Toulouse Lectures

Bargaining, Coalitions, and Externalities

E. Maskin

June 2004

Dedicated to the memory of Jean-Jacques Laffont
Lecture I

What’s Wrong with Cooperative Game Theory (and the Coase Theorem)?
• Game theory central to economics these days
  – taught as part of core curriculum in virtually every economics department
  – used as a tool in almost every applied field
• But vast majority of game theory used by economists is *noncooperative*
  – focus on connection between individual players’ strategies and their payoffs
  – objective: to predict what strategies players will choose
• Mainstream economists make comparatively little use of cooperative game theory, which
  – abstracts from strategy sets
  – focuses on coalitions
  – makes assumptions directly about the payoffs coalitions can attain
  – objective: to predict players’ payoffs (without necessarily saying how these are attained)
• Indeed if you look at modern game theory texts, see relatively little cooperative theory
  – Osborne and Rubinstein devote 2 chapters out of 15 to coalitions
    (3, if you include axiomatic bargaining theory)
  – Myerson has one chapter out of 10
  – Fudenberg and Tirole have no treatment of cooperative theory at all
• These texts accurately reflect the standing of cooperative theory in mainstream economics
• A shame
  – cooperative game theory a beautiful subject
  – potentially very important subject
• Its standing not predictable from early history
  – Von Neumann and Morgenstern devoted about \( \frac{3}{4} \) of book to cooperative theory
    (they clearly thought this was the future of subject)
  – influential later texts, e.g., Luce and Raiffa, Owen have about same balance
• Most surprising thing about absence of cooperative theory:
  coalitions play crucial rule in many important real phenomena
For example:

– economic mergers such as European Union
– military alliances like NATO
– trade agreements such as NAFTA
– international treaties to provide public goods such as Kyoto protocol
– cartels like OPEC
– even political parties
• Noncooperative theory can be used to study these phenomena too

  – but then tied to particular extensive or normal form
  – cooperative theory has potential for a more general perspective
So why hasn’t cooperative theory been more influential?

• perhaps too “normative”
  – axioms interpreted to say what should happen, not what does happen (economists interested in the positive)

• most solution concepts—including core and Shapley value (the 2 most important)—assume that grand coalition (coalition of everybody) forms, and outcome is fully Pareto optimal
But this assumption immediately rules out most applications of interest:

• European Union is not the only economic coalition in world - - - must compete against other such unions

• In democratic societies, there isn’t just one political party (even when coalitions of different parties form, rarely get grand coalition)

• Typically, many countries don’t sign international treaties like Kyoto protocol
• So grand coalition assumption is *empirically* suspect
• But will argue that it is *theoretically* suspect too
  – even when players free to enter into any sort of agreements they want, may well not reach grand coalition
  – in conflict with *Coase theorem*: unrestricted bargaining leads to Pareto optimum
Idea:

- If coalition $S$ generates *positive externalities* i.e., if a player can consume benefits created by $S$ without joining, may gain from *free riding* and committing *not* to negotiate with $S$ (enjoys benefits without incurring costs)
- Thus grand coalition will not form (and outcome not Pareto optimal).
• Propose a cooperative solution concept to handle games with externalities
• Solution concept is “generalization” of Shapley value
  – predicts players’ payoffs, just as Shapley value does
  – also predicts which coalitions form (Shapley value predicts grand coalition)
  – reduces to Shapley value if no externalities (this is always true if \( n \leq 3 \); requires some additional assumptions if \( n > 3 \))
• Like Shapley, will derive solution concept \textit{axiomatically}.

• Most of these axioms quite different from Shapley’s
  – motivated by “noncooperative implementation” of solution
  – reflect idea that Shapley value is natural outcome of “competitive” bargaining.
• Advantage of axiomatic approach:
  – not committed to one particular extensive or normal form (so more general)
  – see clearly what is driving the results

• Disadvantage: may not be apparent (from axioms) what sort of strategic (noncooperative) games correspond to axioms

• Nash program:
  – go back and forth between “general” (cooperative) and “specific” (noncooperative) approaches
  – investigate what kinds of noncooperative games implement given solution concept

• Attempt to carry out Nash program here
Game in characteristic function form

- $N = \{1, \ldots, n\}$ set of players

- $\nu : 2^N \to \mathbb{R}_+$ characteristic function

- For any coalition $S \subseteq N$

  $\nu(S) =$ what coalition $S$ can get on own
  
  = sum of utilities of players in $S$

- Presumes “transferable utility,” i.e., linear utility possibility frontiers.
linearity will hold if underlying utility functions are quasi-linear

\( \nu \) is reduced form for underlying game in which the problem is to choose some social alternative, e.g., allocation of resources among players

- players’ utility functions \( (u_1, \ldots, u_n) \) for social alternative and money give rise coalitions’ utility possibilities
- if \( (u_1, \ldots, u_n) \) quasi-linear (linear in money), then \( S \)'s possibilities can be summarized by number \( \nu(S) \)
• Assume superadditivity:
  if $S_1 \cap S_2 = \emptyset \quad \nu(S_1) + \nu(S_2) \leq \nu(S_1 \cup S_2)$
  (innocuous because $S_1$ and $S_2$ could act as before even if they merged)

• Solution concept:
  Given $\nu$, what payoff vector (or set of payoff vectors) is likely to result from bargaining amongst players?
  $$V \mapsto \phi(\nu),$$
  where $\phi(\nu) \in \mathbb{R}^n$ or $\phi(\nu) \subseteq \mathbb{R}^n$

• many solution concepts in literature
  – here concentrate on core and Shapley value
  – have already suggested Shapley value is “right”
Will start with the core

- partly because of its importance
- partly to discuss its problems
  - some are well-known
  - will emphasize one that is not

\[
\text{core}(v) = \left\{ (x_1, \ldots, x_n) \left| \sum_{i=1}^{n} x_i \leq v(N), \forall S \sum_{i \in S} x_i \geq v(S) \right. \right\}
\]

- first inequality is feasibility condition
- if \( \sum_{i \in S} x_i < v(S) \), \( S \) will “break away”

\[
\text{to get } v(S) \quad \text{- - so will block}
\]
• One well-known problem: core may not make sharp prediction, e.g., if $n = 2$

$$\text{core}(v) = \{(x_1, x_2) \mid x_1 + x_2 = v(1, 2), \ x_i \geq v(i) \quad i = 1, 2\}.$$  

• Another—almost opposite—problem: core may be empty in “well-behaved” games
• Suppose $n = 3$

• $v(i) = 6, \ i = 1, 2, 3,$

$$v(1, 2) = 16, \ v(1, 3) = 16, \ v(2, 3) = 17, \ v(1, 2, 3) = 24$$

• if $(x_1, x_2, x_3) \in \text{core}(v)$

\[
\begin{align*}
    x_1 + x_2 &\geq 16 \\
    x_1 + x_3 &\geq 16 \\
    x_2 + x_3 &\geq 17 \\
    2(x_1 + x_2 + x_3) &\geq 49, \ 	ext{i.e.,} \ (x_1 + x_2 + x_3) \geq \frac{49}{2}, \ \text{violating feasibility}
\end{align*}
\]

• So core is empty
• Problem to be emphasized: core disregards *sequential* nature of bargaining
• Can be seen from noncooperative implementation

• A (noncooperative) game form \( g : S_1 \times \ldots \times S_n \rightarrow A \)

• (Nash) *implements* core if, for all \((u_1,\ldots,u_n) \leftrightarrow v,\)

\[
NE_g(u_1,\ldots,u_n) = \text{core}(v)
\]
• Core is *monotonic*: if social alternative $a$ is in core for $(u_1, \ldots, u_n)$, and then utility functions changed to $(u'_1, \ldots, u'_n)$, so that $a$ does not fall in any player’s preferences (i.e., $u_i(a) \geq u_i(b) \Rightarrow u'_i(a) \geq u'_i(b)$ for all $i$ and $b$), then $a$ in core for $(u'_1, \ldots, u'_n)$

• General theorem from implementation theory implies that core is (Nash) implementable

• But construction of implementing game form invokes no property of core other than monotonicity - - same construction used for *any* monotonic solution concept

• So this noncooperative implementation reveals nothing that is special to *core*
• Want noncooperative game form whose very rules capture the essence of core
• Such a game form devised by Perry and Reny (1992)
Game form:

- continuous time
- at any time $t > 0$, any player $i$ can make proposal $(x, s)$ to coalition $S$,
  where $\sum_{i \in S} x_i \leq v(S)$
- if accepted by members of $S$ (who respond sequentially), proposal becomes binding agreement
- new proposal invalidates any earlier nonbinding proposal
- proposal can include players from $S'$ who have binding agreement, but then must include all of $S'$
- whenever player moves (proposes, accepts, or rejects), minimum time gap $\Delta$ before any subsequent move (implies there exists a first move)
- outcome determined by binding agreements in force at $t = \infty$
Stationary subgame-perfect equilibrium:
- strategies do not depend on previous (invalidated) proposals
- strategies do not depend on time

Proposition: Payoffs $x$ correspond to stationary spe if and only if

$$x \in \text{core} (v)$$
Proof: for $x \in \text{core} \ (v)$, construct equilibrium in which

- player 1 (say) proposes $(x,N)$ and everyone accepts
- if anyone rejects, 1 proposes $(x,N)$ again
- if someone proposes $(y,S)$, must have
  $$y_i < x_i \text{ for some } i \in S$$
  $$(x \in \text{core} \ (v))$$, so $i$ will reject
Suppose have equilibrium with payoffs \( x \notin \text{core}(\nu) \)

- there exist \( y \) and \( S \) such that
  \[
  \sum_{i \in S} y_i \leq \nu(S) \quad \text{and} \quad y_i > x_i \quad \text{for all } i \in S
  \]

- some \( i \in S \) can propose \((y, S)\) \textit{before} any other proposal is made

- everyone in \( S \) will accept
  - \( j \in S \) would reject only if \((y', S')\) proposed and \( j \) does even better than \( y_i \) in continuation equilibrium
  - But then \( j \) could propose \((y', S')\) before any other proposals made and do better than \( x_i \)

- so equilibrium destroyed by \((y, S)\)
• Argument shows that core relies on idea that coalition could always move earlier than everyone else
• But this constitutes weakness in concept: in reality, may be a limit to how early a coalition can move
• This way of viewing core explains why it is so often empty (fails to make prediction):
  – in many bargaining situations, strictly positive advantage to moving first
  – conflicts with core’s presumption that everyone can do so
• Resolution: assume that if everyone wants to move first, order of moves determined by chance
• How does this approach apply to example in which core is empty?

\[ v(1,2) = 16, \ v(1,3) = 16, \ v(2,3) = 17, \ v(1,2,3) = 24 \]

\[ v(1) = v(2) = v(3) = 6 \]

• Suppose player 1 gets to move first, then 2, then 3.

• Think of 1 “arriving” first, then 2, then 3.
What must player 1 offer 2 to induce him to join in a coalition?

- If 2 doesn’t join, will compete with 1 for 3
- 1 will offer up to 16 – 6 = 10
- 2 will offer up to 17 – 6 = 11
- Hence, 2 will win and pay 10
- Thus, 2 must be offered 17 – 10 = 7 to join
• If player 2 signed up, what must \{1, 2\} offer 3 to join?
  – player 3’s only other option is to remain alone -
    - gets 6
  – So \{1, 2\} must offer 6
• All told, player 1 gets 24 – 7 – 6 = 11 from
  signing up 2 and 3
• Because 11 > 6, 1 will do so, i.e., grand
  coalition forms, even though core is empty.
• predicts payoffs (11, 7, 6)
• not in core
  \[ \nu(2, 3) = 17 \]
• core ignores possibility that 1 and 2 may have chance to bargain before 3 does
  – 1 and 2 have binding agreement before 3 arrives
  – so, fact that \( \nu(2, 3) = 17 \) no longer relevant (3 would have to pay player 1 compensation of \( 11 - 6 = 5 \) to lure 2 away)
• Order 1, 2, 3 just one (random) possibility
• 5 other orders
• Can perform same sort of payoff computation for others, e.g., for 2, 1, 3, get (12, 6, 6)
• If each order is equally likely, expected payoffs are \( \left( \frac{45}{6}, 8, \frac{51}{6} \right) \)
  – these are players’ Shapley values
  – no coincidence: will argue that Shapley value is good prediction for competitive bargaining process with random order.
• For any \( v \), Shapley value \( \varphi(v) \) defined so that

\[
\varphi_i(v) = \frac{1}{n!} \sum_{R} \left[ v(S_i(R) \cup \{i\}) - v(S_i(R)) \right],
\]

where \( R \) is ordering of \( N \), and \( S_i(R) \) is coalition before \( i \) in ordering \( R \)

• E.g., if \( n = 3 \), \( i = 1 \) and \( R = 2, 1, 3 \)

\[
v(S_i(R) \cup \{1\}) - v(S_i(R)) = v(1, 2) - v(2).
\]
Shapley value has been axiomatized in several ways

- Shapley’s axiomatization
  - Pareto optimality
    \[ \sum_{i \in N} \varphi_i(v) = v(N) \]
  - Anonymity
    if \( \pi : N \to N \) is one-to-one and \( v_\pi(S) = v(\pi(S)) \) then
    \[ \varphi_{\pi(i)}(v_\pi) = \varphi_i(v) \]
– Dummy property

\[ \text{if } v(S \cup \{i\}) - v(S) = 0 \text{ for all } S, \]

\[ \text{then } \varphi_i(v) = 0 \]

– linearity

\[ \varphi(v + v') = \varphi(v) + \varphi(v') \]
Shapley’s Theorem: function $\varphi : v \mapsto (\varphi_1(v), \ldots, \varphi_n(v))$

satisfies Pareto optimality, anonymity, dummy property, linearity if and only if it satisfies (*)

- first 3 axioms seem plausible properties of competitive bargaining
- hard to see why linearity should hold
- Nevertheless, claim that Shapley value makes good predictions about competitive bargaining in broad class of cases
  - will provide alternative axiomatization tomorrow
Yet Shapley value does not *always* make good predictions

- in previous example, what coalition can get independent of what other players do
- allows representation by characteristic function
- this rules out *externalities*
Public Good game (based on Ray and Vohra (1999))

- \{1,2,3\} can get 24
- if \{1,2\} forms, \{1,2\} gets 12, 3 gets 9
- if \{1,3\} forms, \{1,3\} gets 13, 2 gets 9
- if \{2,3\} forms, \{2,3\} gets 14, 1 gets 9
- if no coalitions forms, all players get 0.

If coalition of two produces public good, remaining player can enjoy benefit without bearing cost
What must player 1 offer 2 to induce him to join?

• if 2 doesn’t join, 1 and 2 will compete for 3
• 2 will offer up to $14 - 9 = 5$
• 1 will offer up to $13 - 9 = 4$
• so 2 will “win” 3 and pay 4
• thus, 1 must offer 2 $14 - 4 = 10$
Next, given that player 2 has joined, what must \{1, 2\} offer 3?
  – 3 can get 9 since \{1, 2\} has formed
  – so \{1, 2\} must offer 9

By signing up 2 and 3, player 1 obtains 24 – 10 – 9 = 5

But, 1 can get 9 by refusing to negotiate with 2 and 3
  – forces 2 and 3 to merge

Thus, predict coalitions \{1\} and \{2, 3\} will form - grand coalition does not arise
• This is an “anti-Coasian” conclusion, so should examine ingredients carefully
• Coase argued that, if players have opportunity to bargain freely, they should reach Pareto optimum
  – everyone better off replacing a non-Pareto optimal agreement with one that Pareto-dominates it
• What interferes with this argument?
(1) Presence of positive externalities

- 2 can free-ride on externality created by \{1, 3\}
- 3 can free-ride on externality created by \{1, 2\}
- so both must be offered a lot to join 1
- leaves 1 with so little, better off free-riding on them
(2) Irreversibility of 1’s decision not to negotiate

• forces 2 and 3 to merge
• if 1 could change mind, 2 and 3 might not merge
  – if did merge, would then sign up 1
    (gross payoff goes up by 10; only need to pay 9 to player 1)
  – 2 and 3 get 24 – 9 = 15
  – so at least one of them gets less than 8
  – could free-ride on others and get 9
• Thus, if I couldn’t commit not to negotiate
  – might cause delay in coalition formation, hence inefficiency
  – but, grand coalition would form eventually
How to interpret commitment not to negotiate?

• suppose bargaining requires
  – communication lines between bargainers
  – diplomatic relations between 2 countries

• decision not to negotiate
  – severing the lines
  – cutting off diplomatic relations

• in practice, communication lines may be reparable, but could be very costly to do so
But if player can commit not to negotiate, why can’t players make a *contingent* agreement to break off relations?

- In public good example, why can’t 1 and 2 agree to offer player 3 payoff of 4 and break off relations with *each other* if he doesn’t accept?
  - would force player 3 to accept 4 (otherwise he—and others—get nothing)
  - 1 and 2 could get 10 each

- But agreement between 1 and 2 not renegotiation-proof
  - suppose 3 broke off relations with 1 and 2
  - 1 and 2 would revise decision to cut line between them
• Thus presence of positive externalities plus ability to commit not to negotiate imply
  – grand coalition may not form
  – outcome may not be Pareto optimal even though bargainers have opportunity to negotiate without constraint

• Argument can be viewed as a qualification to Coase theorem
• Just an example and a noncooperative one
• Tomorrow will develop a more general (cooperative) model for accommodating externalities in bargaining
• Certainly not first to study possibility of grand coalition not forming
• Literature that had particularly strong influence on me:
  Hart and Kurz (1983)
  Aumann and Myerson (1988)
  Greenberg and Weber (1993)
  Chwe (1994)
  Bloch (1996)
  Yi (1997)
• Contribution here: to follow Nash program
  – develop both a cooperative (axiomatic) and noncooperative (strategic) theory of bargaining when externalities may interfere with efficiency