatments are allowed choices as described in section 2.

In the AG treatment, each group is endowed with 40 tokens worth \$3 each in their Group Fund. The decision task of each individual is whether to move tokens to his/her own Individual Fund. Any tokens moved from the Group Fund reduce the value of the Group Fund by \$3, and increase the value of the Individual Fund of the decision maker by \$1. Individual earnings equal the end value of the Individual Fund plus one-fourth of the end value of the Group Fund. The second movers in the BAG and KAG treatments are allowed choices as described in section 2.

The dominant strategy equilibrium for the special case of *homo economicus* preferences would call for each subject to make a zero contribution to the Group Fund in all of the provision treatments. In the appropriation treatments, the equilibrium entails each subject extracting 10 tokens from the Group Fund. In contrast, the group optimum occurs when all tokens are contributed to the Group Fund in a provision game and when no tokens are extracted from the Group Fund in an appropriation game. Unconditional social preferences models predict that the number of tokens retained in Individual Funds in a provision game is the same as the number of tokens moved to Individual Funds in the paired appropriation game.

Data are reported for the number of individual subjects (and four person groups) listed in Table 1. We doubled the initial sample size for the KPG and KAG treatments after observing a striking difference (reported below) between these treatments, in order to ensure that this result was not due to a small sample bias.

The summary presentation of results focuses on: (1) variation in payoffs across the six treatment conditions; and (2) choices by second movers in the four sequential treatment conditions.

4.a. Realized Surplus

The most fundamental issue related to the alternative treatment conditions is the impact of game form on the ability of group members to generate surplus in the three provision games and not to destroy surplus in the three appropriation games. Using each four-member group as the unit of observation, note that both the minimum possible group payoff ($40) and the maximum possible group payoff ($120) are constant across all six treatments. Figure 1 displays average group earnings in the six treatment conditions.

> **Result 1:** Average group earnings across the two baseline conditions (PG and AG) are very similar. Earnings are well above the minimum predicted by the dominant strategy equilibrium for the special case of *homo economicus* preferences (which is $40).

Result 1 is inconsistent with Hypothesis 1 but consistent with Hypothesis 2. The data for the baseline PG treatment are consistent with findings from a large number of linear public goods experiments: the "complete free riding" prediction from the *homo economicus* preferences model fails empirically. Also, the data for the baseline AG treatment are inconsistent with a strong form "tragedy of the commons" prediction, from the self-regarding preferences model, that all available surplus will be destroyed in the appropriation game. This distinct new result is brought to light by our payoff equivalent provision and appropriation games: realized surplus (or efficiency) is nearly the same in the simultaneous provision (public good) and appropriation (common pool) games. In this way we find, in simultaneous games, that free-riding on public good provision is not less nor more of a problem than over-extraction from a common pool.

> **Result 2:** Average earnings are lower in the asymmetric power BPG and BAG treatments than in the symmetric power PG and AG treatments, and are even lower in the asymmetric power KPG and KAG treatments.

The second part of result 2 is consistent with Hypothesis 4. Power asymmetries decrease efficiency (or realized surplus) in both provision and appropriation settings. Low efficiency is especially a feature of the king treatment for the appropriation setting, which is inconsistent with Hypothesis 3 but consistent with Hypothesis 5. Treatment KAG comes closest to manifesting a strong form tragedy of the commons.

> **Result 3:** Pooling across decision groups (n=70), least squares analysis of total allocations to the Group Fund leads to the following results related to selective tests of equality. Group Fund differences between treatments in provision settings are statistically significant for PG vs. BPG and for PG vs. KPG. Group Fund differences between treatments in appropriation settings are significant for AG vs. KAG and for BAG vs. KAG. Group Fund differences are significantly lower for KAG than for KPG.[8]

Lower allocation to the Group Fund in KAG than in BAG is consistent with (an equivalent restatement of) Hypothesis 4. Lower allocation to the Group Fund in KAG than in KPG is inconsistent with (an equivalent restatement of) Hypothesis 3.[9]

4.b. First Mover Decisions

For comparison purposes, the decisions of "first movers" (all subjects in the simultaneous PG and AG games, and those randomly assigned to be first movers in the sequential games) are presented as the dollar amounts allocated to the Group Fund in the provision games or dollar amounts left in the Group Fund in appropriation games. In the notation of section 2, the bar graph in Figure 2 shows the average dollar value across first movers of $3x_j$ in provision games and $30-3z_j$ in appropriation games.

> **Result 4:** Pooling across first mover decisions (n=227), least squares analysis of token allocations to the Group Fund (tokens left in the Group Fund) leads to the following results related to selective tests of equality. Group Fund differences between treatments in provision settings are statistically significant for PG vs. KPG. Group Fund differences between treatments in appropriation settings are significant for AG vs. KAG.[10]

4.c. Second Mover Decisions

Figure 3 displays the decisions of the second movers for the four treatments with sequential decision making. For the boss treatments, decisions are represented as average dollar amounts contributed to the Group Fund (BPG setting) or left in the Group Fund (BAG setting). For the KPG treatment, the bar graph shows the average value across second movers of $3x_4$, where x_4 is a non-negative number of tokens up to 10 or a negative (tokens withdrawn) number up to the maximum number of tokens contributed by the three first movers in her group. For the KAG treatment, the bar graph shows the average across second movers of the amount ($30-3z_4$), where z_4 is the amount withdrawn from the Group Fund.[11]

We next report an analysis of second mover token allocations using treatment dummy variables. Pooling across first mover decision-makers (n=53), a tobit regression of second mover decisions was conducted controlling for the total first mover token allocation FM-SUM to the Group Fund (or left in the Group Fund) and treatment and location dummy variables

> **Result 5:** Only one coefficient estimate is statistically significant, the negative coefficient for the dummy variable for the KAG treatment. The coefficient for the KPG treatment is negative but insignificant.[12]

The significance of the coefficient for the KAG treatment is consistent with implications of reciprocal preferences, as stated in Hypothesis 5, but inconsistent with the implications of fixed preferences stated in Hypothesis 3.[13]

4.d. Economic Significance of Treatment Effects

Table 2 displays average earnings for each treatment condition. Not surprisingly, the actions by second movers in the asymmetric power games are of particular importance in determining final earnings. Consistent with the discussion above, the decisions by second movers in the kings treatments, especially KAG, create a large discrepancy in earnings between first movers and second movers. In the KAG treatment, the average earning of first movers is close to the tragedy of the commons prediction of $10.

In summary, the analysis of data from these experiments suggests that the *opportunity* for second movers to exploit cooperative decisions by first movers: (a) significantly reduces efficiency (or total group payoff); and (b) leads second movers' to exploit this opportunity, in particular in the kings setting.

5. Concluding Remarks

In this paper we report theory and experiments for two types of social dilemmas: provision games and appropriation games. We examine (baseline) symmetric games where all players act at the same time without knowing what others contribute (to a public good) or appropriate (from a common pool) and two types of asymmetric (boss and king) games. In the boss game, three players act first and, subsequently, with knowledge of their decisions the fourth player decides all players' final payoffs by choosing how much to contribute or appropriate. In the king treatment, three players act first, and with knowledge of their decisions, the fourth player decides how much to contribute or take when given the capability of taking everything. While participants do contribute (or refrain from taking) significantly more than predicted for the special-case *homo economicus* interpretation of game theory in symmetric games, average payoffs fall significantly when one player has asymmetric power. The presence of a fourth actor (a "king") who can take resources contributed (to a public good) by others or take resources left (in a common pool) by others has a strong adverse effect on the total payoff in a game. In particular, with a king present in the common-pool appropriation game, we observe an average payoff for first movers that closely approximates the amount predicted for "tragedy of the commons" outcomes.

The experiment provides tests of the different implications of fixed (*homo economicus* or social) preferences models and of a model of reciprocal preferences for behavior in provision (public good) and appropriation (common pool) social dilemmas. The provision and appropriation games in each pair of (baseline, boss, or king) games are, by design, payoff equivalent and have the same Nash equilibria for fixed preferences models. Therefore, such models predict the same outcomes for the provision and appropriation games in each pair of games (although different outcomes across pairs). In contrast, the reciprocity model in revealed altruism theory predicts specific differences in outcomes between the provision and appropriation games in each pair of asymmetric power (boss or king) games. Data from the boss and king treatments are generally consistent with implications of the reciprocity model but inconsistent with implications of the fixed preferences models.

References

Axelrod, Robert (1981). "The Emergence of Cooperation Among Egoists." *American Political Science Review* 75(2): 306-318.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History," *Games and Economic Behavior* 10: 122-142.

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

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

Guth, Werner, Rolf Schmittberger, and Bernd Schwarze, 1982. "An Experimental Analysis of Ultimatum Bargaining, *Journal of Economic Behavior and Organization* 3: 367-388.

Gautschi, Thomas (2000). "History Effects in Social Dilemma Situations." *Rationality and Society* 12(2): 131-162.

Heibing, Dirk, Wenjian Yu, Heiko, Rauhut (2011). "Self-Organization and Emergence in Social Systems: Modeling the Coevolution of Social Environments and Cooperative Behavior." *The Journal of Mathematical Sociology* 35(1): 177-208.

Marshall, Graham R. (2004). "Farmers Cooperating in the Commons? A Study of Collective Action in Salinity Management." *Ecological Economics* 51: 271-286.

Walker, James and Matthew Halloran, 2004. "Rewards and Sanctions and the Provision of Public Goods in One-Shot Settings," *Experimental Economics* 7: 235-247.

Walker, James, Roy Gardner, and Elinor Ostrom, 1990. "Rent Dissipation in Limited Access Common Pool Resource Environments: Experimental Evidence." *Journal of Environmental Economics and Management* 19: 203-211.

Table 1. Number of Individual Subject (and Group) Observations by Treatment

	Simultaneous Games	Boss Games	King Games
Provision Games	32 (8 Groups)	28 (7 Groups)	76 (19 Groups)
Appropriation Games	36 (9 Groups)	32 (8 Groups)	76 (19 Groups)

Table 2. Experimental Earnings by Subject Type and Treatment Condition

	Simultaneous Games	Boss Games		King Games	
		First Mover	Sec. Mover	First Mover	Sec. Mover
Provision Games	$24.19	$19.53	$21.43	$16.43	$21.92
Appropriation Games	$22.39	$20.84	$21.31	$11.04	$22.10

Figure 1. Average Group Earnings by Treatment

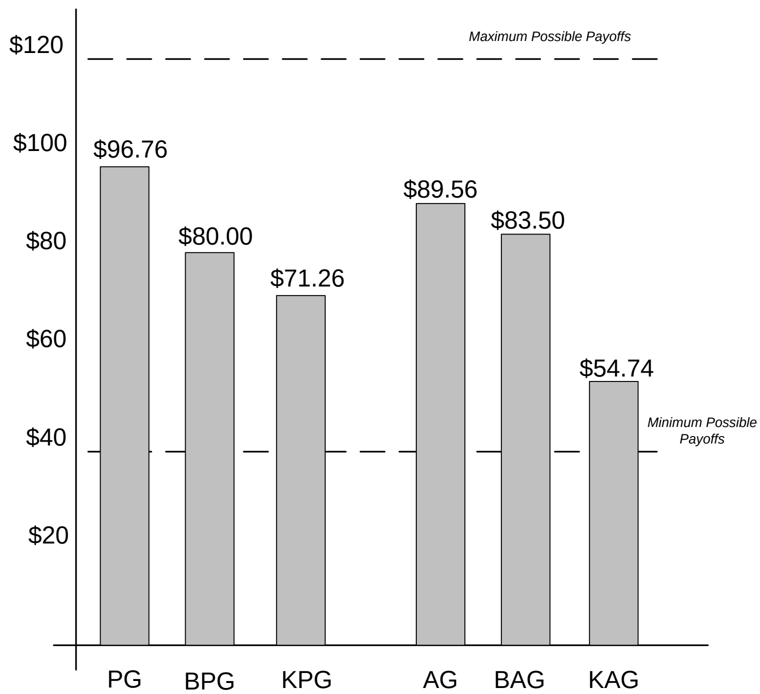


Figure 2. Average First Mover Decisions Represented as $ in Group Fund

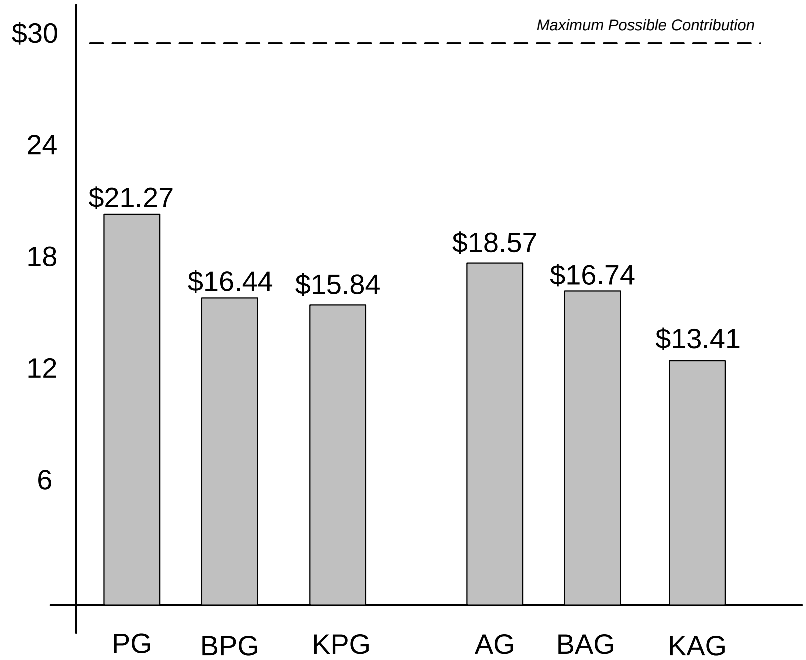
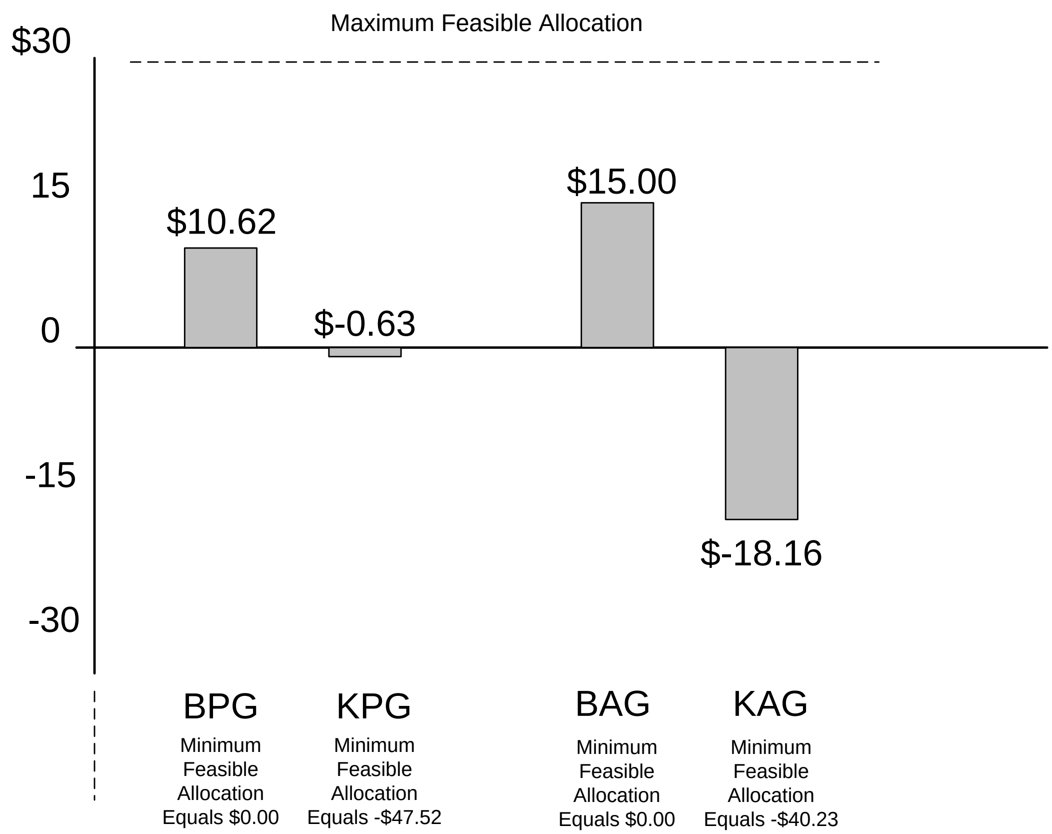


Figure 3. Average Second Mover Decisions Represented as $ in Group Fund



Endnotes

[1] Calculation based on overall baseline VCM, Table 2, page 240.

[2] Calculation based on average amount sent by first movers, Figure 2, page 130.

[3] Calculation based on percent of rejected offers, Tables 4 and 5, page375.

[4] Calculation based on percentage of rents earned in the high endowment setting, Table II, page 208.

[5] Note that the condition is $z_j = e - x_j$, for all j. There is *not* an assumption that x_k and x_j are equal, for $k \neq j$.

[6] Such preferences include common preferences in the literature such as Fehr and Schmidt (1999) and Charness and Rabin (2002).

[7] Complete subject instructions for the experiment are available at http://excen.gsu.edu/jccox/subjects.html .

[8] All linear model analyses are conducted with robust standard errors: PG = BPG, p = .05; PG = KPG, p = .00 ; BPG = KPG, p= .11; AG = BAG, p = .71; AG = KAG, p = .00; BAG = KAG, p= .00; PG = AG, p = .07; BPG = BAG, p = .66; KPG = KAG, p = .04; Lab Location (GSU versus IU), p=.14.

[9] See the discussion in section 3 about the outcome measures G, T, and A for the meaning of "equivalent restatement."

[10] PG = BPG, p = .16; PG = KPG, p = .01; BPG = KPG, p= .52 ; AG = BAG, p = .78; AG = KAG, p = .02; BAG = KAG, p= .13; PG = AG, p = .20; BPG = BAG, p = .92; KPG = KAG, p = .28; Lab Location (GSU versus IU), p=.19.

[11] The maximum amount that can be withdrawn by a second mover in either the KPG or KAG treatment depends on the decisions by first movers in his/her decision making group in the relevant treatment session, as explained in section 2. These amounts are reported below the treatment acronyms at the bottom of Figure 3.

[12] With PG as the baseline setting for the analysis, we find the following parameter estimates and level of significance: DUMBAG: 2.71 (p=.71); DUMKAG: -15.47 (p=.03); DUMKPG: -7.25 (p=.29); DUMIU: 5.99 (p=.21); FMSUM: -0.29 (p=.34); CONSTANT: 10.99 (p=.13).

[13] By design, in KAG, a second mover can make only one decision, how many tokens to remove from the Group Fund. In KPG, a second mover could choose whether to add tokens to the Group Fund from their Individual Fund or remove tokens from the Group Fund. To examine the robustness of our findings, we designed an alternative setting, KPG2. In KPG2, groups began with 10 tokens in the Group Fund rather than zero. Second movers began with zero tokens in their Individual Fund, and second movers made only one decision, how many tokens to remove from the Group Fund. Data from 17 groups was collected. Comparing the data from KPG and KPG2, no significant difference was found in regard to first mover decisions (two sided p-value is 0.102) or second mover decisions (two sided p-value is 0.662).

1 Appendix 1: Proofs of Propositions

Proof of Proposition 1. Let $\beta = M/N$. It follows from the linearity in x_i of the payoff function for individual i

$$u_i(\boldsymbol{\pi}(x_i, \mathbf{x}_{-i})) = v_i(e - x_i + \beta(x_i + \sum_{j \neq i} x_j))$$

and the condition $\beta = M/N < 1$ that for all $\mathbf{x}_{-i}$, $\partial u_i(\boldsymbol{\pi}(x_i, \mathbf{x}_{-i}))/\partial x_i = (\beta - 1)u_i'(\boldsymbol{\pi}(x_i, \mathbf{x}_{-i})) < 0$, for all $x_i \in [0, e]$. Hence, $x_i = 0$ is a dominant strategy for player i in the provision game.

Similarly, referring to payoff specification (2) and the condition $M/N < 1$, one can easily verify that it is dominant to extract e since for all levels of others' extraction $\mathbf{z}_{-i}$, $u_i(\pi_i(z_i, \mathbf{z}_{-i}))$ is increasing in z_i. ■

Proof of Proposition 2. First, note that for all i, the payoff, $\pi_i^a(\mathbf{g})$ of individual i in the appropriation game with extraction vector $\mathbf{g}$ is identical to his payoff, $\pi_i^p(\mathbf{e} - \mathbf{g})$ in the provision game with contribution vector $\mathbf{e} - \mathbf{g}$ as the following shows

$$\begin{aligned} \pi_i^a(\mathbf{g}) &= g_i + \beta(E - \sum_j g_j) = g_i + \beta(Ne - \sum_j g_j) \qquad (1) \\ &= e - (e - g_i) + \beta \sum_j (e - g_j) \\ &= \pi_i^p(\mathbf{e} - \mathbf{g}) \end{aligned}$$

Next, verify that for preferences defined over final payoffs the last statement implies that

$$u_i(\pi_1^a(\mathbf{g}), \pi_2^a(g), ..., \pi_N^a(\mathbf{g})) = u_i(\pi_1^p(\mathbf{e} - \mathbf{g}), \pi_2^p(\mathbf{e} - \mathbf{g}), ..., \pi_N^p(\mathbf{e} - \mathbf{g}))$$

Hence, for all individuals i and for any two vectors of extractions $\mathbf{g}, \mathbf{h} \in [\mathbf{0}, \mathbf{e}]^N$,

$$\mathbf{g} \succeq_i \mathbf{h}$$

in the appropriation game if and only if

$$\mathbf{e} - \mathbf{g} \succeq_i \mathbf{e} - \mathbf{h}$$

in the provision game.

Finally, let $\mathbf{g}^*$ be a Nash equilibrium in the appropriation game. Then, for all i

$$\mathbf{g} \succeq_i (x, \mathbf{g}_{-i}), \text{ for all } x \neq g_i$$

and by the preceding statement

$$\mathbf{e} - \mathbf{g} \succeq_i (e - x, \mathbf{e} - \mathbf{g}_{-i}), \text{ for all } x \neq g_i$$

Thus, $\mathbf{e} - \mathbf{g}$ is a Nash equilibrium in the provision game. Similarly for the reverse implication.

Part (b) follows from appendix statement (1). ■

Proof of Proposition 3. Similar to Proposition 2. ■

Proof of Proposition 4. The proof is straightforward. The opportunity set of the boss is the set of all nonnegative options of the king's opportunity set, hence for any given vector of contributions, $\mathbf{x} \in \mathbf{R}^{N-1}$ of the first movers, if the best reply, $br^K(\mathbf{x})$ in the king game is negative then the best reply, $br^B(\mathbf{x})$ in the boss game is larger, $br^B(\mathbf{x}) > br^K(\mathbf{x})$. For unconditional social preferences (which satisfy the axiom of independence of irrelevant alternatives), if the best reply $br^K(\mathbf{x})$ in the king game is nonnegative then $br^B(\mathbf{x}) = br^K(\mathbf{x})$. Thus, for any given $\mathbf{x} \in \mathbf{R}^{N-1}$ the total group payoffs from $\mathbf{b} = (\mathbf{x}, br^B(\mathbf{x}))$ and $\mathbf{k} = (\mathbf{x}, br^K(\mathbf{x}))$ in the boss and king games satisfy $\sum b_i \geq \sum k_i$. ■

Proof of Proposition 5. Assume reciprocal preferences. Let the vector of appropriations $\mathbf{g}_{-N} \in [0, e]^{N-1}$ be given. Let $\mathbf{e} - \mathbf{g}_{-N} \in [0, e]^{N-1}$ be the vector of contributions in the provision game. It follows from appendix statement (1) that in the payoff space the opportunity sets of a second mover at information set $\mathcal{I}(\mathbf{g}_{-N})$ in the appropriation game and at information set $\mathcal{I}(\mathbf{e} - \mathbf{g}_{-N})$ in the provision game are identical. We need to show that $br^{\mathcal{A}}(\mathbf{g}_{-N}) \geq e - br^{\mathcal{P}}(\mathbf{e} - \mathbf{g}_{-N})$ for an agent with reciprocal preferences. Let g_N be the most preferred choice of a second mover at information set $\mathcal{I}(\mathbf{g}_{-N})$ when the set is chosen by nature. Let the corresponding vector of payoffs be $\boldsymbol{\pi} = (\pi_1(\mathbf{g}), ..., \pi_N(\mathbf{g}))$. In the appropriation game, $\mathcal{I}(\mathbf{0})$ MGT $\mathcal{I}(\mathbf{g}_{-N})$ as follows. First, the largest possible payoff for the second mover is larger at $\mathcal{I}(\mathbf{0})$ than at $\mathcal{I}(\mathbf{g}_{-N})$; it follows from $\overline{\pi}_N(\mathbf{g}_{-N}) = e + \beta(E - \sum_{j<N} g_j - e) < e + \beta(E - e) = \pi_N(\mathbf{0})$. Second, for any first mover $j < N$, $\overline{\pi}_N(\mathbf{g}_{-j}, 0) - \overline{\pi}_N(\mathbf{g}_{-j}, g_j) = \beta g_j \geq \overline{\pi}_j(\mathbf{g}_{-j}, 0) - \overline{\pi}_j(\mathbf{g}_{-j}, g_j) = (\beta - 1) g_j$. By Axioms S and R, the choice $\mathbf{g}_{-N}$ induces less altruistic preferences on the second mover, which requires that the second mover (weakly) decrease the payoffs of others, $\pi_{-N}(\mathbf{g})$; the second mover can do so by (weakly) increasing his (optimal according to unconditional preferences) level of appropriation, hence

$$br^{\mathcal{A}}(\mathbf{g}_{-N}) \geq g_N.$$

On the other hand, in the provision game, $\mathcal{I}(\mathbf{e}-\mathbf{g}_{-N})$ MGT $\mathcal{I}(\mathbf{0})$ and Axioms S and R require that the second mover further (weakly) increase payoffs of others, $\pi_{-N}(\mathbf{e}-\mathbf{g})$; the second mover can do so by increasing his level of contribution

$$br^{\mathcal{P}}(\mathbf{e}-\mathbf{g}_{-N}) \geq e - g_N.$$

The last two inequalities imply that $br^{\mathcal{A}}(\mathbf{g}_{-N}) \geq e - br^{\mathcal{P}}(\mathbf{e}-\mathbf{g}_{-N})$, which completes the proof. ■