

Georgia State University

ScholarWorks @ Georgia State University

ExCEN Working Papers

Experimental Economics Center

8-28-2012

Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas

James Cox
Georgia State University

Elinor Ostrom
Arizona State University

Vjollca Sadiraj
Georgia State University

James Walker
Indiana University

Follow this and additional works at: https://scholarworks.gsu.edu/excen_workingpapers

Recommended Citation

Cox, James; Ostrom, Elinor; Sadiraj, Vjollca; and Walker, James, "Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas" (2012). *ExCEN Working Papers*. 61.
https://scholarworks.gsu.edu/excen_workingpapers/61

This Article is brought to you for free and open access by the Experimental Economics Center at ScholarWorks @ Georgia State University. It has been accepted for inclusion in ExCEN Working Papers by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

Provision versus Appropriation in Symmetric and Asymmetric Social Dilemmas*

James C. Cox,¹ Elinor Ostrom,² Vjollca Sadiraj,¹ and James M. Walker³

August 28, 2012

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 theoretically equivalent provision game. This outcome can be rationalized by reciprocal preference theory but not by models of unconditional social preferences.

¹Experimental Economics Center and Department of Economics, Georgia State University

²Deceased founder of the Workshop in Political Theory and Policy Analysis, Indiana University and the Center for the Study of Institutional Diversity, Arizona State University

³Department of Economics and Workshop in Political Theory and Policy Analysis, Indiana University

* 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; Gauthier, 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 strategy and outcome spaces 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 for models of self-regarding (or “economic man”) preferences and models of unconditional social preferences. As we explain, payoff equivalence implies

isomorphism and efficiency equivalence for these models; therefore, we often simply refer to pairs of “equivalent” provision and appropriation games.

In the appropriation game in an 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 private 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 exist within larger economic and political contexts that often 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. This 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, isomorphic, and efficiency equivalent for unconditional preference theories. Self-regarding (or “economic man”) 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 (Baseline) 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 private 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 a private fund and can choose an amount x_j from the set $X = \{0, 1, 2, \dots, e\}$ 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 private 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 , $j = 1, 2, \dots, N$, from the feasible set $\{0, 1, 2, \dots, e\}$. The dollar payoff to agent i equals the amount of her endowment that is *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:

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

2.b 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. Each agent can choose an amount z_j from the set $Z = \{0, 1, 2, \dots, e\}$ to extract from the group fund. Extractions from the group fund destroy surplus; each token removed from the group fund increases the private 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 surplus of $\$MNe$. Each token extracted from the group fund increases the value of the private 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 , $j=1,2,\dots,N$, from the feasible set $\{0,1,2,\dots,e\}$ to extract from the group fund. The dollar payoff to agent i equals the end value of his private fund plus an equal $(1/N)$ share of the remaining value of the group fund after the extractions by all agents:

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

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 , $j=1,2,\dots,N$, 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 , $j=1,2,\dots,N$ 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 , $j=1,2,\dots,N-1$ 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 $X_{-N} = \sum_{j=1}^{N-1} x_j$. The king can

choose an amount x_N (to take or contribute) from the feasible set

$$K_{PG} = \{X_{-N}, X_{-N} + 1, X_{-N} + 2, \dots, e\}.$$

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 $z_j, j=1,2,\dots,N-1$ 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

$$K_{AG} = \{0,1,2,\dots,E - \sum_{j=1}^{N-1} z_j\}.$$

3. Theory of Provision and Appropriation Games

We use Γ to denote game $(N, \times_i G_i, \{\succeq_i\})$ where N denotes the set of players, G_i denotes the (opportunity) strategy set of player $i, i \in N$, and \succeq_i represents the preferences of player i over $G = \times_i G_i$. As conventionally done, vectors will be represented by letters written in bold font.⁵

3.1 Definitions of Equivalence

We focus on three ways in which the provision and appropriation games could be equivalent: they could be payoff equivalent, isomorphic, or efficiency equivalent. Terms such as “efficiency” and “isomorphism” are frequently used in the literature, although with varying specific interpretations. We begin by providing definitions of terms that we will use in the subsequent theoretical and empirical analysis; these terms include “payoff equivalence,” “isomorphism” and “efficiency equivalence.”

We say that two games, Γ and Φ are payoff equivalent if for each player $i \in N$ there exists some bijection, P_i of strategy set G_i in game Γ to strategy set F_i in game Φ such that the vector of such bijections preserves all players’ payoffs. Thus if we let $\mathbf{P}(\mathbf{g})$ denote $(P_j(g_j) : j \in N)$ one has the following formal definition of payoff equivalence:

Definition 1: Payoff Equivalence. *Two games Γ and Φ are payoff equivalent if for all $i \in N$ there exist bijections, $P_i: G_i \rightarrow F_i$ such that for all $\mathbf{g} \in G, \pi_i(\mathbf{g}) = \pi_i(\mathbf{P}(\mathbf{g}))$.*

It will help to illustrate this definition using the N -player games in this paper. As above, let e be the private fund endowment of each player in the provision game and E be the endowment of the group fund in the appropriation game; by design $E = Ne$. Also, let x_j denote the contribution by agent j to the group fund in the provision game and z_j denote that agent's extraction from the group fund in the appropriation game. It can be easily verified that mappings, P_i of x_i to z_i such that (*) $e - x_i = z_i$ for $i \in N$, that is $P_i(x_i) = e - x_i$ are bijections that satisfy the payoff equalities in Definition 1. In words, the bijection for agent i specifies that the amount $e - x_i$ left in the private fund by agent i in the provision game equals the amount z_i appropriated from the group fund by that agent in the appropriation game. It is straightforward to show (see statements (1) and (2)) that, if (*) $e - x_i = z_i$ for all $i \in N$, then each player receives the same payoff in the provision game as she does in the appropriation game (although different players may receive different payoffs).

Next, we give a formal definition of isomorphic games; these are games for which there exist bijections that preserve preferences of each individual player:

Definition 2: Isomorphism. *Two games Γ and Φ are isomorphic if for all $i \in N$ there exist bijections, $I_i: G_i \rightarrow F_i$ such that for all $\mathbf{g}, \mathbf{h} \in G$, if $\mathbf{g} \succeq_i^\Gamma \mathbf{h}$ then $I(\mathbf{g}) \succeq_i^\Phi I(\mathbf{h})$.*

For games included in this paper, bijections $P_i(x_i) = e - x_i$, also work as illustrations of Definition 2.⁶ It is straightforward to show (see appendix 1) that isomorphism preserves pure strategy Nash equilibria, in the sense that \mathbf{g} is a pure strategy Nash equilibrium in one game if and only if $\mathbf{P}(\mathbf{g})$ is a pure strategy Nash equilibrium in the other game.

Comparisons across pairs of (baseline, boss, or kings) games raise the question of efficiency. Unlike the previous two definitions that are not related to equilibria, the definition of efficiency equivalence applies to equilibrium outcomes. We say that two games are efficiency equivalent if the most efficient equilibria preserve the total payoff. Let the most efficient equilibrium, \mathbf{h}^* in a game be the Nash equilibrium with the largest total payoff to all players, that is

$$\sum_{i=1}^N \pi_i(\mathbf{h}^*) \geq \sum_{i=1}^N \pi_i(\mathbf{g}), \quad \forall \mathbf{g} \in \Omega(\Gamma)$$

where $\Omega(\Gamma)$ is the set of all pure strategy Nash equilibria of Γ .

Definition 3: Efficiency Equivalence. Games Γ and Φ are efficiency equivalent if, for the most efficient equilibria \mathbf{g}^* and \mathbf{h}^* in Γ and Φ , $\sum_{i \in N} \pi_i(\mathbf{g}^*) = \sum_{i \in N} \pi_i(\mathbf{h}^*)$.

3.2 Equivalence of the Simultaneous-Move Provision and Appropriation Games

Statements (1) and (2) imply that if $z_j = e - x_j$ for $j \in N$ then agent i 's payoff in the provision game (π_i^{PG}) equals her payoff in the appropriation game (π_i^{AG}), that is $\pi_i^{PG}(\mathbf{x}) = \pi_i^{AG}(\mathbf{e} - \mathbf{x})$, and this is true for $i \in N$. Since both, (fixed) social preferences and *homo economicus* preferences, are defined on money payoff space, $\mathbf{g} \succeq_i^{PG} \mathbf{h}$ in the provision game if and only if $\mathbf{e} - \mathbf{g} \succeq_i^{AG} \mathbf{e} - \mathbf{h}$ in the appropriation game, as stated in Proposition 1, part b. It is expected from the discussion below Definition 2 then isomorphism and payoff equivalence imply efficiency equivalence.

Proposition 1. The simultaneous-move provision and appropriation games are:

- a. payoff equivalent;
- b. isomorphic; and
- c. efficiency equivalent.

Proof: See appendix 1.

3.3 Equivalence and Non-equivalence of the Sequential-move Games

Theoretical properties of sequential-move (boss and king) provision and appropriation games depend on the distinction between fixed preferences and reciprocal preferences. 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. For such preferences the boss (resp. king) provision game is equivalent to the boss (resp. king) appropriation game, as stated in the following proposition.

Proposition 2. For fixed (*homo economicus* and social) preferences, the sequential-move boss (resp. king) provision and boss (resp. king) appropriation games are:

- d. payoff equivalent;
- e. isomorphic; and
- f. efficiency equivalent.

Proof: See appendix 1.

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. Hence, the cardinality of the king's set of strategies is strictly larger than the cardinality of the boss's set of strategies. This implies that boss and king (provision or appropriation) games are neither isomorphic nor payoff-equivalent because bijections over finite sets preserve cardinality of the set. The remaining question is how they rank with respect to efficiency.

This question can be answered for agents whose preferences satisfy a type of translation invariance in the space of money payoffs. We say that a translation is an "inequality-reducing" one if it decreases the money payoff of the richest players and increases the money payoffs of others. It is "costless" if the total amount by which the richest individuals' payoffs are reduced is not greater than the sum of increased payoffs of others. The translation is "order-preserving" if identity of the richest individuals is preserved by the translation. Let $W(\pi)$ denote the set of richest individuals in

an allocation π and let $W^C(\pi)$ denote the complement of $W(\pi)$ in the set of players N that contains $N-R$ agents. We use notation in which R is the number of richest players (the cardinality of set $W(\pi)$) and, hence, the number of other players is $N-R$ (the cardinality of set $W^C(\pi)$). A formal definition follows.

Definition 4: (utilitarian translation) For any given payoff vector π , a translation $\mathbf{v} = (\mathbf{a}, \mathbf{b})$ such that for all $i \in W^C(\pi)$ and all $j \in W(\pi)$

$$(inequality-reducing) \quad v_i = a > 0 > b = v_j, \forall i, j$$

$$(efficiency-increasing) \quad a(N-R) + bR > 0$$

$$(rank-preserving) \quad a + \max_{i \in W^C(\pi)} \pi_i \leq b + \pi_j$$

is called π -utilitarian translation. The set of such translations is denoted by $T(\pi)$.

We say that agents prefer utilitarian translations (PUT) if for all translations $\mathbf{v} \in T(\pi) \cap T(\pi')$, for all $\pi, \pi' \in R^N$, and for all $i \in W(\pi) \cap W(\pi')$

$$(a) \quad \pi + \mathbf{v} \succ_i \pi$$

$$(b) \quad \pi \succ_i \pi^* \text{ implies } \pi + \mathbf{v} \succ_i \pi^* + \mathbf{v}.$$

This property of preferences means that for any individual who is not among the richest: (a) an allocation that results from a translation $\mathbf{v} \in T(\pi)$ is preferred to the original allocation; and (b) if allocation of payoffs π is preferred to allocation π' then a translation $\mathbf{v} \in T(\pi)$ of π is preferred to the translation $\mathbf{v} \in T(\pi')$ of π' . Preferences that satisfy this property include the inequality-averse preferences in Fehr and Schmidt (1999) and the quasi-maximin preferences in Charness and Rabin (2002).

Proposition 3 below states that king games can't be more efficient than boss games for preferences with property PUT.

Proposition 3. For sequential-move finite (provision or appropriation) games one has:

- a. boss and king games are neither payoff equivalent nor isomorphic; and
- b. for preferences with property (IUT), boss games are (weakly) more efficient than king games.

Proof: See appendix 1.

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 willingness to pay (WTP) amounts of one's own material payoff ("my income", m) to change another's material payoff ("your income", y). A partial ordering of preferences is defined for WTP, as follows. If $WTP_A(m,y) \geq WTP_B(m,y)$ for all (m,y) in some specified domain then preference ordering A is "more altruistic than" preference ordering B on that domain. In that case, one writes $A \text{ MAT } B$. Two different preference orderings, A and B can represent the preferences of two different agents or the preferences of the same agent in two different situations (that differ, for example, with respect to reciprocity-eliciting actions of someone else).

The other partial ordering is of opportunity sets. In sequential games a first mover's ("your") action affects a second mover's ("my") opportunities. If a first mover's 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. 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 $G \text{ MGT } F$ then $A_G \text{ MAT } A_F$.

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 common-pool endowment (i.e. they appropriate nothing for their private 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 $\mathbf{e} - \mathbf{g}_{-N}$ be the vector of amounts left in their private funds (i.e. *not* provided to the group fund) by first movers in a sequential provision game and let \mathbf{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 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 response in the sequential provision game (br^{DPG}) and in the sequential appropriation game (br^{DAG}), where $D = K$ (for king) or B (for boss).

Proposition 4. Assume reciprocal preferences characterized by Axioms R and Axiom S. At any given information set $I(\mathbf{g}_{-N})$ determined by first movers' choices \mathbf{g}_{-N} , the second mover's optimal choice is more efficient in the king (resp. boss) provision game, DPG than in the corresponding king (resp. boss) appropriation game, DAG. That is, for all $\mathbf{g}_{-N} \in \times_{i < N} G_i$, $e - br^{DAG}(\mathbf{e} - \mathbf{g}_{-N}) \leq br^{DPG}(\mathbf{g}_{-N})$, where $D = K$ or B .

Proof: See appendix 1.

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 and remained in these groups throughout the session. 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 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 choice 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 Nash 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 contributed in a provision game is the same as the number of tokens *not* extracted 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 three primary behavioral characteristics of the experiments: (1) efficiency or variation in payoffs across the six treatment conditions; (2) choices by first movers in all treatments; and (3) 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 Nash equilibrium for the special case of *homo economicus* preferences (which is \$40).

Result 1 is consistent with Proposition 1. Also, 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.

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 first part of result 2 is inconsistent with theoretical implications for fixed preferences models, as stated in Proposition 2. 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 consistent with Proposition 3. 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 group allocations to the Group Fund leads to the following results related to selective tests of equality. Earnings differences between treatments in provision settings are statistically significant for PG vs. BPG and for PG vs. KPG. Earnings differences between treatments in appropriation settings are significant for AG vs. KAG and for BAG vs. KAG. Earnings are significantly lower for KAG than for KPG.⁸

4.b. Type X Decisions

For comparison purposes, the decisions of Type X subjects (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 shows the average value across Type X subjects of $3x_j$ in provision games and $30 - z_j$ in appropriation games.

Result 4: Pooling across Type X decision makers (n=227), least squares analysis of Type X 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.⁹

4.c. Type Y Decisions

Figure 3 displays the decisions of the second movers (Type Y) 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 Type Y subjects 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 Type X players in her group. For the KAG treatment, the bar graph show the average across Type Y subjects of the amount $(30 - 3z_4)$, where z_4 is the amount withdrawn from the Group Fund, chosen from the feasible set

$$K_{CPR} = \{0, 3, 6, \dots, 120 - \sum_{j=1}^3 z_j\}.$$
¹⁰

We next report an analysis of Type Y token allocations using treatment dummy variables. Pooling across Type X decision makers (n=53), a tobit regression of Type Y decisions was conducted controlling for the total Type X token allocation (XSUM) to the Group Fund (or total tokens 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. After controlling for treatment effects, the amount first movers place in the Group Fund in provision settings or leave in the Group Fund in appropriation settings is not a significant determinant of second mover choice.¹¹

The insignificance of the coefficient for first movers' contributions is weakly consistent with *homo reciprocans* theory but also consistent with fixed (non-reciprocal) social preferences models. In contrast, the significance of the coefficient for the KAG treatment is consistent with reciprocal preferences, as stated in Proposition 4, but inconsistent with the implications of fixed preferences as stated in Proposition 2.¹²

4.d. Type X and Type Y Earnings

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 treatment, especially KAG, creates a large discrepancy in earnings between first and second movers.

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 first movers' level of cooperation (and resulting efficiency); and (b) leads second movers' to exploit this opportunity, in particular in the kings setting.

5. Concluding Remarks

In this paper we report differences in behavioral outcomes in 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) more than predicted in the special-case *homo economicus* interpretation of game theory in the symmetric condition, 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 surplus appropriation setting, we observe outcomes that closely approximate the “tragedy of the commons.”

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, isomorphic, and efficiency equivalent 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 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

	Symmetric Power	Boss Asymmetric Power	King Asymmetric Power
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

	Symmetric Power Type X	Boss Asymmetric Power Type X Type Y		King Asymmetric Power Type X Type Y	
Provision Games	\$24.19	\$19.53	\$21.43	\$16.43	\$21.92
Appropriation Games	\$22.39	\$20.84	\$21.31	\$11.04	\$22.10

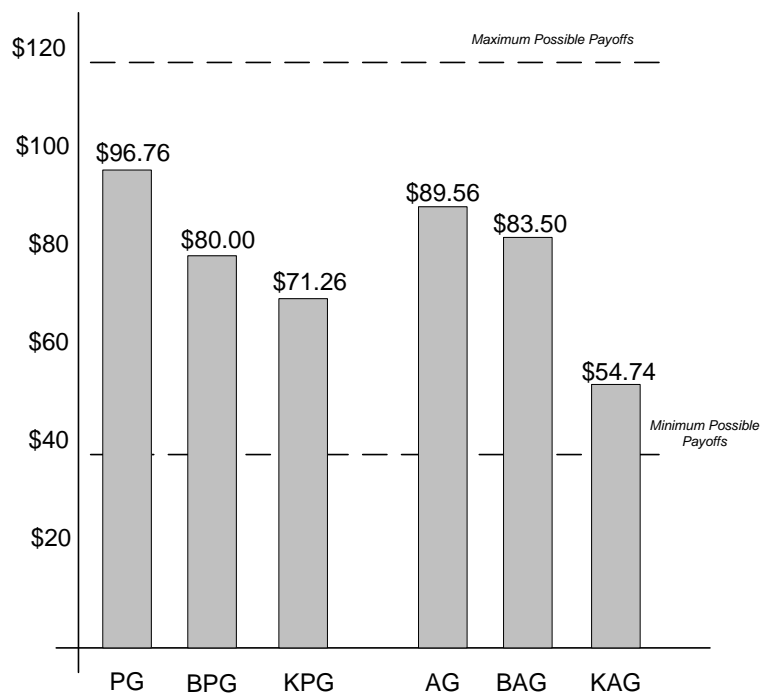
Figure 1. Total Earnings for Groups by Treatment

Figure 2. Average Individual Type X Decisions Represented as \$ in Group Fund

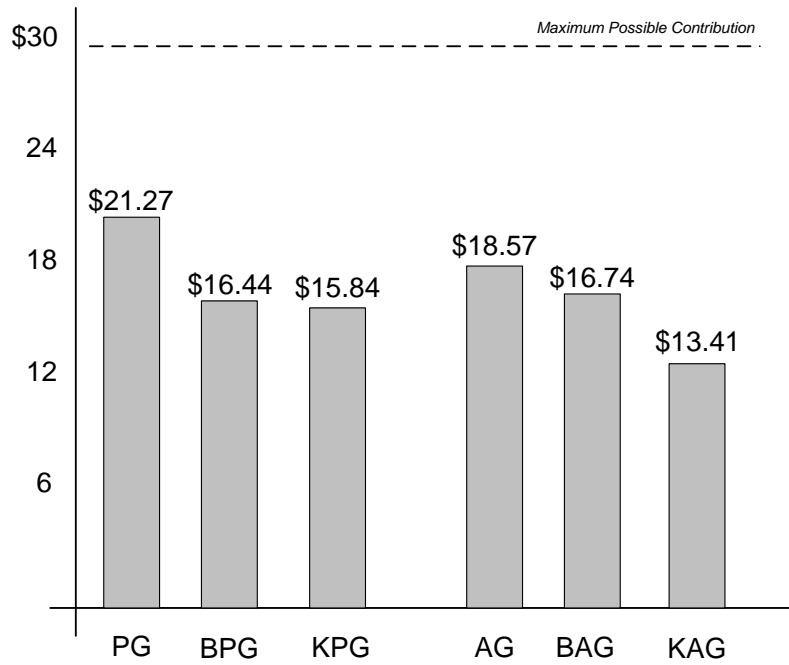
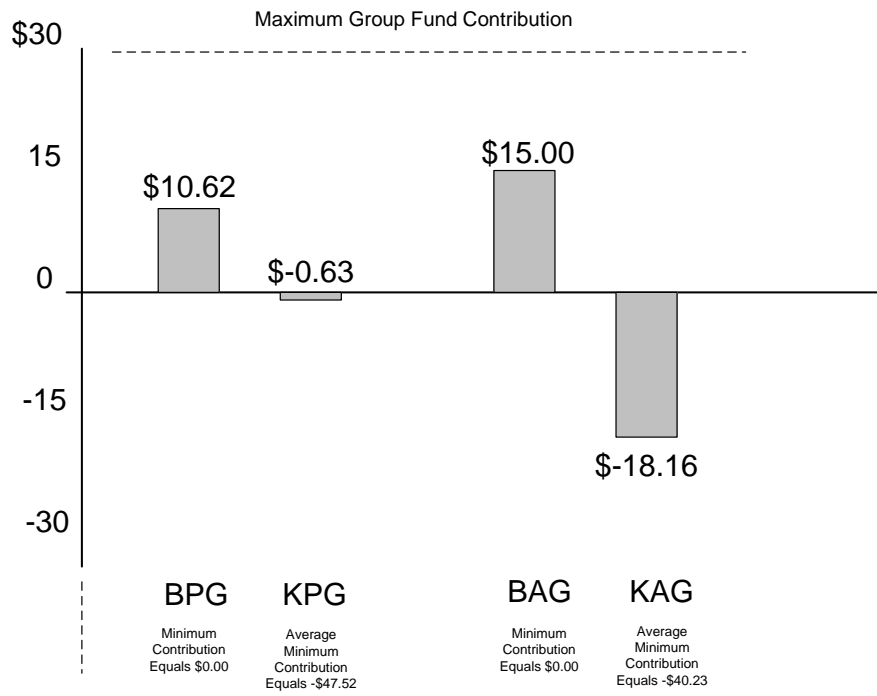


Figure 3. Average Individual Type Y Decisions Represented as \$ in Group Fund



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, page 375.

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

⁵ We will use N also as a notation for the cardinality of the set of players.

⁶ It is easy to show that isomorphic games are not necessarily payoff-equivalent. One example would be splitting \$10 and splitting \$20 on a continuous space; or in a discrete space, splitting \$10 in a one-cent unit of divisibility and splitting \$20 in a two-cent unit of divisibility. The splitting \$10 and \$20 games are clearly isomorphic but not payoff-equivalent.

⁷ Complete subject instructions for the experiment are available at http://excen.gsu.edu/prov_approp_games/.

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 VCM = BVCM, $p = .16$; VCM = KVCM, $p = .01$; BVCM = KVCM, $p = .52$; CPR = BCPR, $p = .78$; CPR = KCPR, $p = .02$; BCPR = KCPR, $p = .13$; VCM = CPR, $p = .20$; BVCM = BCPR, $p = .92$; KVCM = KCPR, $p = .28$; Lab Location (GSU versus IU), $p = .19$.

¹⁰ The maximum amount that can be withdrawn by a Type Y subject in either the KPG or KAG treatment depends on the decisions by Type X subjects in his/her decision making group in the relevant treatment session. The Type X data reported in Figure 2 show the average constraint faced by the Type Y subjects in Figure 3.

¹¹ With PG as the baseline setting for the analysis, we find the following parameter estimates and level of significance: DUMBCPR: 2.71 ($p = .71$); DUMKCPR: -15.47 ($p = .03$); DUMKVCM: -7.25 ($p = .29$); DUMIU: 5.99 ($p = .21$); XSUM: -0.29 ($p = .34$); CONSTANT: 10.99 ($p = .13$).

¹² By design, in KAG, Type Y subjects make only one decision, how many tokens to remove from the Group Fund. In KPG, Type Y subjects chose 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 initial findings, we designed an alternative setting, KPG2. In KPG2, groups began with 10 tokens in the Group Fund rather than zero. Type Y subjects began with zero tokens in their Individual Fund, and Type Y subjects 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 Type X decisions (two sided p -value is 0.102) or Type Y decisions (two sided p -value is 0.662).

1 Appendix 1: Proofs of Propositions

Lemma 1 *Isomorphism preserves pure Nash equilibria.*

Proof. Suppose that games \mathcal{G} and \mathcal{F} are isomorphic: $\mathcal{G} \triangleq \mathcal{F}$. Then, there exists bijective functions, $I_i : G_i \rightarrow F_i$ that preserve preference ordering. Let $\mathcal{N}(\mathcal{G})$ and $\mathcal{N}(\mathcal{F})$ be the sets of pure strategy Nash equilibria of games \mathcal{G} and \mathcal{F} . We need to show that $\mathbf{g} \in \mathcal{N}(\mathcal{G})$ if and only if $\mathbf{f} = \mathbf{I}(\mathbf{g}) \in \mathcal{N}(\mathcal{F})$. Since we are dealing with bijections it suffices to show in one direction. Let $i \in N$ and $h \in G_i$ be given. We show that $\mathbf{f} \succeq_i(h, \mathbf{f}_{-i})$. By surjectivity of $\{I_j\}$ there exists $\mathbf{g}' \in G$ for which $\mathbf{I}(\mathbf{g}') = (h, \mathbf{f}_{-i})$ and by injectivity of $\{I_j\}$, $\mathbf{g}'_j = g_j$ for all $j \neq i$. The last statement and $\mathbf{g} \in \mathcal{N}(\mathcal{G})$ imply that $\mathbf{g} \succeq_i \mathbf{g}'$ and by the definition of isomorphism $\mathbf{I}(\mathbf{g}) \succeq_i \mathbf{I}(\mathbf{g}')$, hence $\mathbf{f} \succeq_i(h, \mathbf{f}_{-i})$. ■

Lemma 2 *For unconditional preferences, payoff equivalence implies isomorphism and efficiency equivalence.*

Proof. Suppose that games \mathcal{G} and \mathcal{F} are payoff equivalent. Then for all $i \in N$ there exists bijections $P_i : G_i \rightarrow F_i$ such that $\mathbf{P} = (P_1, \dots, P_N)$ preserves the payoffs.

We need to show that for all $\mathbf{g}, \mathbf{h} \in G$, $\mathbf{g} \succeq_i \mathbf{h}$ iff $\mathbf{P}(\mathbf{g}) \succeq_i \mathbf{P}(\mathbf{h})$. It follows from payoff equivalence, that for $\mathbf{f} \in \{\mathbf{g}, \mathbf{h}\}$, payoff distribution of $\mathbf{P}(\mathbf{f})$ is identical to the payoff distribution of \mathbf{f} . Since unconditional preferences are defined solely over payoff distributions, for all $i \in N$, $\mathbf{g} \succeq_i^{\mathcal{G}} \mathbf{h}$ and $\{\pi_i(\mathbf{f})\}_{i \in N} = \{\pi_i(P(\mathbf{f}))\}_{i \in N}$, $\mathbf{f} \in \{\mathbf{g}, \mathbf{h}\}$ imply that $\mathbf{P}(\mathbf{g}) \succeq_i \mathbf{P}(\mathbf{h})$.

Next, let \mathbf{g}^* be one of the most efficient Nash equilibria in \mathcal{G} . Then, by Lemma 1, $\mathbf{P}(\mathbf{g}^*)$ is a Nash equilibrium in \mathcal{F} and by payoff equivalence, $\sum_i \pi_i(\mathbf{g}^*) = \sum_i \pi_i(\mathbf{P}(\mathbf{g}^*))$. Thus, there exists a Nash equilibrium in \mathcal{F} with same total payoff as \mathbf{g}^* , hence \mathcal{G} cannot be more efficient than \mathcal{F} . To show that \mathcal{F} cannot be more efficient than \mathcal{G} take $I_i = P_i^{-1}$, $i \in N$ and apply the same arguments. ■

Proof of Proposition 1. Let \mathcal{P} and \mathcal{A} denote a provision game and its corresponding appropriation game. For both games the strategy sets, P and A for each player i satisfy $P = [0, e] = A$. For all $i \in N$, consider function $T_i : P \rightarrow A$ such that $T_i(x) = e - x$.

To show payoff-equivalence consider $\mathbf{T} : \times_i P \rightarrow \times_i A$ such that $\mathbf{T}(\mathbf{g}) = (T_i(g_i)_{i \in N})$. This function is well defined because $(T(g_i))_i \in A^N$. We need to verify that transformation \mathbf{T} is a bijection that preserves payoffs. To show injectivity, take any $\mathbf{g}, \mathbf{h} \in \times_i P$ such that $\mathbf{g} \neq \mathbf{h}$. This means that there exists some i such that $g_i \neq h_i$, hence $T(g_i) = e - g_i \neq e - h_i = T(h_i)$, so $\mathbf{T}(\mathbf{g}) \neq \mathbf{T}(\mathbf{h})$. To show surjectivity for any $\mathbf{f} \in \times_i A$ take $\mathbf{g} = (e - f_1, e - f_2, \dots, e - f_N)$ and verify that $\mathbf{T}(\mathbf{g}) = \mathbf{f}$. Finally, \mathbf{T} preserves payoffs for each player i since by

construction

$$\begin{aligned}
\pi_i(\mathbf{T}(\mathbf{g})) &= T_i(g_i) + \beta(E - \sum_j T_j(g_j)) \\
&= e - g_i + \beta \sum_j (e - (e - g_j)) = e - g_i + \beta \sum_j g_j \\
&= \pi_i(\mathbf{g})
\end{aligned}$$

It follows from payoff equivalence and Lemma 2 that these two games are isomorphic and neither is more efficient.

The proof for the sequential versions of the games is similar. ■

Proof of Proposition 2. Consider a provision game. We show that sequential provision game, \mathcal{P}^s is (weakly) more efficient than the simultaneous provision game, \mathcal{P} , which is the same efficient as the corresponding appropriation game; that is

$$\mathcal{P}^s \geq \mathcal{P} \approx \mathcal{A}$$

Let \mathbf{h}^* be a most efficient (pure) Nash equilibrium of \mathcal{P} . We show that there exists a Nash equilibrium of the sequential game that is the same efficient (in the sense that the total payoff is the same) as \mathbf{h}^* . Without any loss of generality let Y-player be player N and consider the following strategy, y : contribute h_N^* at all information sets. Then, the strategy profile (\mathbf{h}_{-N}^*, y) is a Nash equilibrium in the sequential game and it is the same efficient as, \mathbf{h}^* . Hence, the simultaneous game cannot be more efficient than the sequential provision game.

Similarly,

$$\mathcal{A}^s \geq \mathcal{A} \approx \mathcal{P}$$

Thus, for fixed preferences, sequential games are (weakly) more efficient than simultaneous games. ■

Proof of Proposition 3. Part (1) follows from strategy sets of Y-player being discrete and having different cardinality in the king and boss game whereas bijections preserve cardinality. We show part 2 for the provision games; apply isomorphism (or similar arguments to the ones below) to show that the results hold for the appropriation games as well.

Let a most efficient Nash equilibrium in the king game be

$$\mathbf{K}^* = (\mathbf{x}^*, F^*(.))$$

where $F^*(.)$ is the strategy function of the Y-player. If we let $X(\mathbf{x})$ denote $\sum_{i \leq N} x_i$ then at any information set $\mathcal{I}(\mathbf{x})$ in the king game

$$F^*(\mathbf{x}) \in [-X(\mathbf{x}), e]$$

Let y^* denote Y-player's contribution $F^*(\mathbf{x}^*)$ at the information set $\mathcal{I}(\mathbf{x}^*)$ (i.e. on the equilibrium path) and consider profile

$$\mathbf{k}^* = (\mathbf{x}^*, \kappa(\mathbf{x}))$$

with $\kappa(\cdot)$ such that $\forall \mathbf{x} \in [0, e]^{N-1}$,

$$\begin{aligned}\kappa(\mathbf{x}) &= y^*, \text{ if } y^* \geq -X(\mathbf{x}) \\ &= F^*(\mathbf{x}), \text{ otherwise}\end{aligned}$$

Profile \mathbf{k}^* differs from profile \mathbf{K}^* with respect to Y-choices only in information sets $\mathcal{I}(\mathbf{x})$ in which y^* is a feasible choice; in such sets Y-choices are all y^* . By construction, \mathbf{k}^* is efficiency-equivalent to \mathbf{K}^* and a Nash equilibrium in the king game.

Next, consider strategy profile $\mathbf{b} = (\mathbf{x}^*, \psi(\mathbf{x}))$ in the boss game that differs from profile \mathbf{k}^* only with respect to Y-strategy which is defined as follows $\psi(\mathbf{x}) = \max\{\kappa(\mathbf{x}), 0\}$, $\forall \mathbf{x} \in [0, e]^{N-1}$. The total payoff of \mathbf{b} is not smaller than the total payoff of \mathbf{k}^* since by construction, $\sum \pi_i(\mathbf{b}) \geq \sum \pi_i(\mathbf{k}^*)$. To complete the proof we need to show that for preferences with property (xxxx), \mathbf{b} is a Nash equilibrium in the boss game.

We distinguish between two possible cases: $y^* \geq 0$ and $y^* < 0$. In the former case, $\kappa(\mathbf{x}) \equiv y^* \equiv \psi(\mathbf{x})$, so $\mathbf{b} \equiv \mathbf{k}^*$ is a Nash equilibrium in the boss game as well, which completes the proof.

Suppose that $y^* < 0$. To show that Y-player cannot do strictly better by deviating it suffices to verify that at information set $\mathcal{I}(\mathbf{x}^*)$, $\psi(\mathbf{x}^*) = 0$ is optimal. It follows from \mathbf{k}^* being a Nash equilibrium in the king game that (for well-behaved preferences of Y-player, i.e. $u_N \circ \pi$ concave on player Y's contribution y) the Y-player's utility must be decreasing on $(y^*, e]$. So, if $y^* < 0$ then $0 (= \psi(\mathbf{x}^*))$ is the best choice Y-player can make at $\mathcal{I}(\mathbf{x}^*)$ when the feasible choice set is $[0, e]$. Next, we show that player 1's choice x_1^* is optimal. Let $\pi(g, \mathbf{x}_{-1}, \kappa(g, \mathbf{x}_{-1}))$ denote the payoff vector generated by the vector of contributions $(g, \mathbf{x}_{-1}, F(g, \mathbf{x}_{-1}))$. Formally one has

$$\pi(g, \mathbf{x}_{-1}, \kappa(g, \mathbf{x}_{-1})) = (c + \Phi(g) - g, c + \Phi(g) - \mathbf{x}_{-1}, c + \Phi(g) - \kappa(g, \mathbf{x}_{-1})) \quad (1)$$

where $c = e + \beta \sum_{j=2..N-1} x_j$ and $\Phi(g) = \beta(g + \kappa(g, \mathbf{x}_{-1}))$. Strategy profile \mathbf{k}^* is a

Nash equilibrium in the king game, so

$$(\mathbf{x}^*, y^*) \succeq_1 (g, \mathbf{x}_{-1}^*, \kappa(g, \mathbf{x}_{-1}^*)), \forall g \in [0, e] \quad (2)$$

We need to show that

$$(\mathbf{x}^*, 0) = (\mathbf{x}^*, \psi(\mathbf{x}^*)) \succeq_1 (g, \mathbf{x}_{-1}^*, \psi(g, \mathbf{x}_{-1}^*)), \forall g \in [0, e] \quad (3)$$

since by construction, $\kappa(\mathbf{x}^*) = y^* (< 0)$ and $\psi(\mathbf{x}^*) = 0$. Let $g \in [0, e]$ be given. There can be two cases only: (a) $\kappa(g, \mathbf{x}_{-1}^*) = y^*$ and (b) $\kappa(g, \mathbf{x}_{-1}^*) > y^*$.

Case (a): $\kappa(g, \mathbf{x}_{-1}^*) = y^* (< 0)$. Statement (2) can be equivalently rewritten as for all $g \in [0, e]$

$$(\mathbf{x}^*, y^*) \succeq_1 (g, \mathbf{x}_{-1}^*, y^*), \text{ for all } g \in [0, e]$$

whereas statement (3) is simply

$$(\mathbf{x}^*, 0) \succeq_1 (g, \mathbf{x}_{-1}^*, 0), \text{ for all } g \in [0, e] \quad (4)$$

Let $\mathbf{v}(y) = (-\beta y, \dots, -\beta y, (1-\beta)y) \in R^N$, and use (1) to verify that $\pi(\mathbf{x}^*, 0) = \pi(\mathbf{x}^*, y^*) + \mathbf{v}(y^*)$, $\pi(g, \mathbf{x}_{-1}^*, 0) = \pi(g, \mathbf{x}_{-1}^*, y^*) + \mathbf{v}(y^*)$, and apply supposition PUT¹ to show that $(\mathbf{x}^*, y^*) \succeq_1 (g, \mathbf{x}_{-1}^*, y^*)$ implies $(\mathbf{x}^*, 0) \succeq_1 (g, \mathbf{x}_{-1}^*, 0)$.

Case (b): $\kappa(g, \mathbf{x}_{-1}^*) > y^*$. If $\kappa(g, \mathbf{x}_{-1}^*) \leq 0$ then

$$(\mathbf{x}^*, \kappa(g, \mathbf{x}_{-1}^*)) \succeq_1 (\mathbf{x}^*, y^*) \succeq_1 (g, \mathbf{x}_{-1}^*, \kappa(g, \mathbf{x}_{-1}^*)) \quad (5)$$

where the first weak preference statement follows from $\pi(\mathbf{x}^*, \kappa(g, \mathbf{x}_{-1}^*)) = \pi(\mathbf{x}^*, y^*) + \mathbf{v}(y^* - \kappa(g, \mathbf{x}_{-1}^*))$ and supposition PUT whereas the second one follows from statement (2). In this case $\psi(g, \mathbf{x}_{-1}^*) = 0$ and verify that $\pi(\mathbf{x}^*, 0) = \pi(g, \mathbf{x}_{-1}^*, \kappa(g, \mathbf{x}_{-1}^*)) + \mathbf{v}(\kappa(g, \mathbf{x}_{-1}^*))$, $\pi(g, \mathbf{x}_{-1}^*, 0) = \pi(g, \mathbf{x}_{-1}^*, \kappa(g, \mathbf{x}_{-1}^*)) + \mathbf{v}(\kappa(g, \mathbf{x}_{-1}^*))$, and apply supposition PUT² to show that statement 5 implies statement (3),

$$(\mathbf{x}^*, 0) \succeq_1 (g, \mathbf{x}_{-1}^*, 0) = (g, \mathbf{x}_{-1}^*, \psi(g, \mathbf{x}_{-1}^*)).$$

If $\kappa(g, \mathbf{x}_{-1}^*) > 0$ then $\kappa(g, \mathbf{x}_{-1}^*) = \psi(g, \mathbf{x}_{-1}^*)$ and one has

$$(\mathbf{x}^*, 0) \succeq_1 (\mathbf{x}^*, y^*) \succeq_1 (g, \mathbf{x}_{-1}^*, \kappa(g, \mathbf{x}_{-1}^*)) = (g, \mathbf{x}_{-1}^*, \psi(g, \mathbf{x}_{-1}^*))$$

where the first preference follows from $\pi(\mathbf{x}^*, 0) = \pi(\mathbf{x}^*, y^*) + \mathbf{v}(y^*)$ and supposition PUT whereas the second one follows from statement (2).

Thus, player 1 cannot do strictly better by deviating. Similarly for all other X -players. ■

Proof of Proposition 4. Assume reciprocal preferences. Let $\mathbf{g}_{-N} \in \prod_{i < N} G_i$, be given. In the payoff space, the opportunity sets of Y -player at information set $\mathcal{I}(\mathbf{g}_{-N})$ in the provision game and at information set $\mathcal{I}(\mathbf{e} - \mathbf{g}_{-N})$ in the appropriation game are identical. We need to show that $br^{\mathcal{P}}(\mathbf{g}_{-N}) > \mathbf{e} - br^{\mathcal{A}}(\mathbf{e} - \mathbf{g}_{-N})$. Let g_N be the most preferred choice of Y -player when information set $\mathcal{I}(\mathbf{g}_{-N})$ is not chosen by X -players, that is g_N is the optimal choice for fixed preferences. Let the corresponding vector of payoffs be $\boldsymbol{\pi} = (\pi_1(\mathbf{g}), \dots, \pi_N(\mathbf{g}))$. In the provision game, $\mathcal{I}(\mathbf{g}_{-N})$ *MGT* $\mathcal{I}(\mathbf{0})$, because (a) the largest possible payoff for Y -player is larger at $\mathcal{I}(\mathbf{g}_{-N})$ than at $\mathcal{I}(\mathbf{0})$, $\bar{\pi}_N(\mathbf{g}_{-N}) = e + \beta \sum_i g_i > e = \pi_N(\mathbf{0})$ and that (b) for any X -player $j < N$, $\bar{\pi}_N(\mathbf{g}_{-j}, g_j) - \bar{\pi}_N(\mathbf{g}_{-j}, 0) = \beta g_j \geq \bar{\pi}_j(\mathbf{g}_{-j}, g_j) - \bar{\pi}_j(\mathbf{g}_{-j}, 0) = (\beta - 1)g_j$. By Axioms *S* and *R*, choice \mathbf{g}_{-N} induces *MAT* preferences on Y -player, which require that he (weakly) reduces his optimal payoff (according to unconditional preferences), $\pi_N(\mathbf{g})$ to (weakly) increase the payoffs, $\pi_i(\mathbf{g})$ of X -player i . Y -player can do so by increasing his contribution g_N , hence,

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

¹ Definition (4) applies since $y^* \leq 0$ and $-y^*(N-1)\beta + y^*(1-\beta) \geq 0$.

² Definition (4) applies since $\kappa(g, \mathbf{x}_{-1}^*) \leq 0$ and $-\kappa(g, \mathbf{x}_{-1}^*)(N-1)\beta + \kappa(g, \mathbf{x}_{-1}^*)(1-\beta) \geq 0$.

On the other hand, in the appropriation game, $\mathcal{I}(\mathbf{e} - \mathbf{0}) \text{ MGT } \mathcal{I}(\mathbf{e} - \mathbf{g}_{-N})$, and Axioms S and R require that player Y (weakly) decrease the payoffs of others, $\pi_{-N}(\mathbf{g})$; Y -player can do so by increasing his appropriation

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

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