

Power and efficiency in pillage games

by

J.S. Jordan¹

The Pennsylvania State University

November 11, 2007

¹ Helpful comments by Colin Rowat are gratefully acknowledged. This draft has cryptic exposition and omits some more recent results

Abstract

A pillage game is a formal model of Hobbesian anarchy as a coalitional game. The technology of pillage is specified by a power function that determines the power of each coalition as a function of its members and their wealth. A coalition can despoil any other coalition less powerful than itself. The present paper studies the problem of achieving an efficient allocation of resources when the required reallocation changes the distribution of power. For example, land redistribution may increase total production, but may also deprive the original owners of the power they need to compel compensation. In this case the original owners would block the redistribution. Previous work on pillage games has focused on the stable set (von Neumann-Morgenstern solution) as a representation of a stable balance of power. However, the balance of power is typically too delicate to support a efficient allocations. In particular stable sets are necessarily finite. The present paper shows that for certain power functions, a recently developed extension of the stable set, called the legitimate set, can be rich enough to support efficient reallocation with enforceable compensation.

1. Introduction

Pillage games are introduced in Jordan (2006a) to provide a formal model of Hobbesian anarchy as a coalitional game without commitment. In the economic environment of a pillage game, there is only one commodity, wealth, which is allocated among a finite number of players. Since wealth is desired by all, reallocation can only be effected by force. The technology of force is specified by a power function that determines the power of each coalition as a function of its members and their wealth. A coalition can take the wealth of any coalition less powerful than itself. Pillage is costless and certain, but due to the absence of a commitment technology, it can also be treacherous. A pillage game is a natural setting for the study of power, but the zero-sum property limits its economic interest. In particular, every allocation in a pillage game is trivially Pareto efficient.

Pillage games are extended to a more interesting economic environment by Polemarchakis and Rowat (2007). They add a second commodity to the environment, allowing players to engage in Pareto improving exchange as well as pillage. However, the exchange of goods can change the distribution of power. Polemarchakis and Rowat (2007) show that the maintenance of a balance of power can constrain the achievement of Pareto efficiency.

The present paper extends the pillage game environment by introducing production. Instead of consuming wealth directly, the players use wealth as an input in the production of consumption. Moreover, consumption can be reallocated without reallocating wealth, which implies that efficiency is only achieved when wealth is allocated to maximize aggregate output. For example, if the production technology available to each player is identical and strictly concave, efficiency requires wealth to be allocated equally among the players. However, the power of each coalition is assumed to depend on wealth, as in the basic pillage game environment, rather than consumption. This can inhibit reallocations of wealth needed to achieve efficiency. For example, suppose an initial allocation gives all wealth to one player, who consumes the entire output. This player would be willing to reallocate the wealth equally in return for the assurance of receiving consumption equal to the original output plus a portion of the efficiency gain. However, this could require the other players to give up most of their post-reallocation output. Depending on the power function, this may be unenforceable. In particular, an example is given in Section 3 showing that an economic pillage game may fail to have an efficient stable set (von Neumann - Morgenstern solution).

A stable set of a zero-sum pillage game is shown in Jordan (2006a) to have a dynamic representation as a farsighted core under rational expectations, which answers a critique of the stable set concept due to Harsanyi (1974). However, stable sets of economic pillage games may fail to have this property, as shown by an example in Section 3, making them vulnerable to the Harsanyi critique. Zero-sum pillage games are extended in Jordan (2006b) by adding extrinsic variables that can record historical information, which makes possible a much richer class of farsighted cores. A set of allocations that can be represented as a farsighted core in this fashion is called *legitimate*. It is shown in section 4 that for a large class of power functions, the entire set of efficient allocations is legitimate. However, the economic pillage game that is shown in Section 3 to have no efficient stable set is also shown in Section 4 to have no efficient legitimate set.

2. Economic pillage games

The pillage games introduced in Jordan (2006) involve a single commodity, wealth, to be allocated among a finite number of players. The total amount of wealth is fixed, so every allocation is trivially Pareto efficient. In order to introduce the issue of economic efficiency, the present paper extends the basic pillage game in two ways. First, wealth, instead of being consumed directly, is assumed to be an input in the production of the single consumption good. In particular, each player is assumed to possess a technology that produces consumption as a strictly increasing function of the player's individual wealth. Since individual production is strictly increasing in individual wealth, this extension by itself merely reinterprets the basic pillage game without changing it. The second extension allows consumption to be reallocated without reallocating wealth. The implication of these two extensions is that Pareto efficiency requires wealth to be allocation to maximize aggregate production.

2.1 Allocations: The set of players is the finite set $I = \{1, \dots, n\}$, where $n \geq 2$. There is a fixed amount of wealth, normalized to unity, to be allocated among the players. A wealth allocation is a vector $w = (w_1, \dots, w_n) \in \mathbb{R}_+^I$ satisfying $\sum_i w_i = 1$. Each player i has a strictly increasing production function $f_i : [0, 1] \rightarrow \mathbb{R}_+$ that produces consumption as a function of player i 's allocation of wealth. We assume that $f_i(0) = 0$ for each i . Any allocation among the players of the total consumption produced is feasible. Thus the set of allocations is $A = \{(w, x) \in \mathbb{R}_+^I \times \mathbb{R}_+^I : \sum_i x_i = \sum_i f_i(w_i)\}$. For each wealth allocation w , let $f(w)$ denote the vector of outputs $f(w) = (f_1(w_1), \dots, f_n(w_n))$. Then $(w, f(w)) \in A$. An allocation (w, x) is (Pareto) *efficient* if there does not exist an allocation (w', x') satisfying $x'_i > x_i$ for all i . Let P denote the set of efficient allocations. It follows that $P = \{(w, x) \in A : \sum f_i(w_i) \geq \sum_i f_i(w'_i) \text{ for all } w' \in \mathbb{R}_+^I \text{ satisfying } \sum_i w'_i = 1\}$.

Since production is a function of individual wealth, there are no externalities in production among players. Wealth can be interpreted as land, and each player i can be interpreted as having a fixed amount of labor, which is applied without disutility to player i 's allocation of land, w_i . The resulting output, $f_i(w_i)$, can reflect productivity differences among players.

Wealth and consumption can be reallocated voluntarily or by force. The possibilities of reallocation by force are determined by a technology of power, which is represented by a power function as defined below.

2.2 Power: Subsets of I will be called *coalitions*. Let \mathcal{C} denote the set of coalitions. A *power function* is a function $\pi : \mathcal{C} \times \mathbb{R}_+^I \rightarrow \mathbb{R}_+$ satisfying

- p.1a) for all w , if $C \subset C'$ then $\pi(C', w) \geq \pi(C, w)$;
- p.1b) for all w , if $C \subset C'$ and there is some $i \in C' \setminus C$ such that $\pi(\{i\}, w) > 0$, then $\pi(C', w) > \pi(C, w)$;
- p.2a) if $w'_i \geq w_i$ for all $i \in C$ then $\pi(C, w') \geq \pi(C, w)$; and
- p.2b) if $C \neq \emptyset$ and $w'_i > w_i$ for all $i \in C$ then $\pi(C, w') > \pi(C, w)$.

We will also assume for convenience that $\pi(\emptyset, w) = 0$ for all w . Together with the above properties, this implies that $\pi(C, w) > 0$ if C contains at least one player i with $w_i > 0$.

The power of a coalition depends on the allocation (w, x) through the allocation of wealth, w , but not through the allocation of consumption. Since production is strictly increasing in wealth, this is equivalent to assuming that power depends on the output vector $(f_1(w_1), \dots, f_n(w_n))$, which is the interpretation we will commonly adopt. This reflects the interpretation that player i exercises direct control over both w_i and the output $f_i(w_i)$. This interpretation is made more precise in Definition 2.5 below.

Power functions defined in Jordan (2006) are required to satisfy (p.1a), and (p.2a,b) but not (p.1b). This additional requirement is used in Theorem 4.?? The function $\pi(C, w) = \max\{w_i : i \in C\}$ is an example of a power function that violates (p.1b) but satisfies the other three conditions.

An economic pillage game is completely specified by the player set I , the production functions f_1, \dots, f_n and the power function π .

2.3 Reallocation by force: An allocation (w', x') *dominates* an allocation (w, x) by *pillage*, written $(w', x') \succ_p (w, x)$, if

$$\pi(W, w) > \pi(L, w),$$

where $W = \{i : x'_i > x_i\}$, and $L = \{i : x'_i < x_i\} \cup \{i : w'_i \neq w_i \text{ and } x'_i = x_i\}$.

Any transformation of π that leaves the pillage relation \succ_p unchanged has no effect on the solution concepts considered here.

The coalition of winners, W , consists of all players whose consumption is increased by the act of pillage. The coalition of losers, which unsuccessfully opposes the act of pillage, includes all players whose consumption is decreased. In addition, we assume that any player whose wealth is changed will oppose the pillage in the absence of an increase in consumption. This reflects our assumption throughout that player i will voluntarily accept a change in the allocation (w_i, x_i) only if the change increases x_i . A player whose allocation remains unchanged is assumed to remain neutral. This *neutrality assumption* will be modified below as appropriate when players are assumed to anticipate the future consequences of any change in the current allocation.

Pillage is interpreted as an act of force, but the domination relation \succ_p formally includes changes in the allocation that are unopposed because everyone involved is made better off, that is, $L = \emptyset$. It will be useful to formally define this subset of the pillage relation.

2.4 Reallocation by voluntary transaction: An allocation (w', x') *dominates* an allocation (w, x) by *voluntary transaction*, written $(w', x') \succ_t (w, x)$, if $w' \neq w$, $x' \geq x$ and $x'_i > x_i$ for each i such that $w'_i \neq w_i$. To verify that $\succ_t \subset \succ_p$, let $(w', x') \succ_t (w, x)$. The definition of \succ_t implies that $L = \emptyset$. There must be at least one player i with $w_i > 0$ and $w'_i \neq w_i$, so $i \in W$. Hence $\pi(w, w) > \pi(L, w)$, so $(w', x') \succ_p (w, x)$.

As mentioned above, the assumption that power depends on wealth rather than consumption is based on our assumption that player i exercises direct control over the output $f_i(w_i)$. For example, consider an allocation (w, x) at which $x_i < f_i(w_i)$. This will be interpreted as player i paying the difference $f_i(w_i) - x_i$ as rent to another player or players. This could result from the use or threat of force by others against player i , or because player i received part of the wealth w_i as a voluntary exchange from other players in return for the rent. However, we assume that, as a technological matter, player i has the ability to retain the entire output $f_i(w_i)$.

2.5 Reallocation by retaining rent: Given an allocation (w, x) , define the set $R = \{i : x_i < f_i(w_i)\}$. An allocation (w', x') *dominates* an allocation (w, x) *by rent retention*, written $(w', x') \succ_r (w, x)$, if $w' = w$, $x' \neq x$, $f_i(w_i) \geq x'_i \geq x_i$ for all $i \in R$ and $x_i \geq x'_i \geq f_i(w_i)$ for all $i \notin R$. In particular, for any allocation (w, x) , if $x \neq f(w)$ then $(w, f(w)) \succ_r (w, x)$.

2.6 Domination: If $(w', x') \succ_p (w, x)$ or $(w', x') \succ_r (w, x)$, we say that (w', x') *dominates* (w, x) , written $(w', x') \succ (w, x)$.

In the domination relation $(w', x') \succ (w, x)$, the allocation (w, x) is interpreted as the status quo allocation, which determines the power of each coalition and the ability of any coalition to change the status quo to their own benefit. The statement " $(w', x') \succ (w, x)$ " is interpreted to mean that at the status quo allocation (w, x) , there is a coalition that would benefit from changing the allocation to (w', x') and has the ability to make this change.

The domination relation \succ yields the traditional coalitional game solution concepts of the *core* and *stable set* (Lucas (1992)).

2.7 Core: The core of an economic pillage game is the set $K = \{(w, x) \in A : \text{there is no } (w', x') \in A \text{ such that } (w', x') \succ (w, x)\}$.

The core is the set of allocations that no coalition has the ability to change to its benefit. As in Polemarchakis and Rowat (2007), core allocations are necessarily efficient. An allocation (w, x) is undominated by voluntary transaction only if it is efficient, and undominated by rent retention only if $x = f(w)$. An efficient allocation $(w, f(w))$ is undominated by pillage only if each player i with positive consumption has enough power to prevent the coalition consisting of all other players from taking it. These observations are summarized by the following.

2.8 Proposition: The core of an economic pillage game is

$$K = \{(w, x) \in P : x = f(w) \text{ and } \pi(\{i\}, w) \geq \pi(I \setminus \{i\}, w) \text{ for each } i \text{ with } w_i > 0\}.$$

For example, suppose the power of each coalition is given by the coalition's total output, that is,

$$\pi(c, w) = \sum_i f_i(w_i),$$

an example that will be considered more fully in Section 3. Moreover, suppose that the production technology is the same for all players, that is, there is a strictly increasing

production function $f : [0, 1] \rightarrow \mathbb{R}_+$ satisfying $f(0) = 0$ and $f_i = f$ for all i . If f is strictly convex, then the core consists of the n allocations $K = \{(w, x) : w_i = 1 \text{ and } x_i = f(1) \text{ for some } i\}$. In contrast, if f is strictly concave, then efficiency requires that $w = (\frac{1}{n}, \dots, \frac{1}{n})$. If $n \geq 3$, this wealth allocation leaves every player vulnerable to pillage by the coalition of all others, so the core is empty.

This example indicates that the core is always small and frequently empty. A *stable set* (von Neumann-Morgenstern solution) can be somewhat larger.

2.9 Definitions: For any set of allocations E , let $U(E)$ denote the set of allocations undominated by E , that is, $U(E) = \{(w, x) \in A : \text{no } (w', x') \in E \text{ dominates } w\}$. For example, $K = U(A)$. A set of allocations E is *internally stable* if no allocation in E is dominated by an allocation in E , that is,

$$\text{IS) } E \subset U(E).$$

A set of allocations E is *externally stable* if every allocation not in E is dominated by some allocation in E , that is,

$$\text{ES) } U(E) \subset E.$$

A set of allocations E is *stable* if it is both internally and externally stable, that is,

$$\text{S) } E = U(E).$$

The core always exists, but can be empty. Since $U(\emptyset) = A$, external stability (ES) implies that stable sets cannot be empty, but they can fail to exist. External stability also implies that a stable set must contain the core. The set of efficient allocations is externally stable, since any inefficient allocation is dominated by an efficient allocation by voluntary transaction. However, the set of all efficient allocations is not internally stable, since any efficient allocation that imposes rent payments is dominated by the allocation without rent payments by rent retention.

Von Neumann and Morgenstern (1947) interpret a stable set as a self-enforcing “standard of behavior.” A stable allocation (w, x) may be dominated by an allocation (w', x') , but external stability ensures that (w', x') is in turn dominated by a stable allocation (w'', x'') , and internal stability implies that (w'', x'') does not dominate (w, x) . Thus players are dissuaded from moving to (w', x') by the prospect of a subsequent move to (w'', x'') . However, Harsanyi (1974) observes that the players who are made better off by the move to (w', x') might be still further improved by the subsequent move to (w'', x'') , so the expectations implicit in the definition of a stable set can actually be destabilizing. To address the Harsanyi critique of the stable set, it is necessary to make players’ expectations explicit.

2.10 Definitions: An *expectation* is a function $\eta : A \rightarrow A$ with the property that for each $(w, x) \in A$, $\eta(\eta(w, x)) = \eta(w, x)$. Given an expectation η , let S_η denote the set of *stationary* allocations, that is, $S_\eta = \{(w, x) : \eta(w, x) = (w, x)\}$. Note that $S_\eta = \eta(A)$.

An expectation η represents the common belief of all players that if an allocation (w', x') is reached, it will then be replaced by the allocation $\eta(w', x')$. The requirement

that $\eta(\eta(w, x)) = \eta(w, x)$ ensure that players do not have to anticipate subsequent moves. Given this expectation, all players will evaluate a move from (w, x) to (w', x') not in terms of the consumption allocation x' , but in terms of x'' instead, where $(w'', x'') = \eta(w', x')$.

2.11 Definitions: Let η be an expectation. An allocation (w', x') *dominates* an allocation (w, x) *in expectation* if either $\eta(w', x') \succ_p (w, x)$ or $(w', x') \succ_r (w, x)$ and $\{i : x'_i > x_i\} \subset \{i : x''_i > x_i\}$, where $(w'', x'') = \eta(w', x')$. Define the *farsighted core* as the set of allocations that are undominated in expectation, that is, $K_\eta = \{(w, x) : \text{no } (w', x') \in A \text{ dominates } (w, x) \text{ in expectation}\}$.

The only restriction on an expectation η imposed by Definition 2.10 is $\eta^2 = \eta$. For example, let η be the identity function on A . Then $S_\eta = A$ and $K_\eta = K$. However, the expectation η is irrational since any allocation $(w, x) \notin K_\eta$ invites a change to an allocation that dominates it in expectation. Thus allocations that are dominated in expectation should not be expected to be stationary. Similarly, an allocation (w, x) should not be expected to be succeeded by an allocation $\eta(w, x)$ that does not dominate (w, x) in expectation. An expectation that satisfies these rational expectations conditions is said to be *consistent*.

2.12 Definition: An expectation η is *consistent* if for each $(w, x) \in A$, either

- i) $\eta(w, x)$ dominates (w, x) in expectation; or
- ii) (w, x) is undominated in expectation and $\eta(w, x) = (w, x)$.

If η is a consistent expectation, then $S_\eta = K_\eta$. The farsighted core of a consistent expectation is the set of allocations that satisfy the Harsanyi critique. Jordan (2006, Theorem 6.4) shows that in the basic pillage game, every stable set is a far sighted core for a consistent expectation, and every farsighted core for a consistent expectation is also a stable set. Thus, in the basic pillage game, stable sets are the only sets of allocations that satisfy the Harsanyi critique. The following result shows that in economic pillage games, it is also true that only stable sets can be farsighted cores for consistent expectations. However, an example is given in Section 3 showing that not all stable sets in economic pillage games can be farsighted cores for consistent expectations.

2.13 Theorem: If f is a consistent expectation, then $S_\eta = K_\eta$. A set of allocations, S , is stable if there exists a consistent expectation η such that $S = S_\eta = K_\eta$.

Proof: The first assertion is an immediate consequence of Definition ???. To prove the second assertion, suppose that η is a consistent expectation satisfying $S = K_\eta = S_\eta$. Let $(w, x) \in S$. Since $(w, x) \in K_\eta$, (w, x) is undominated in expectation, so (w, x) is undominated by any $(w', x') \in S_\eta$, and hence (w, x) is undominated by any $(w', x) \in S$. Therefore S is internally stable. Let $(w, x) \notin S$. Then $(w, x) \notin K_\eta$, so (w, x) is dominated in expectation by some (w', x') . Thus (w, x) is dominated by $\eta(w', x') \in S_\eta = S$, so S is externally stable.

3. Examples

This section studies two classes of economic pillage games. In both the production function is the same strictly concave function for all players.

3.1 Assumption: Let $f : [0, 1] \rightarrow \mathbb{R}_+$ be a strictly increasing, continuous and strictly concave function satisfying $f(0) = 0$. Assume that $f_i = f$ for all i . Let $w^* = (\frac{1}{n}, \dots, \frac{1}{n})$ and let $y^* = nf(\frac{1}{n})$. Then $P = \{(w, x) : w = w^* \text{ and } \sum_i x_i = y^*\}$.

The traditional majority game (e.g., Lucas (1992)) is a coalitional game in which any coalition of more than half the players can dictate the allocation of a fixed amount of consumption. This can be modified to make an economic pillage game by specifying the power of each coalition as depending on the size and total output of the coalition, with size weighted much more heavily. Then wealth, through total coalitional output, is relevant only in comparing the power of two coalitions of the same size.

3.2 Majority pillage: Assume that n is odd and $n \geq 3$. Define the *majority power function* $\pi_m(C, w) = v\#C + \sum_{i \in C} f(w_i)$ where $v > f(1)$. The majority power function yields the pillage relation

$$(w', x') \succ_p (w, x) \quad \text{if} \quad \begin{cases} \#W > \#L; \text{ or} \\ \#W = \#L \quad \text{and} \quad \sum_W f(w_i) > \sum_L f(w_i), \end{cases}$$

where $W = \{i : x'_i > x_i\}$ and $L = \{i : x'_i < x_i\} \cup \{i : w'_i \neq w_i \text{ and } x'_i = x_i\}$.

The traditional majority game has a stable set consisting of all allocations in which total wealth is divided evenly among a minimal majority coalition. The analogue for an economic pillage game, in which minimal majority coalitions divide the efficient total output, is shown to be stable by Proposition 3.4.

3.3 Definition: Let $k = (n + 1)/2$, and define

$$S_m = \{(w, x) \in P : w = w^* \text{ and } x \text{ is any permutaion of } (\frac{y^*}{k}, \dots, \frac{y^*}{k}, 0, \dots, 0)\}.$$

3.4 Proposition: S_m is a stable set of the majority game.

Proof: To verify internal stability, let (w^*, x) and (w^*, x') be distinct allocations in S_m . By the definition of S_m , $(w^*, x) \not\prec_r (w^*, x')$. Let $W = \{i : x'_i > x_i\}$ and let $L = \{i : x'_i < x_i\}$. By the definition of S_m , $\#W = \#L$ and $\sum_{i \in W} f(w_i^*) = \sum_{i \in L} f(w_i^*)$, which implies that $(w^*, x) \not\prec_p (w^*, x)$ and thus that S_m is internally stable. To verify external stability, let $(w, x) \notin S_m$. Then $\#\{i : x_i < \frac{y^*}{k}\} \geq k$. Let $(w^*, x') \in S_m$ satisfy $\{i : x'_i = \frac{y^*}{k}\} \subset \{i : x_i < \frac{y^*}{k}\}$, so that $\#\{i : x'_i > x_i\} = k > \frac{n-1}{2} \geq \#\{i : x'_i < x_i\}$. Hence $(w^*, x') \succ_p (w, x)$, so S_m is externally stable.

The stability of the set S_m is unaffected by the possibility of rent retention, which is not present in either the traditional majority game or the majority pillage game studied in Jordan (2006) and Jordan and Obadia (??). However, rent retention prevents S_m from being the farsighted core of a consistent expectation. For example, let $n = 3$. Then $S_m = \{(w, x) : w = w^* \text{ and } x_i = 0 \text{ for some } i = 1, 2, 3\}$. Each of the three stable allocations is dominated by the no-rent allocation $(w^*, x^o) = (w^*, (\frac{y^*}{3}, \frac{y^*}{3}, \frac{y^*}{3}))$. Let η be a consistent expectation such that $\eta : A \rightarrow S_m$. Suppose that $\eta(w^*, x^o) = (w^*, (\frac{y^*}{2}, \frac{y^*}{2}, 0))$. Then at the stable allocation $(w^*, (0, \frac{y^*}{2}, \frac{y^*}{2}))$, which requires player 1 to pay his entire output as rent to the other two players, player 1 can retain his output to change the allocation to (w^*, x^o) , which under the expectation η results in the stable allocation $(w^*, (\frac{y^*}{2}, \frac{y^*}{2}, 0))$ which benefits player 1. Any other choice of $\eta(w^*, x^o)$ creates the incentive for the rent-paying player to retain the rent at one of the three stable allocations. Thus the Harsanyi critique is fatal to the stable set S_m . The following Proposition formalizes and extends this reasoning to $n > 3$.

3.5 Proposition: There is no consistent expectation for which S_m is a farsighted core of the majority game.

Proof: Suppose by way of contradiction that $\eta : A \rightarrow S_m$ is a consistent expectation with $K_\eta = S_m$. Let $x' = (\frac{y^*}{n}, \dots, \frac{y^*}{n})$, so that $(w^*, x') \notin S_m$. Then $\eta(w^*, x') \in S_m$. Let $(w^*, x'') = \eta(w^*, x')$. Permuting the player indices if necessary, we can assume that $x'' = (\frac{y^*}{k}, \dots, \frac{y^*}{k}, 0, \dots, 0)$. Let $x = (0, \dots, 0, \frac{y^*}{k}, \dots, \frac{y^*}{k})$. Then $(w^*, x) \in S_m$. Moreover, $(w^*, x') \succ_r (w^*, x)$ (players $1, \dots, k-1$ retain their rent) and $\{i : x'_i > x_i\} = \{1, \dots, k-1\} = \{i : x''_i > x_i\}$, so (w^*, x') dominates (w, x) in expectation by Definition ???. Since $(w^*, x) \in S_m$, this proves that $K_\eta \neq S_m$.

The power function of the majority game, π_m , place much grater weight on coalitional size than on coalitional wealth or output. At the opposite extreme is the power function that depends only on coalitional output.

3.6 Wealth is power: Define the *wealth-is-power* (WIP) power function $\pi_w(C, w) = \sum_{i \in C} f(w_i)$. The WIP power function yields the pillage relation

$$(w', x') \succ_p (w, x) \quad \text{if} \quad \sum_{i \in W} f(w_i) > \sum_{i \in L} f(w_i),$$

where $W = \{i : x'_i > x_i\}$ and $L = \{i : x'_i < x_i\} \cup \{i : w'_i \neq w_i \text{ and } x'_i = x_i\}$.

The following result shows that the three player WIP game has no efficient stable set.

3.7 Proposition: Let $n = 3$ and let $S \subset P$. If $y^* - f(1) < f(\frac{1}{2})$, then S is not a stable set of the WIP game.

Proof: Suppose that S is externally stable. Let $w^1 = (1, 0, 0)$ and $x^1 = (f(1), 0, 0)$. Then $(w^1, x^1) \notin P$, so there is some $(w^*, x) \in S$ such that $(w^*, x) \succ (w^1, x^1)$. Since no rent

is paid at the allocation (w^1, x^1) , it must be that $(w^*, x) \succ_p (w^1, x^1)$. By the definition of π_w , this implies that $x_1 > f(1)$. Let $w^2 = (0, \frac{1}{2}, \frac{1}{2})$ and $x^2 = (0, f(\frac{1}{2}), f(\frac{1}{2}))$. Since $(w^2, x^2) \notin P$, there must be some $(w^*, x') \in S$ such that $(w^*, x') \succ (w^2, x^2)$. Here again, it must be that $(w^*, x') \succ_p (w^2, x^2)$. This implies that $x'_2, x'_3 \geq f(\frac{1}{2})$. If $y^* - f(1) < f(\frac{1}{2})$, then since $x_1 > f(1)$ and $\sum_i x_i = y^*$, it follows that $x_2, x_3 < f(\frac{1}{2})$. This implies that $(w^*, x') \succ_p (w^*, x)$, so S cannot be internally stable.

4. Legitimacy

An expectation $\eta : A \rightarrow A$, as defined in 2.10, is limited in its ability to dissuade players from pillage. This is why the stable set S_m in the majority game fails to be a farsighted core. At an allocation $(w^*, x) \in S_m$, the players with zero consumption can retain their rents to change the allocation to $(w^*, x^*) = (w^*, (\frac{y^*}{n}, \dots, \frac{y^*}{n}))$. This can only be discouraged if the expectation $\eta(w^*, x^*) = (w^*, x)$, restoring the previous stable allocation, but the allocation (w^*, x^*) does not provide the necessary historical information. This can be remedied by extending the game as in Jordan (2006b) to add variables that can record historical information, along with a procedure for updating the historical information in response to players' actions.

4.1 Definitions: A *dynamic extension* of an economic pillage game consists of a set H and a recording function $\delta : A \times H \times \mathcal{C}^2 \times A \rightarrow H$. An *expectation* is a function $\eta : A \times H \rightarrow A \times H$ satisfying $\eta^2 = \eta$. Since $\eta^2 = \eta$, the set $\eta(A \times H)$ is the set of *stationary states*. Given an expectation η and states $((w, x), h)$ and $((w', x'), h')$, let $((w'', x''), h'') = \eta((w', x'), h')$. The state $((w', x'), h')$ *dominates* $((w, x), h)$ *in expectation* if there exist coalitions F (the *forcing coalition*) and R (the *resisting coalition*) satisfying

- i) $\delta((w, x), h, F, R, (w', x')) = h'$;
- ii) $F \cap R = \emptyset$; and either
- iii.a) $\pi(F, w) > \pi(R, w)$, $F \subset \{i : x''_i > x_i\}$ and $\{i : x''_i < x_i\} \cup \{i : x''_i = x_i \text{ and either } w'_i \neq w_i \text{ or } w''_i \neq w_i\} \subset R$; or
- iii.b) $(w', x') \succ_r (w, x)$, and $F = \{i : x'_i > x_i\} \subset \{i : x''_i > x_i\}$.

Note that (iii.a) implies $(w'', x'') \succ_p (w, x)$. The forcing coalition has the ability to force the move from (w, x) to (w', x') and, anticipating the subsequent move to (w'', x'') , also has the incentive to do so. The definitions of consistent expectations and farsighted cores in Section 2 extend directly. A set of allocations that can be stabilized as the farsighted core of a consistent expectation is said to be *legitimate*.

4.2 Definitions: An expectation η is *consistent* if for each $((w, x), h)$, either

- i) $\eta((w, x), h)$ dominates $((w, x), h)$ in expectation; or
- ii) $((w, x), h)$ is undominated in expectation and $\eta((w, x), h) = ((w, x), \delta((w, x), h, \emptyset, \emptyset, (w, x)))$.

Given a consistent expectation η , the *farsighted core* is the set of states $K_\eta \subset A \times H$ that are undominated in expectation. If K_η is the farsighted core for a consistent expectation η , the set $\{(w, x) : ((w, x), h) \in K_\eta \text{ for some } h \in H\}$ is said to be *legitimate*. Note that since $\eta^2 = \eta$, (i-ii) imply that $\eta(A \times H) \subset K_\eta$.

As in Jordan (2006b), the inclusion of historical information can make the class of legitimate sets much larger than the class of farsighted cores without historical information, which must in particular be stable sets. Legitimate sets need not be stable. Unfortunately, there remain some pillage games with no efficient legitimate sets. In particular, the three-player WIP game considered in Proposition 3.7 fails to have an efficient legitimate set for essentially the same reasons that it fails to have an efficient stable set.

4.3 Proposition: Let $n = 3$ and let $S \subset P$. If $y^* - f(1) < f(\frac{1}{2})$, then S is not a legitimate set of the WIP game.

Proof: Let (H, δ) be a dynamic extension and let η be a consistent expectation. Suppose by way of contradiction, $S = \{(w, x) : ((w, x), h) \in K_\eta \text{ for some } h\}$. Let $w^1 = (1, 0, 0)$ and $x^1 = (f(1), 0, 0)$, and let $h^1 \in H$. Then $(w^1, x^1) \notin P$, so there is some $(w^*, x) \in S$ and $h \in H$ such that $((w^*, x), h) = \eta((w^1, x^1), h^1)$. Since η is consistent and no rents are paid at (w^1, x^1) , 4.1(iii.a) implies that $(w^*, x) \succ_p (w^1, x^1)$, so $x_1 > f(1)$. Let $w' = (0, \frac{1}{2}, \frac{1}{2})$ and $x' = (0, f(\frac{1}{2}), f(\frac{1}{2}))$. Since $(w', x') \notin P$, there must be some $(w^*, x'') \in S$ and $h'' \in H$ such that $(w^*, x'') = \eta(w', x')$. Here again, it must be that $(w^*, x'') \succ_p (w', x')$. This implies that $x''_2, x''_3 \geq f(\frac{1}{2})$. If $y^* - f(1) < f(\frac{1}{2})$, then since $x_1 > f(1)$ and $\sum_i x_i = y^*$, it follows that $x_2, x_3 < f(\frac{1}{2})$. Let $F = \{2, 3\}$ and $R = \{1\}$, and let $h' = \eta((w^*, x), h, F, R, (w', x'))$. Then $\pi_w(F, w^*) > \pi_w(R, w^*)$, so by 4.2 (iii.a), $((w', x'), h')$ dominates $((w^*, x), h)$ in expectation, so $((w^*, x), h) \neq \eta((w^1, x^1), h^1)$, and this contradiction completes the proof.

At an efficient allocation in the three-player WIP game, two players can pillage the third and split the total wealth equally between them, making each of them individually invulnerable to punishment. This is not possible in the three-player majority game, in which no single player can become as powerful as the coalition of the other two. The following definition generalizes this property of π_m .

4.4 Definition: For each i , let e^i denote the wealth allocation that gives everything to player i , that is, $e^i_i = 1$. A power function π is said to be *nontyrannical* if
 NT) $\pi(\{i\}, e^i) < \pi(I \setminus \{i\}, e^i)$ for all i .

The majority power function π_m is nontyrannical under the assumption that $n \geq 3$, but the WIP power function π_w is not, since $\pi_w(\{i\}, e^i) = f(1) > 0 = \pi_w(I \setminus \{i\}, e^i)$.

The *dynamic citizenship extension* introduced in Jordan (2006b) can be adapted to economic pillage games to make the entire set of efficient allocations legitimate if π is nontyrannical. The basic idea of the citizenship extension is the following. Suppose that each player can be designated as either a *citizen* or an *outlaw*. A social state then consists of an allocation together with a given designation for each player. The player designations change with pillage in the following way. If any citizen is pillaged, all players who benefit

from the pillage become outlaws. However, if only outlaws are pillaged, then all players become citizens. This social norm protects the property of citizens, provided that the no-tyranny condition is satisfied. Any player who participates in the pillage of a citizen becomes an outlaw, and is thus left open to pillage by anyone. Of course, even under the no tyranny condition, it may not be possible for the remaining citizens to pillage all of the outlaws. For example, in the three-player majority pillage game, suppose that player 1 is a citizen and is pillaged by the coalition of the other two players. Then player 1 will need to enlist the aid of one of the two outlaws, say player 2, to get some of his property back. In this case, player 2 benefits a second time by betraying his fellow outlaw, but this prospect discourages player 3 from participating in the original pillage.

4.5 Definition: The *dynamic citizenship extension* is defined as follows. Let $H = C$ and let $((w, x), C, F, R, (w', x')) \in A \times H \times C^2 \times A$. Otherwise, let

$$\delta((w, x), C, F, R, (w', x')) = \begin{cases} I \setminus F & \text{if } R \cap C \neq \emptyset; \\ I & \text{if } R \neq \emptyset \text{ and } R \cap C = \emptyset; \\ C & \text{if } R = \emptyset \text{ and } \{i : x_i > 0\} \subset C; \\ I & \text{if } R = \emptyset \text{ and } \{i : x_i > 0\} \not\subset C; \end{cases}$$

Define $S = \{((w, x), C) : (w, x) \in P \text{ and } \{i : x_i > 0\} \subset C\}$.

4.6 Theorem: Suppose that π is nontyrannical. In the dynamic citizenship extension, an expectation $\eta : A \times H \rightarrow A \times H$ can be defined as follows. Let $((w, x), C) \in A \times H$. We will define $((w', x'), C') = \eta((w, x), C)$ by proceeding through the cases below. For each case in which $((w', x'), C') \neq ((w, x), C)$, we will also define coalitions F and R .

- i) $((w, x), C) \in S$. In this case, let $((w', x'), C') = ((w, x), C)$.
- ii) $((w, x), C) \notin S$ and $(w, x) \in P$. In this case, there is some $i^o \in I \setminus C$ such that $x_{i^o} > 0$. Let x' satisfy $\sum x'_i = \sum f_i(w_i)$, $x'_{i^o} = 0$ and $x'_i > x_i$ for all $i \neq i^o$. Let $w' = w$ and $C' = I$. Then $(w', x') \in P$, and $((w', x'), C') \in S$. Since P is dynamically protected, Remark 4.5 implies that $(w, x) \in V$, so $(w', x') \succ_p (w, x)$. Let $F = I \setminus \{i^o\}$ and $R = \{i^o\}$.
- iii) $((w, x), C) \notin S$, $(w, x) \notin P$ and $C = I$. In this case, let $C' = I$, let $(w', x') \in P$ such that $(w', x') \succ_t (w, x)$, let $F = \{i : x'_i > x_i\}$ and let $R = \emptyset$.
- iv) $((w, x), C) \notin S$, $(w, x) \notin P$ and there is some $i^o \notin C$. In this case, let $C' = I$ and $(w', x') \in P$ such that $(w', x') \succ_p (w, x)$ and $x_{i^o} = 0$. Let $F = I \setminus \{i^o\}$. and let $R = \{i^o\}$.

Then η is consistent and S is the farsighted core. In particular, the set P is legitimate.

Proof: Since π is nontyrannical, η is well-defined. It is clear that $\eta(A \times H) = S$. The construction of η ensures that η is consistent, so $K_\eta = S$.

References

- Harsanyi, J. (1974), "An equilibrium-point interpretation of stable sets and an alternative definition," *Management science* **20**, 1472-1495.
- Jordan, J. (2006a), Pillage and property, *Journal of economic theory*, **131**, 26-44.
- Jordan, J. (2006b), "Power and legitimacy in pillage games," mimeo.
- Jordan, J. and D. Obadia (2005), "Stable sets in majority pillage games," mimeo.
- Lucas, W. (1992), "Von Neumann-Morgenstern stable sets," in *Handbook of game theory*, Aumann, R. and S. Hart (eds), Elsevier Science Publishers B.V., 543-590.
- Polemarchakis, H. and C. Rowat (2007) "An exchange economy with pillage," draft.
- Von Neumann, J. and O. Morgenstern (1947), *Theory of games and economic behavior*, John Wiley & Sons, Inc., New York.