Contents

Chapter 8: Bargaining and Cooperative Game Theory

This chapter introduces several different models of a classical question – how to carve up a "pie" when there are several claimants. It begins with three standard (non-cooperative) game theoretic models, one due to John Nash, and the others to Ariel Rubinstein. Then it looks at three ways to axiomatize the bargaining process. This approach is traditionally known as cooperative game theory, and is of interest in its own right. The solution concepts we introduce — Shapley Value and Core — have many applications beyond bargaining.

1 The Nash Demand Game

- There is a fixed total amount (pie or stake) of a divisible resource, say 100 units, and two or more players. For simplicity, we assume just 2 players.
- Each player makes a confidential demand, i.e., it is a simultaneous move game.
- If the sum of demands is less than 100, then each player gets their demand, otherwise they get nothing.
- Let 1's demand be x and 2's demand be y.
- If $x + y \leq 100$, $\pi_1 = x$, then $\pi_2 = y$; otherwise $0 = \pi_1 = \pi_2$.
- There are 101 NE with integer demands, any $x + y = 100$ is an NE.
  - Symmetric NE $(50, 50)$ seems to be a focal point, even though allocations such as $(87, 13)$ are also NE.

2 Alternating Offers Game

The Rubinstein model (Ariel Rubinstein, 1982).

- Two players take turns making offers until an offer is accepted,
- However, there is a time cost to bargaining.
  - A discount factor $\delta < 1$ shrinks the pie after each time period.
- It turns out that there is a slight advantage related to δ to the player making the first offer.
- Normalize the pie to 1 and then use a variant on IDDS to solve
- We will see that in equilibrium player 1 obtains $\frac{1}{1+\delta}$, and player 2 obtains $\frac{\delta}{1+\delta}$.
  - Thus, if $\delta = 0.9$ then player 1 obtains $\frac{10}{19}$ and player 2 obtains $\frac{9}{19}$,
  - while if $\delta = 0.5$ then player 1 obtains $\frac{2}{3}$ and player 2 obtains $\frac{1}{3}$.
  - So, the payoff advantage to the first player is increasing in the discount rate and decreasing in the discount factor δ.
- The solution can be found by taking the pie and "chopping off the ends."

Figure 1: The first two chops.

- If player 1 offered more than δ then Accept would be a dominant strategy for player 2 in the subsequent subgame, since in that subgame the entire remaining pie is worth at most δ.
  - Hence for player 1, any strategy involving an initial offer exceeding δ is (weakly) dominated by a strategy involving an initial offer of δ or less. A SPNE outcome can't lie in that segment, so we can chop it off, as in Figure 1.

- By the same logic, when it is her turn, player 2 won't be willing to keep less than than $1-\delta$ of the remaining pie, or $\delta(1-\delta)$ of the original pie, so she offers player 1 at most $1-\delta(1-\delta)=1-\delta+\delta^2$.
- This implies that no SPNE outcome will lie in the initial segment $\delta(1-\delta)$, so we can chop that off.
- Given the reduced range of possible outcomes, Player 1 shouldn't offer more than ...
- The upshot is that, if both players have beliefs consistent with SPNE, then the first player will keep $\lim_{n\to\infty} 1-\delta+\delta^2-\delta^3+...+(-\delta)^n=\frac{1}{1+\delta}$, and will offer player 2 the rest, $\frac{\delta}{1+\delta}$, in the very first round, and player 2 will accept.
- See any good book on game theory for a more rigorous proof.

- If the two players have different discount rates then there is an advantage to the more patient player.
- Example: A union and management bargaining over a wage contract (share of profits) during a strike, i.e. there is a time cost.
- If the union is less (or more) patient than management, then it is predicted to gain a smaller (or larger) settlement.
- The model, contrary to frequent experience, predicts immediate settlement. Both players lose by delay, and in SPNE they have nothing to gain from delay, since they correctly anticipate subsequent offers and counteroffers.
- So the model is missing an important ingredient, such as imperfect information or agency problems, or (more behaviorally) players who care (positively or negatively) about how the other player does.

- Even if not complete, the Rubinstein bargaining model seems like an important benchmark and perhaps an important component of a more realistic model.

[[Next draft consider (a) writing out the solution when the discount factors differ, and (b) mentioning the baron-ferejohn model.]]

Ultimatum game.

There is no last period in the full alternating offers game just described. One can also truncate the game after T alternating offers, and get similar equilibria for large finite T.

But how about small T? The most interesting case is $T = 1$. (See exercises for $T = 2, 3, 4$.) This extreme case is known as the Ultimatum Game.

- Pie = 100,
- player 1 offers y to 2.
- If 2 accepts, then $\pi_1 = 100 - y, \ \pi_2 = y$. If 2 rejects, then $0 = \pi_1 = \pi_2$.
- The essentially unique SPNE is to demand 99, leaving 1 for the other player, which they accept since this strictly dominates reject.
- However, this is not what often happens.
  - Many times the player receiving 1, or another sufficiently small amount, will reject and earn zero, to punish the other player for being greedy.
- This departure from SPNE has triggered a vast literature on ultimatum game experiments, second only perhaps to repeated prisoner's dilemma. See Behavioral Considerations section below.

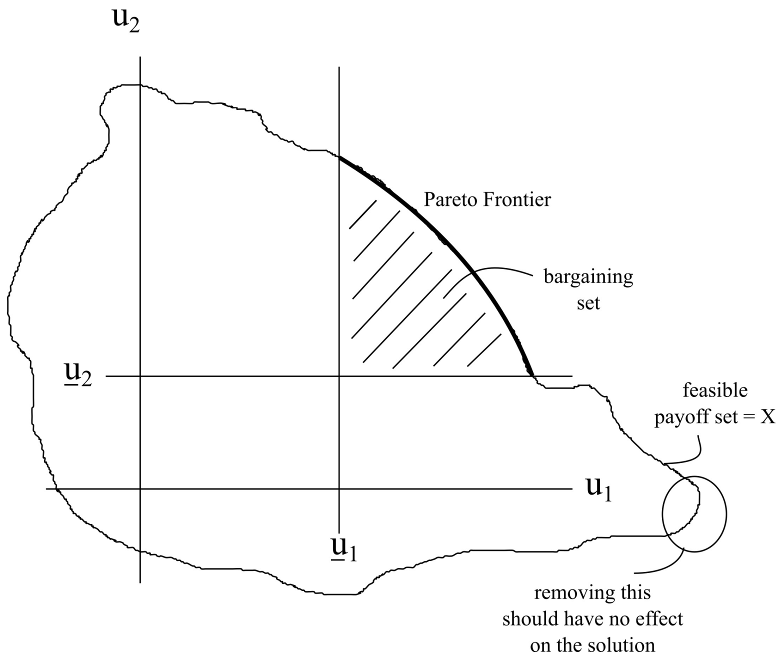

Figure 2: Nash Bargaining Setup. Diagram shows feasible set (not quite convex in the current version) , bargaining set, threat points, and Pareto Frontier.

3 Nash Bargaining Solution (NBS)

The NBS (Nash, 1950, 1953) uses an axiomatic approach. This means that we don't worry about the details of the process, just the results. That is, the payoff vector, however obtained, is assumed to satisfy certain axioms. These axioms concern the feasible payoff set $X \subset R^n$, obtained by mapping all strategy profiles (pure and mixed) to the resulting payoff vectors. In view of mixed strategies, we will assume that X is convex.

For simplicity, for the moment we assume just $n = 2$ players.

- Let $\underline{u}_1$ and $\underline{u}_2$ represent the threat points, security level, reservation payoffs, disagreement payoff or minimax payoff (where these terms are interchangeable).
- This is the maximum payoff that player i can guarantee herself.
- To avoid trivialities, we assume that $(\underline{u}_1, \underline{u}_2) \in X$, so the threat points are mutually feasible.

The Nash Bargaining Solution (u_1^*, u_2^*) is characterized by the following four axioms:

1. It is invariant to linear rescaling.
    - That is, utility is cardinal in the usual sense of von Neumann & Morgenstern.
2. It is Pareto efficient.
    - That is, the solution is on the Pareto Frontier.
3. It is IIA: "Independent of irrelevant alternatives" aka "Invariant to irrelevant alternatives."
    - That is, deleting points not close to (u_1^*, u_2^*) should have no impact on the solution.
    - That is, if $A \succ B$ is true for some feasible choices, then IIA asserts that it will still be true when other alternatives are included or excluded.
    - This sounds innocent, but is where the bodies are buried. See the Behavioral Considerations section for further discussion.
4. Anonymity or symmetry.
    - The solution is preserved if the players are relabeled.

Theorem: (Nash, 1950): Axioms 1-4 imply that (u_1^*, u_2^*) is the solution to:

$$\max_{u_1, u_2 \in X} (u_1 - \underline{u}_1)(u_2 - \underline{u}_2) \tag{1}$$

where X is the set of feasible payoffs and $\underline{u}_1$ and $\underline{u}_2$ are threat points.

It is straightforward to verify that the solution to (1) satisfies each of the four axioms. The other direction — that if (u_1^*, u_2^*) satisfies Axioms 1-4 for every convex X and every $\underline{u} \in X$, then it must solve (1) — takes a bit more work, but no new ideas. See standard textbooks for proofs, e.g., MCWG p. 843-844.

- Axiom 4 can be relaxed so that weights are assigned to the bargainers, and equation (1) is replaced by:

$$\max_{u_1,u_2\in X}(u_1-\underline{u}_1)^{\alpha}(u_2-\underline{u}_2)^{1-\alpha} \tag{2}$$

  Or more generally, with $n \geq 2$ bargainers, by

$$\max_{u_i\in X}\sum_{i=1}^{N}\alpha_i \ln(u_i-\underline{u}_i) \text{ where } \sum \alpha_i = 1. \tag{3}$$

  Example: Consider the two player bargaining game with a "pie" of 100.

  - The threat points are $\underline{u}_1 = 20$ and $\underline{u}_2 = 30$.
  - The problem is:

$$\max_{u_1,u_2\in X}(x-20)(y-30) \text{ such that } x+y=100 \tag{4}$$

  - Since the constraint will hold with equality (we could have $x+y \leq 100$, but we know the constraint will bind) we have:

$$\begin{aligned} \max_x f(x) &= (x-20)(70-x) \\ \max_x f(x) &= -x^2+90x-1400 \\ \frac{\partial f(x)}{\partial x} &= -2x+90=0 \\ x^* &= 45, \quad y^*=55 \end{aligned} \tag{5}$$

  - Note that the SOC is $-2 < 0$ so we have a max.

- One way to interpret the solution is each player gets $\underline{u}_i + \frac{1}{n}(pie-$ sum of the threat points), their threat point plus an equal share of the available surplus (the size of the pie less that already earmarked via threat points).

- With unequal weights a player receives their threat point $\underline{u}_i + \alpha_i(surplus)$.

  - Thus, each player does better if they can increase their threat point.

– The solution is usually unique (except certain knife-edge cases).

– Choosing the threat point may be a pregame game.

- Graphically, the NBS is the highest rectangular hyperbola that meets X.

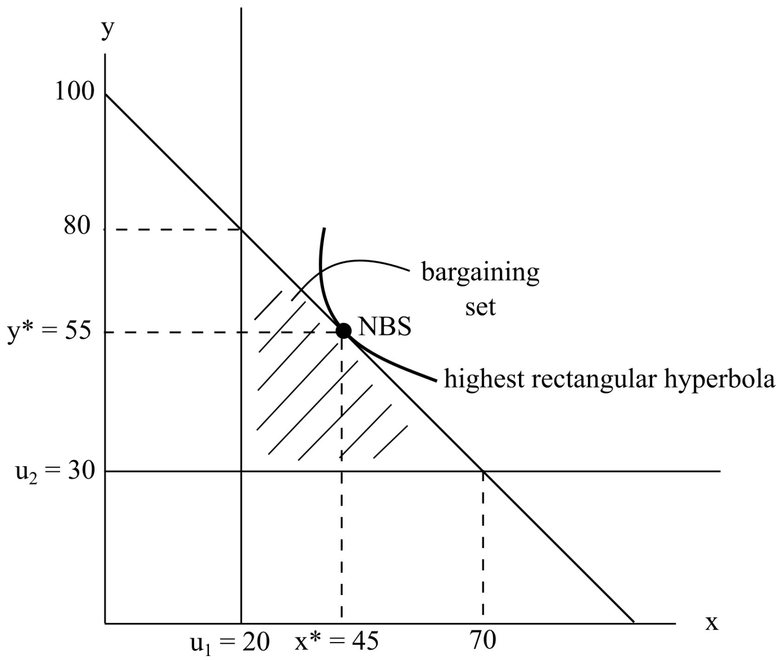

Figure 3: NBS, bargaining set and highest rectangular hyperbola.

4 Characteristic Function Games

Non-cooperative extensive form (EFG) is the most complete representation for a game, showing actions, information sets and timing, but quickly becomes cumbersome.

- NFG streamlines the specification, but even it can be cumbersome with more than 2 players and/or more than 4 strategies each.

- A game specified by a characteristic function (CFG) further streamlines. Ignore the specifics of strategies and simply give the total payoff achievable (the "worth") by any subset of players.

- Change notation, N is now a set of players with $|N| = n$ elements.
  - Any subset of N is called a coalition and denoted by K, where $K \subset N$. The "grand coalition" consists of all players is denoted by N.
  - There are 2^n possible coalitions, including the null coalition $\emptyset$, and the grand coalition N.
- Usually, it is assumed implicitly that there is free communication and binding agreements (i.e., a commitment technology) that are costless to enforce within any coalition.
  - It is possible to include monitoring costs.
- All of this information is reflected in a summary number called the worth v of a coalition.
  - The worth is the total payoff available to the coalition, thus we are considering transferable utility games.
  - That is, payoffs are like money, and units can be transferred 1-1 among players.

A *characteristic function* is a mapping:

$$v : 2^n \to \Re \tag{6}$$

such that:

1. $v(\emptyset) = 0$.
   - The worth of the null coalition is often (not always) normalized to zero.
2. $v(K) \leq v(N) < \infty \quad \forall\ K \in 2^n$.

- Every coalition gets no more than the worth of the grand coalition, which is finite.

3. $v(K) + v(L) \leq v(K \cup L) \quad \forall$ disjoint $K, L \in 2^n$.

   - Disjoint means K and L have no elements in common: $K \cap L = \varnothing$.
   - This property is called **superadditivity,** meaning the whole has worth at least as great as the sum of the parts. It is a monotonicity property.
   - This makes sense, since the whole coalition $K \cup L$ could always separate into its constituent parts with the associated guarantees, but it may be able to do even better together.

Let $X \in \Re^n$ be the set of feasible payoff vectors for N according to the characteristic function v.

- Feasibility with transferrable utility implies

$$x = (x_1, ..., x_n) \in X \implies \sum_{i=1}^{n} x_i \leq v(N) \tag{7}$$

- Non-transferable utility results also exist, but the definitions are messier.
- Most economic applications utilize transferable utility, that is, side-payments are possible.
- The characteristic function looks just at the results, not the process.

5 Core

What can we say about the solution of a CFG? It turns out that there are dozens of candidates, none of which satisfies all desiderata (are you surprised?)

The first solution concept, and arguably still the most important, is called the core. The key idea is called blocking, or eliminating allocations that a coalition could veto. The core is what remains after all vetoed payoff vectors have been eliminated. Think of an apple, and each coalition takes out a bite. What remains is the core.

More formally,

1. A payoff vector $x = (x_1, ..., x_n) \in \Re^n$ is **blocked** by coalition K if:

$$\sum_{i \in K} x_i < v(K) \tag{8}$$

   - Coalition K vetoes any allocation that gives its members less than it can guarantee them.
   - Below we will occasionally write $\beta(K)$ to denote the set of blocked vectors, i.e., feasible allocations satisfying (8).
   - This is a notion of group rationality. But does it make sense to be in an arbitrary coalition?

2. A payoff vector $x = (x_1, ..., x_n) \in \Re^n$ is **individually rational** if

$$x_i \geq v(\{i\}) \quad \forall i \in N \tag{9}$$

   - That is, the payoff vector is not blocked by any of the singleton coalitions.
     - Note this differs for games with externalities.
   - The NBS is individually rational since each player gets at least their threat-point.

3. Pareto optimality. A payoff vector $x = (x_1, ..., x_n) \in \Re^n$ is **Pareto optimal** if:

$$\sum_{i \in N} x_i \geq v(N) \tag{10}$$

  - A payoff vector is Pareto optimal if it is unblocked by the grand coalition.
  - Obviously, this will hold with equality when we restrict our attention to feasible allocations.
  - An **imputation** is an efficient allocation of the worth of the grand coalition and requires only (9) and that (10) holds with equality.

- The core is what remains after eliminating all blocked payoff vectors:

$$\text{imputation } x \in core(v) \text{ iff } x \text{ is not blocked by any } K \in 2^N. \tag{11}$$

  – Thus, if $\sum_{i \in K} x_i \geq v(K) \quad \forall\ K \in 2^N$, then x is in the core.

  – Alternatively, the core is: $X \diagdown \cup_{K \in 2^N} \beta(K)$, that is, the set of feasible payoffs minus the union across all coalitions of blocked payoff vectors.

- Example: Edgeworth box.

  – The core is that portion of the contract curve that is within the lens formed by the endowment point.

  – Why? In this 2 person game, the only coalitions are the grand coalition, which blocks everything not on the contract curve, and the single players, which block anything not preferred to their endowments.

  – Caveat: we don't necessarily have transferrable utility here, but the core is defined the same way even in with NTU.

- **Example.** Consider the following three-player characteristic function game:

$$\begin{aligned} v(\{1\}) &= v(\{2\}) = v(\{3\}) = 0 \\ v(\{1,2\}) &= v(\{1,3\}) = 0.2 \\ v(\{2,3\}) &= 0.3 \\ v(\{1,2,3\}) &= 1 \end{aligned}$$

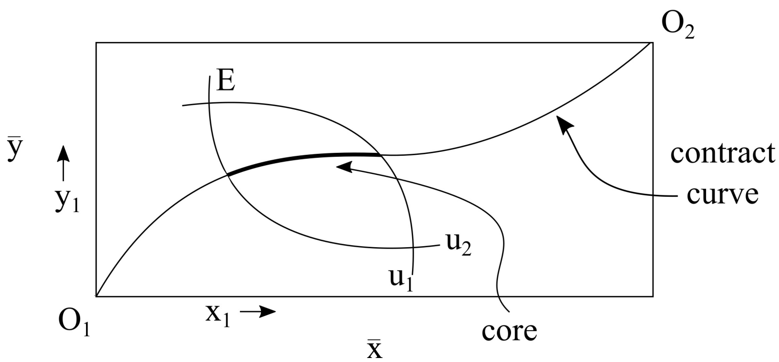

Figure 4: Edgeworth box, the contract curve, reservation utility levels and the core.

- The worth of the grand coalition is often normalized to 1.
- Story: three adjacent colleges want to build a science park. None can do it on their own. Working together, 1 and 2 (or 1 and 3) can build a dinky park, and 2 and 3 can do a little better, but all three combined can create a very nice park.
- In realistic applications, the hard part is usually estimating a reasonably accurate characteristic function.

- Find the core.
  - The core is the set of all payoff vectors that are feasible, individually rational, Pareto optimal and unblocked by proper coalitions.
  - The feasibility and PO conditions say that we are working in a simplex. With 3 players the set of imputations is a triangle.
  - We'll use barycentric coordinates. Graphically, imagine a perpendicular from each edge to the opposite vertex with length 1.

Since $x_1 + x_2 + x_3 = 1$, we can rewrite the blocking condition
$x_1 + x_2 \geq 0.2$ as $x_3 \leq 0.8$, etc.

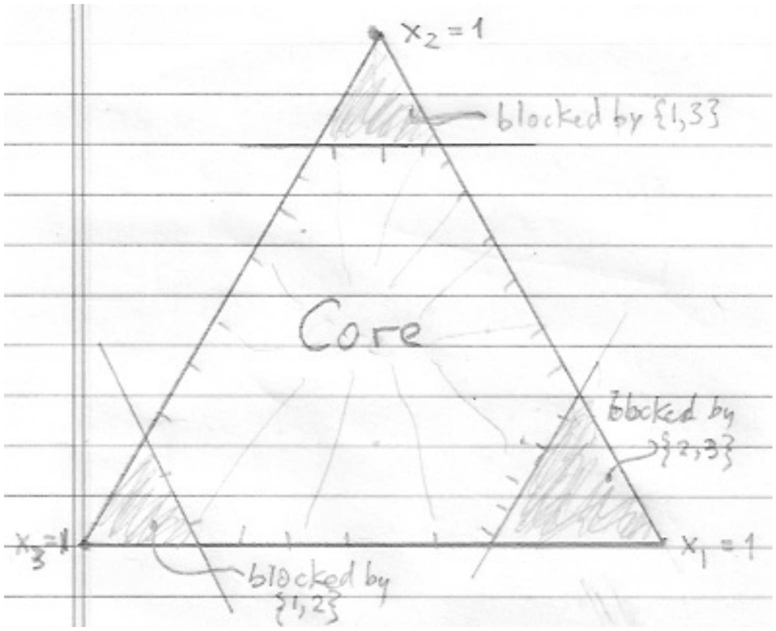

Figure 5: Blocked imputations and the core for given CFG.

Thus the core is the set of all points satisfying:

$$
\begin{aligned}
0 &\leq x_1 \leq 0.7 \\
0 &\leq x_2 \leq 0.8 \\
0 &\leq x_3 \leq 0.8 \\
x_1 + x_2 + x_3 &= 1
\end{aligned}
$$

- So, one point in the core is: $x = (0.2, 0.2, 0.6)$.
- But there are many other points, indeed, in this case it consists of more than half of the simplex.
- In terms of the story, the lesson is that here there are lots of ways to share the facility that nobody (or group) can rationally veto.

Miscellaneous comments about the core:

1. The core may be large, as in the last example.

2. The core can be empty; examples are not hard to construct. Hence Desiderata 2 and 3 from Chapter 4 are unsatisfied.

3. An empty core is interpreted as an unstable situation — there is no bargaining outcome that is stable against defection by some group of players.

4. PO and IR properties of the core ensure that it satisfies Desiderata 1 and 4.

5. Do you think that it is reasonable (Des. 5)?

6. Desideratum 6 is that it predicts well. There haven't been many aggressive tests of the core's predictive power, but in those that we know about (from Fiorina and Plott, 1980 to Yan et al., 2016) it does reasonably well.

7. Reduced forms, such as CHF games may miss subtleties that are more apparent in the extensive form. The core implictly presumes that:

   - any coalition can be formed as easily as any other.
   - Costless enforcement of internal agreement in each coalition.
   - No externalities: the payoff to a player outside a coalition does not change when that coalition is formed.

6 Shapley value (SV)

Like the NBS, the Shapley value is axiomatic. Like the core, it offers a solution to CFGs. Unlike the core, it enforces existence and uniqueness (Desid. 2 and 3). It turns up in applications ranging from Accounting (allocation of joint costs) to Political Science (the Banzhaf power index).

Given a CFG with characteristic function v, suppose that there is a solution vector $\phi(v) = (\phi_1(v), ..., \phi_n(v))$ which satisfies the following Axioms:

1. Pareto optimality.

2. If player i contributes the same worth increment v_i to every possible coalition, then $\phi_i(v) = v_i$.

3. Symmetry. Switching the names of players has no impact on their payoffs.

   - E.g., if v' is v except the names of players 1 and 2 have been switched, then $\phi_3(v) = \phi_3(v'), \phi_1(v) = \phi_2(v')$ and $\phi_2(v) = \phi_1(v')$.

4. Additive over combined games.

   - Suppose $v + \mu = \lambda$, thus λ is a combined game.
   - Then $v(K) + \mu(K) = \lambda(K) \quad \forall K \in 2^n$, that is, the worth of a coalition in the combined game is the sum of its worths in the separate games.

Shapley's Theorem: Any function $\phi(v)$ mapping CHFs to payoff vectors that satisfies Axioms 1. through 4. must be the Shapley value, defined constructively as follows:

$$\phi_i(v) = \sum_{K \in 2^n} [v(K) - v(K \diagdown \{i\})] \left[\frac{(\#K-1)!(\#N-\#K)!}{\#N!} \right] \tag{12}$$

- The first factor in (12) is the marginal contribution of player i to coalition K.
  - It is how much player i adds to the worth of the coalition.
- The second term is a binomial coefficient:
  - $\#N$ refers to the cardinality of the grand coalition $|N| = n$, or the number of players.
  - $\#K$ is the cardinality of the coalition $|K| = k$, i.e. the number of coalition members.
  - The ! denotes the factorial operator.

- The Shapley value is a weighted sum.
- It is the sum of a player's marginal contributions across all possible coalitions, under the assumption that each coalition formation sequence is equally likely.
- The weights are the probabilities that player i joins $K \setminus \{i\}$ as a new member.

- Another way to write the Shapley value is

$$\phi_i(v) = \frac{1}{n!} \sum_{\rho} MC_i(\rho), \tag{13}$$

where ρ is a permutation (reordering) of $\{1, ..., n\}$ and $MC_i(\rho)$ is the marginal contribution $v(K) - v(K \setminus \{i\})$ of player i when he first joins the coalition built up in the order specified by ρ.

- Equation (13) is more compact than (12) but less explicit. It accounts for all possible coalitions by starting with a random player, adding one remaining player at a time to reach the grand coalition, and considering all possible such orderings ρ. We will see that (13) provides a handy way to compute the Shapley value.

Shapley value example

Recall the science park example introduced earlier: $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$; $v(\{1,2\}) = v(\{1,3\}) = 0.2$; $v(\{2,3\}) = 0.3$; $v(\{1,2,3\}) = 1$. There are $2^n = 2^3 = 8$ different coalitions and 6 different coalition formation sequences ρ: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1).

We compute the Shapley value in Table 1. In the line for $\rho = 213$, for example, the sequence starts with player 2. His marginal contribution is 0 because he is just a singleton coalition, which have 0 worth in this example. Next, player 1 joins, raising the worth to $v(\{1,2\}) = 0.2$, so that is his marginal contribution in this sequence. Finally, player 3 joins, raising the worth to $v(\{1,2,3\}) = 1$, so his marginal contribution is $\text{MC}_3 = 1 - .2 = .8$. Doing this calculation for each sequence, we obtain all the possible

Table 1: Shapley value for the science park example. Here there are $n = 3$ players, so the sum of marginal contributions in the next to last line is divided by $n! = 3 \cdot 2 \cdot 1 = 6$ to obtain the Shapley value given in the last line.

ρ	MC_1	MC_2	MC_3
123	0	.2	.8
132	0	.8	.2
213	.2	0	.8
231	.7	0	.3
312	.2	.8	0
321	.7	.3	0
Σ	1.8	2.1	2.1
ϕ_i	0.3	0.35	0.35

marginal contributions for each player, with the proper weightings. Adding them up and dividing by $n! = 6$ we obtain the Shapley value $\phi(v) = (.3, .35, .35)$.

General Properties.

The Shapley value:

1) Always exists.

2) Always is unique.

3) Lies in the core if the game is convex.

A CHF game is called *convex* if its characteristic function is supermodular, that is, if

$$v(K \cup L) + v(K \cap L) \geq v(K) + v(L) \quad \forall K, L \subseteq N \tag{14}$$

This property is similar to superadditivity but is stronger in that $K \cap L$ need not be $\varnothing$. In a convex game, the marginal contribution of a player increases with the size of the coalition to which that player joins.

7 Behavioral Considerations

- See several Binmore articles on alternating offers games and Nash demand games.

A few words on the vast Ultimatum Game literature. The stylized facts: responder (player 2) often engages in negative reciprocity. Proposer (player 1) choices are generally close to BR to responder empirical distribution. Zillions of variants, e.g., involving 3rd parties, ...

Independence of Irrelevant Alternatives seems innocuous for neoclassical individual choice, but certainly is open to question when applied to multi-player choice such as in bargaining. Recall that even majority rule is not even transitive in general.

Even for individual choice, there are behavioral exceptions to IIA. Restaurants often include a ridiculously pricy wine in the list, thinking that it will cause some customers to pick a fairly pricy wine that they would not have picked if the more expensive wine were omitted.

In next version of these Notes, use the example from McGinty, Milam and Gelves experiment (ERE 2012) that has Croson et al 2004 (J of Finance) game to show that superadditivity does not mean that the grand coalition is stable under Shapley, NBS, etc. since the game is concave, even if it is superadditive.

Also can use this to show internal and external stability of the coalitions from d'Aspremont, et. al (1983) *CJE*.

Maybe discuss Fiorina and Plott, 1980 and Yan et al., 2016.

8 Further Reading

Nash, J. (1950). "The Bargaining Problem," *Econometrica*, 18(2): 155-62.

Nash, J. (1953). "Two-Person Cooperative Games," *Econometrica*, 21(1): 128-40.

Shapley, L. (1953). "A Value for n-Person Games," in *Contributions to the Theory of*

Games II (Annals of Mathematics Studies 28), H. W. Kuhn and A. W. Tucker (eds.), Princeton University Press, 307–17.

Winter, E. (2002). "The Shapley Value" in *The Handbook of Game Theory*, eds. R. J. Aumann and S. Hart, North-Holland.

Chander, P. and H. Tulkens (1997). "The Core of an Economy with Multilateral Environmental Externalities," *International Journal of Game Theory*, 26: 379-401.

A. Rubinstein (1982) *Econometrica* finish reference if included.

Henrich et al. (2004) ultimatum game played in 14 different societies.

Yan, Friedman and Munro, 2016