Cooperative Game Theory

Cooperative games are often defined in terms of a characteristic function, which specifies the outcomes that each coalition can achieve for itself.

For some games, outcomes are specified in terms of the total amount of dollars or utility that a coalition can divide. These are games with transferable utility.

For other games, utility is non-transferable, so we cannot characterize what a coalition can achieve with a single number. For example, if a coalition of consumers can reallocate their endowments between themselves, then the "utility possibilities frontier" in general is not linear.

Aumann (and others) think of cooperative games as differing from non-cooperative games only in that binding agreements are possible before the start of the game. The primative notions defining a cooperative game are the set of players, the action sets, and the payoffs. The value to a coalition is what it can achieve by coordinating their actions.

For the market game where players decide how much of their endowments to trade, the players outside a coalition cannot affect the trading within a coalition, so the action sets and payoffs uniquely determine the characteristic function of the cooperative game.

For a Cournot game, where the players are the firms who choose quantities, what a coalition can achieve depends on what is assumed about the behavior of the players outside the coalition.

Definition 257.1: A coalitional game (game in characteristic function form) with transferable payoff consists of

–a finite set of players, N

–a function v that associates with every nonempty subset S of N a real number $v(S)$, the "worth" of the coalition S.

The interpretation of $v(S)$ is the amount of money or utility that the coalition can divide between its members.

Definition 287.1: A coalitional game $< N, v >$ with transferable payoff is cohesive if we have

$$v(N) \geq \sum_{k=1}^{K} v(S_k)$$

for every partition $\{S_1, ..., S_K\}$ of N.

Note: cohesiveness is a special case of the stronger assumption, superadditivity, which requires

$$v(S \cup T) \geq v(S) + v(T)$$

for all disjoint coalitions S and T.

Cohesiveness guarantees that the equilibrium coalition is the one that should form. The worth of other coalitions will influence how $v(N)$ will be divided among the players.

Definition (feasibility): Given a coalitional game $< N, v >$ with transferable payoff, for any profile of real numbers $(x_i)_{i \in N}$ and any coalition S, let $x(S) \equiv \sum_{i \in S} x_i$. Then $(x_i)_{i \in S}$ is an S-feasible payoff vector if $x(S) = v(S)$. An N-feasible payoff vector is a feasible payoff profile.

Definition 258.2: The **core** of a coalitional game with transferable payoff $< N, v >$ is the set of feasible payoff profiles $(x_i)_{i \in N}$ for which there is no coalition S and S-feasible payoff vector $(y_i)_{i \in S}$ such that $y_i > x_i$ holds for all $i \in S$.

For games with transferable payoff, the core is the set of feasible payoff profiles such that $v(s) \leq x(S)$ holds for all coalitions, $S \subseteq N$. Thus the core is the set of points in payoff space that satisfies a finite number of linear inequalities. The core is a closed, convex set.

Example 259.1 (A three player majority game)

Three players together can obtain \$1 to share, any two players can obtain α, and one player by herself can obtain zero.

Then $v(N) = 1, v(\{1, 2\}) = v(\{1, 3\}) = v(\{2, 3\}) = \alpha, v(\{1\}) = v(\{2\}) = v(\{3\}) = v(\varnothing) = 0$.

The core of this game is

$$\begin{aligned}\{(x_1, x_2, x_3) &\in \mathbb{R}^3_+ : x_1 + x_2 \geq \alpha, \\ x_1 + x_3 &\geq \alpha, \text{ and} \\ x_2 + x_3 &\geq \alpha\}\end{aligned}$$

If $\alpha > \frac{2}{3}$, the core is empty.

If $\alpha = \frac{2}{3}$, the core is the single point, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

If $\alpha < \frac{2}{3}$, the core has a continuum of points.

Exercise 261.1 (simple games) A coalitional game $< N, v >$ with transferable payoff is a simple game if $v(S)$ is either 0 or 1 for every coalition S, and $v(N) = 1$. A player who belongs to every winning coalition is a veto player.

(a) If there is no veto player, the core is empty. If player i receives a positive payoff, x_i, there must be a winning coalition with a subset of the remaining players. They could form a coalition and increase their payoffs by dividing x_i (and the payoffs of the other players not in the winning coalition) amongst each other. Thus, any payoff profile in which anyone receives a positive payoff can be excluded from the core.

(b) If there is at least one veto player, then the core is exactly the set of nonnegative feasible payoffs that gives zero to all of the non-veto players. Part (a) shows that anything that does not give everything to the veto players cannot be in the core. To show that any way of allocating the payoff to the veto players is in the core, notice that any coalition S that improves upon the feasible payoff x must include all of the veto players. But then the total amount given to the veto players must exceed 1, which is impossible.

When is the core nonempty?

Given a coalition, S, define the vector $\mathbf{1}_S \in \mathbb{R}^N$ by

$$\begin{aligned}(\mathbf{1}_S)_i &= 1 \quad \text{if} \quad i \in S \\ &= 0 \quad \text{otherwise.}\end{aligned}$$

A collection of numbers between zero and one (weights for each coalition), $(\lambda_S)_{S \in 2^N}$ is a **balanced collection of weights** if we have

$$\sum_{S \in 2^N} \lambda_S \mathbf{1}_S = \mathbf{1}_N.$$

A coalitional game $< N, v >$ with transferable payoff is **balanced** if for every balanced collection of weights, we have

$$\sum_{S \in 2^N} \lambda_S v(S) \leq v(N).$$

Interpretation: λ_S is the fraction of time that each member of coalition S must devote to that coalition. The payoff received by the coalition is then $\lambda_S v(S)$. Requiring a balanced collection of weights is a feasibility condition on the players' time. A game is balanced if there is no allocation of time that yields the players more than $v(N)$.

Proposition 262.1: A coalitional game $< N, v >$ with transferable payoff has a nonempty core if and only if it is balanced.

Example: For the three player majority game, let each player devote half of her/his time to each of her/his 2-player coalitions. The only nonzero weights are

$$\lambda_{\{1,2\}} = \lambda_{\{1,3\}} = \lambda_{\{2,3\}} = \frac{1}{2}.$$

It is easy to see that this collection of weights is balanced. For these weights, we have

$$\sum_{S \in 2^N} \lambda_S v(S) = \frac{3\alpha}{2}.$$

Thus, if $\alpha > \frac{2}{3}$, the game is not balanced, and we have an empty core. If $\alpha \leq \frac{2}{3}$, it is easy to see that the game is balanced, because the weights given above are the most "profitable" weights.

Definition 268.2: A coalitional game with non-transferable payoff consists of

–a finite set of players, N

–a set of consequences, X

–a function V that associates with every nonempty subset S of N a set of possible consequences, $V(S) \subseteq X$

–for each player $i \in N$ a preference relation on X.

Definition 268.3: The core of the coalitional game $\langle N, V, X, (\succsim_i)_{i \in N} \rangle$ is the set of all $x \in V(N)$ for which there is no coalition S and $y \in V(S)$ for which we have $y \succ_i x$ for all $i \in S$.

Example: An exchange economy.

$X = \{(x_i)_{i \in N} : x_i \in \mathbb{R}^{\ell}_{+} \text{ for all } i \in N\}$

$V(S) = \{(x_i)_{i \in N} \in X : \sum_{i \in S} x_i = \sum_{i \in S} \omega_i$ and $x_j = \omega_j$ for all $j \in N \backslash S\}$

$(x_j)_{j \in N} \succsim_i (y_j)_{j \in N}$ if and only if $x_i \succsim_i y_i$

Assume that endowments are strictly positive and that preferences are increasing, continuous, and strictly quasi-concave.

Proposition 272.1: Every competitive allocation is in the core. [Note: Since our assumptions are enough to guarantee that a C.E. exists, we know that the core is nonempty.]

Let E be an exchange economy with n agents, and let kE be the exchange economy derived from E, with a total of kn agents, k copies or replicas of each agent in E.

Lemma 272.2: (equal treatment) Let E be an exchange economy and let k be a positive integer. In any core allocation of kE, all agents of the same type obtain the same bundle.

Proof sketch: If not, collect one agent of each type who is treated the worst in the core allocation. This group can give each agent the average bundle received by that type in the core allocation, making at least one type better off. That type is still better off if we take a tiny amount of consumption from that agent and give some to the other agents in the coalition, contradicting the fact that we were in the core.

Note: Because of equal treatment, the core allocations stay in the same space, $\mathbb{R}^{n\ell}_+$, as we replicate the economy.

Proposition 273.1: Let E be an exchange economy and let x be an allocation in E. If for every positive integer k, the allocation in kE in which every agent of type t receives the bundle x_t is in the core of kE, then x is a competitive equilibrium allocation of E.

The core shrinks as we replicate, so that eventually every allocation except the CE allocations are excluded.

Example: Two types and two goods

$$\begin{aligned} u_i(x_i^1, x_i^2) &= x_i^1 x_i^2 \\ \omega_1 &= (2,1) \text{ and } \omega_2 = (1,2)\ . \end{aligned}$$

By equal treatment, and the fact that core allocations must involve equal marginal rates of substitution, we can characterize the allocation by a scalar α

$$(x_1^1, x_1^2, x_2^1, x_2^2) = (\alpha, \alpha, 3-\alpha, 3-\alpha).$$

For $k = 1$, the minimum value of α consistent with being in the core provides agent 1 with her endowment utility, so

$$\alpha^2 = 2 \; or \; \alpha \simeq 1.4.$$

For general k, the minimum value of α consistent with being in the core of kE is found by considering a coalition with k type 1 agents and $k - 1$ type 2 agents. To prevent the incentive for this coalition to form, the type 1 consumers should not be able to give themselves a utility greater than α^2 and the type 2 consumers a utility of at least $(3 - \alpha)^2$.

Solve the problem

$$\begin{aligned} & \max x_1^1 x_1^2 \\ & \text{subject to} \\ kx_1^1 + (k-1)x_2^1 &= 3k - 1 \\ kx_1^2 + (k-1)x_2^2 &= 3k - 2 \\ x_2^1 x_2^2 &= (3 - \alpha)^2 \end{aligned}$$

Using the resource constraints to solve for type 2 consumption, we have the equivalent problem

$$\max x_1^1 x_1^2$$

$$\text{subject to}$$

$$\left(\frac{3k-1-kx_1^1}{k-1}\right)\left(\frac{3k-2-kx_1^2}{k-1}\right) = (3-\alpha)^2.$$

The first order conditions yield

$$x_1^1 = \left(\frac{3k-1}{3k-2}\right) x_1^2$$

which can be substituted into the constraint to eliminate x_1^1. To prevent this deviation from being advantageous, we must have $x_1^1 x_1^2 \leq \alpha^2$, which holds with equality at the minimum α consistent with the core of kE. This allows us to eliminate x_1^2, giving us the lower boundary of the core, as a function of k.

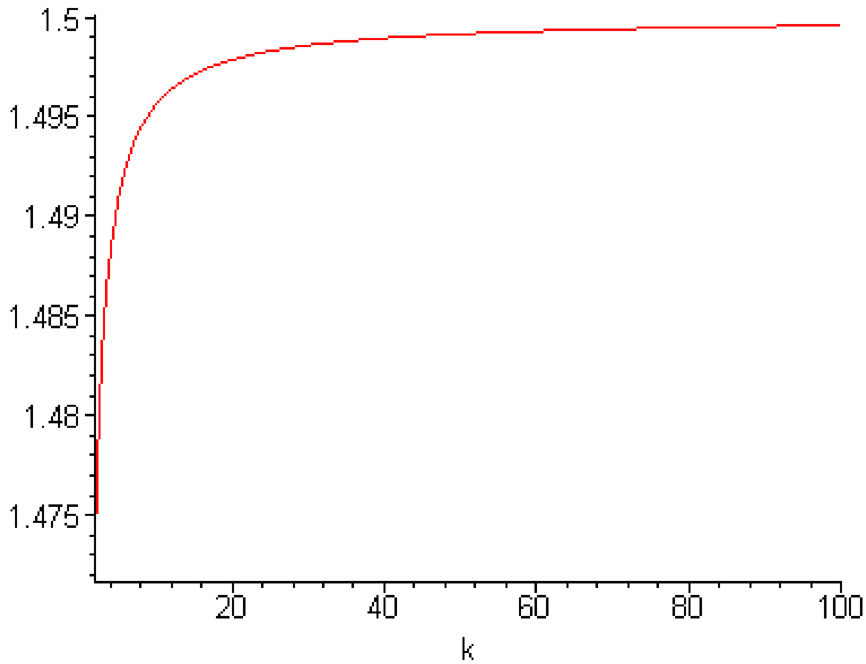

Figure 1:

By symmetry, the upper endpoint of the core is $(3 - \alpha, 3 - \alpha)$. In the limit, the core shrinks to the unique C.E., where everyone consumes $(1.5, 1.5)$.

The Shapley Value

The core suffers from the problems that it is sometimes empty, and that when the core is nonempty it usually admits a continuum of solutions.

The Shapley value always exists and is unique. It can be motivated as the only solution that satisfies the axioms of symmetry, additivity, and a dummy player axiom that says a player that does not add anything to any coalition receives what he can achieve for himself.

Define the *marginal contribution* of player i to a coalition S (not containing i) in the game $\langle N, v\rangle$ to be

$$\triangle_i(S) = v(S \cup \{i\}) - v(S)$$

Definition 291.2: The Shapley value φ is defined by

$$\varphi_i(N, v) = \frac{1}{|N|!} \sum_{R \in \Re} \triangle_i(S_i(R))$$

for each $i \in N$, where $\Re$ is the set of all $|N|!$ orderings of N and $S_i(R)$ is the set of players preceding i in the ordering R.

Suppose that the players will be arranged in random order, with all orderings equally likely. Then the Shapley value is the expected marginal contribution that player i brings to the set of players preceding him.

Notice that, if we fix an ordering, the sum of the marginal contributions must be $v(N)$, so the sum of the expected marginal contributions must also sum to $v(N)$. Thus, the Shapley value is feasible.

α-core and β-core

If a cooperative game is a strategic game in which binding commitments can be reached before the play of the game, then the characteristic function is **not** uniquely determined when the amount a coalition S can achieve depends on the behavior of those in $N \backslash S$.

Under the α-core, we assume that players in $N \backslash S$ choose actions to minimize the payoff to coalition S. In other words, S moves first to maximize their minimum payoff, and $N \backslash S$ responds by minimizing the payoff to S, given the actions of coalition S. Because of the commitment to punish and the ability to react, the α-core forces a coalition to consider the worst-case scenario, which yields a large core (harder to improve upon an allocation).

Under the β-core, the players in $N \backslash S$ choose actions first to minimize the maximum payoff to S, and then the players in S choose a best response. Thus, the value of a coalition is larger than in the α-core (easier to improve upon an allocation), so the β-core is smaller than the α-core.

Maybe both cores are too big, because $N \backslash S$ may threaten an action that is not credible. In a Cournot game, $N \backslash S$ can threaten to produce so much output that the price is zero.