EC516 Contracts and Organisations, LSE

3. Lecture Notes on Binding Agreements

Debraj Ray

1. Coalition Formation in Symmetric Games

1.1. **Standard Equilibrium and the Algorithm.** This is the theorem with which we ended the previous lecture:

Theorem **1.** *There exists $\delta^* \in (0,1)$ such that for all $\delta \in (\delta^*, 1)$, any standard equilibrium must be of the following form. At a stage in which any substructure π has left the game (with associated numerical substructure $\mathbf{n}$), the next coalition that forms is of size $t(\mathbf{n})$ and the payoff to a proposer is*

$$a(\mathbf{n}, \delta) \equiv \frac{v(t(\mathbf{n}), c(\mathbf{n}))}{\delta t(\mathbf{n}) + (1-\delta)]}. \tag{1}$$

In particular, the numerical coalition structure corresponding to any such equilibrium is $\mathbf{n}^*$.

Proof. To prove Theorem 1 our first task is to fix δ^*. This is done by the help of the following result.

Lemma **1.** *There exists $\delta^* \in (0,1)$ such that for any $\delta \in (\delta^*, 1)$, and any substructure $\mathbf{n}$, $t(\mathbf{n})$ uniquely maximizes*

$$\frac{v(t, c(\mathbf{n}.t))}{1+\delta[t-1]}. \tag{2}$$

Proof. Fix such a substructure and consider the maximization of the expression in (2). By a standard argument (the maximum theorem), the set of maximizers is uhc in δ, and so all limit points of sequences from this set must lie in the set of maximizers of

$$\frac{v(t, c(\mathbf{n}.t))}{t},$$

as $\delta \to 1$. But the maximizers are just integers from some finite set, so all this limit stuff can be dropped, and in fact the set of maximizers of (2) — call the set $\mu(\mathbf{n}, \delta)$ — must become a *subset* of $\mu(\mathbf{n}, 1)$ for all δ large enough; bigger than some $\delta^* \in (0,1)$.

So for every $\delta > \delta^*$ and for every maximizer t of (2) the value of

$$\frac{v(t, c(\mathbf{n}.t))}{t},$$

is just the same. Comparing this expression and (2), it is obvious, then, that the maximizer of (2) must indeed exhibit the largest value of

$$\frac{t}{1+\delta[t-1]}$$

among all values of t that achieve the maximum of $v(t, c(\mathbf{n}.t))/t$. But this fraction strictly increases in t, so we've proved that for all $\delta > \delta^*$ there is only one maximizer of (2), and that is $t(\mathbf{n})$ as given by the algorithm. $\square$

LEMMA **2.** *Consider a stage in which π has left the game and S is the set of active players. Let $\mathbf{n}$ denote the numerical coalition structure corresponding to π, and let $(x_i)_{i \in S}$ denote the equilibrium payoffs to each active player if she is the proposer at this stage. Suppose that for any $t \in \{1, \ldots, n - K(\mathbf{n})\}$ the numerical coalition structure following $(\mathbf{n}.t)$ is $c(\mathbf{n}.t)$. Then, if i makes an acceptable proposal to a coalition T (that includes herself) with positive probability, $x_i \leq x_k$ for all $k \in S$.*

Proof. Let i makes an acceptable proposal to T (which includes herself) with positive probability. Pick $k \neq i$. Suppose first that $k \notin T$. Imagine that k makes an offer to $\{T - i\} \cup k$, and gives everyone slightly more than what i was giving them; then this is strictly acceptable to all, because after acceptance the continuation numerical structure is exactly the same, by the assumption of the lemma. Therefore $x_k \geq x_i - \epsilon$ for all $\epsilon > 0$, which is just another way of saying that $x_k \geq x_i$.

Now, if $k \in T$, she too can propose T acceptably (again, an ϵ-argument of the sort above will suffice). Consequently, writing t for the cardinality of t,

$$\begin{aligned} x_k &\geq v(t, c(\mathbf{n}.t)) - \delta \sum_{j \in T; j \neq k} x_j \\ &= v(t, c(\mathbf{n}.t)) - \delta \sum_{j \in T; j \neq i} x_j + \delta x_k - \delta x_i \\ &= x_i + \delta x_k - \delta x_i, \end{aligned}$$

where the last line follows from the fact that i's proposal to t does attain her equilibrium payoff. Now rearrange this inequality to see that

$$x_k - x_i \geq \delta(x_k - x_i),$$

which simply means that $x_k \geq x_i$. $\square$

The above Lemma is crucial. Notice that if i does not make a proposal to a coalition that includes herself with positive probability, we cannot use the argument above to draw the conclusion that $x_k \geq x_i$ for all $k \neq i$. Indeed, that conclusion is false; we will soon see an example of this. But first let us finish the proof of the theorem.

Fix an equilibrium as described in the statement of the theorem, and let $\delta \in (\delta^*, 1)$, with δ^* as in Lemma 1. We proceed by induction on the cardinality of the set of active players, following the departure of any collection of players. If there is one active player left, then there is nothing to prove. Inductively, suppose that the theorem is valid at every stage with $K(\mathbf{n}(\pi)) = m + 1, \ldots, n - 1$ for some $m \geq 0$.

Consider, now, a stage with $K(\mathbf{n}(\pi)) = m$. Let S be the set of active players, and let $\{x_j\}_{j \in S}$ denote the vector of equilibrium payoffs to player j if j is the proposer at this stage. Let T^* be a coalition that forms at this stage (with cardinality t^*), and let i be the proposer. We need to prove that

$$t^* = t(\mathbf{n}(\pi)). \tag{3}$$

Since every player in S makes an acceptable proposal to some coalition with positive probability, it follows immediately from the induction hypothesis and Lemma 2 that $x_j = x_i = x$ for all $i, j \in S$. It follows from the induction and the optimality of the proposal that

$$x = v(t^*, c(\mathbf{n}(\pi).t^*)) - \delta(t^* - 1)x \geq v(t, c(\mathbf{n}(\pi).t) - \delta(t-1)x,$$

for all $t \in \{1, \ldots, n - K(\mathbf{n}(\pi))\}$.

But this observation implies that x must also be the maximum value of the expression in (2), as t varies over the set $\{1, \ldots, n - K(\mathbf{n}(\pi))\}$. Using Lemma 1, we may conclude that $t^* = t(\mathbf{n})$. Of course, the payoff to a proposer is $a(\mathbf{n}, \delta)$, as defined in (1). This completes the proof of Theorem 1. □

1.2. **An Example.** Theorem 1 is useful, in that it links the equilibria of the bargaining game to the coalition structure predicted by us. But the link is made by a particular kind of equilibrium, those in which an acceptable proposal is made in each stage to a coalition containing the proposer; the so-called "standard equilibria". How standard are these standard equilibria?

To get a sense of this, consider the following example:

$$\mathbf{v}(4,1) = (6,2), \qquad \mathbf{v}(3,2) = (3,8), \qquad \mathbf{v}(2,1,1,1) = (0.1,3,3,3),$$
$$\mathbf{v}(3,1,1) = (10,0,0), \qquad \mathbf{v}(\pi) \simeq 0 \text{ for all other } \pi.$$

Apply algorithm: $\mathbf{n}^* = (4,1)$.

There isn't actually a standard equilibrium in this example. For if there is one, then four-person coalitions are always made offers and proposer payoffs are given by (1). It will then pay a respondent to actually make offers to the remaining four persons, get them out of the way, and pick up $\delta 2$ instead of approximately 1.5.

Yet there is still an equilibrium yielding the structure (4,1), which is asymmetric: one player makes proposals to the other four and the other four make proposals to one another. Under this equilibrium, the intransigent player receives 2δ whenever it is his turn to propose to the grand coalition. The others receive only $\frac{6}{1+3\delta}$ in their roles as proposer.

[Exercise: Check that this is indeed part of an equilibrium strategy.]

Yet there is *also* an equilibrium with coalition structure $(3,2)$. It is constructed as follows. Players 1, 2 and 3 make acceptable offers to each other and the other two instruct the coalition $\{123\}$ to form by equally dividing their worth, and they accept. Let $\bar{x}_i$, the equilibrium payoff to i if i starts the game, be defined as

$$\bar{x}_i = \frac{3}{1+2\delta} \text{ for i =1,2,3}$$

$$\bar{x}_j = \frac{8\delta}{1+\delta} \text{ for j=4, 5.}$$

For δ close to 1, players 1, 2 and 3 get approximately 1 while players 4 and 5 get approximately 4. Clearly, player i, $i = 1, 2, 3$ cannot do better by including player 4 or 5, since $v(4, 1) = (6, 2)$. Given the strategies of the others, i cannot do better by making some other proposal. It is also easy to see that players 4 and 5 do not have a profitable deviation. Thus, the above strategies (together with obvious specifications for non-equilibrium subgames) constitute an equilibrium.

1.3. **Existence and Uniqueness of Standard Equilibrium.** These examples raise the following questions:

1. When does a standard equilibrium exist (so that $\mathbf{n}^*$ is always *a* prediction)?[1]

2. When is $\mathbf{n}^*$ the *only* predicted outcome?

1.3.1. *Existence.* Recall that for each numerical substructure $\mathbf{n}$,

$$a(\mathbf{n}) \equiv \frac{v(t(\mathbf{n}), c(\mathbf{n}.t(\mathbf{n}))}{t(\mathbf{n})}.$$

The numbers $a(\mathbf{n})$ can, of course, be directly computed from the primitives of the model.

Say that algorithmic average worth is *weakly nonincreasing* if

$$a(\mathbf{n}) \geq a(\mathbf{n}.t(\mathbf{n}))$$

for every substructure $\mathbf{n}$ such that $\mathbf{n}.t(\mathbf{n})$ is also a substructure.

The following theorem settles the existence question.

Theorem **2.** *A pure strategy standard equilibrium exists for discount factors close to one if and only if algorithmic average worth is weakly nonincreasing.*

Proof. The proof will reply on the following technical lemma.

Lemma **3.** *Assume that average algorithmic worth is weakly nonincreasing. Then there exists* $\hat{\delta} \in (\delta^*, 1)$ *such that for all numerical substructures* $\mathbf{n}$ *and positive integers* $t_1, \ldots, t_k$ *with* $\mathbf{n}.t_1.\ldots.t_k$ *also a substructure,*

$$a(\mathbf{n}, \delta) > \delta a(\mathbf{n}.t_1.\ldots.t_k, \delta) \text{ for all } \delta \in (\hat{\delta}, 1), \tag{4}$$

where $a(\mathbf{n}, \delta)$*, it will be recalled, is defined in (1).*

Without getting into the proof of the lemma just try and understand what it means; that's half the battle won. Recall that $a(\mathbf{n}, \delta)$ is like a perturbation of the average worth — it is the proposer's payoff in a standard equilibrium — which converges to the maximal average worth available at the substructure $\mathbf{n}$ as $\delta \to 1$. Likewise for $a(\mathbf{n}.t_1.\ldots.t_k, \delta)$. Now, just because the inequality

$$a(\mathbf{n}) \geq a(\mathbf{n}.t_1.\ldots.t_k) \tag{5}$$

[1]Note: question 1 slides a more fundamental question under the rug, which is this: is n^* always *an* equilibrium outcome, standard or not standard? [Recall the first equilibrium in the above example, which is not standard and yet generates $\mathbf{n}^*$.] This is an open question.

holds *at* $\delta = 1$ (this follows from weak nonincreasingness) does *not* imply — right away — that the same inequality will hold in the "δ-version of these expressions, even in the form (4) where we are helped along by having the right hand side multiplied by δ. This needs to be proved.

Observe that if the inequality in (5) is strict, this is guaranteed simply by continuity.

For the proof of this lemma, see Ray and Vohra (1999), page 316. Now return to the proof of the theorem.

First assume that algorithmic average worth is weakly nondecreasing. Pick any $\delta \in (\hat{\delta}, 1)$, where $\hat{\delta}$ is given by Lemma 3.

Consider any stationary strategy σ as follows: In every subgame following the departure of π, player i makes a proposal to a coalition of size $t(\mathbf{n}(\pi))$ that contains himself. He offers to each partner a payoff $\delta a(\mathbf{n}, \delta)$ in the event that the numerical coalition structure $c(\mathbf{n})$ is formed, and any other payoff division otherwise. All such offers are accepted by respondents (other responses are described in the obvious way: for a description, see (ii) and (iii) in Ray and Vohra (1999), page 312). We will show that σ is an equilibrium.

To this end, consider any stage described by π. Along the proposed strategy profile σ a proposer receives $a(\mathbf{n}(\pi), \delta)$. Therefore, the only way that a proposer can possibly deviate gainfully is by making an unacceptable proposal, or a proposal to a coalition that does not include him. Given the strategies of the other players, this will result in the formation of coalitions of cardinalities $t(\mathbf{n}), t(\mathbf{n}.t(\mathbf{n}))$, and so on. Thus the deviant proposer will ultimately receive a payoff that is bounded above by $\delta a(\mathbf{n}.t_1 \ldots . t_k, \delta)$, where $t_1 \ldots . t_k$ is a finite string of the form $t(\mathbf{n}).t(\mathbf{n}.t(\mathbf{n}))...$. Applying the nonincreasing average worth condition repeatedly, we see that

$$a(\mathbf{n}) \geq a(\mathbf{n}.t_1 \ldots . t_k).$$

But then, by Lemma 3 and the fact that $\delta > \hat{\delta}$, we conclude that (4) holds. This means that the deviation cannot be profitable.

It is now easy to see that as a responder, a player cannot gainfully deviate from σ. Consequently, σ is an equilibrium.

Now we prove the converse: that the existence of a standard equilibrium for all δ close enough to 1 must imply that average algorithmic worth is weakly nonincreasing.

Suppose, on the contrary, that there is $\hat{\delta} \in (0, 1)$ such that for all $\delta \in (\hat{\delta}, 1)$ there exists a pure strategy standard equilibrium but the algorithmic worth condition fails. This means that there exists a numerical substructure $\mathbf{n}$ such that $\mathbf{n}.t(\mathbf{n})$ is also a substructure, and such that $a(\mathbf{n}.t(\mathbf{n})) > a(\mathbf{n})$. It follows that there exists $\bar{\delta} \in (0, 1)$ such that

$$a(\mathbf{n}.t(\mathbf{n}), \delta) > a(\mathbf{n}, \delta) \text{ for all } \delta \in (\bar{\delta}, 1). \tag{6}$$

Consider any $\delta > \max\{\hat{\delta}, \bar{\delta}\}$, and fix some pure strategy standard equilibrium σ. Consider any subgame where π has left, where $\mathbf{n}(\pi) = \mathbf{n}$. Let i be the first proposer in this subgame. Since σ is a pure strategy standard equilibrium, and $\delta \geq \hat{\delta} \geq \delta^*$, i makes an acceptable proposal to some determinate coalition T of size $t(\mathbf{n})$ that includes him. Because $\mathbf{n}.t(\mathbf{n})$ is also a substructure, there must exist a player j who is not included in the proposal by player i, and thereafter picks up a present value of $a(\mathbf{n}.t(\mathbf{n}), \delta)$ in the very next stage.

Now consider another subgame (in the same stage) so that exactly the same set of players have left (and in the same structure), but j is the first proposer instead of i. Because σ is pure strategy standard, j is also supposed to make an acceptable proposal to a coalition of size $t(\mathbf{n})$, picking up $a(\mathbf{n}, \delta)$. However, suppose that she deviates by proposing that the coalition T (which i proposed in the last para, and that did not include j) form, using exactly the same proposal that i used, or even equal division. This offer should be accepted. By stationarity, we are then in the precise situation of the preceding paragraph. Thus by making an acceptable proposal to player i's would-be coalition, as it were, j receives a present value of $a(\mathbf{n}.t(\mathbf{n}), \delta)$. By (6) this deviation is profitable. This contradicts the fact that we have an equilibrium, and completes the proof of the theorem. □

1.3.2. *Uniqueness.* Notice that our theorem only asserts the *existence* of a standard equilibrium under some conditions. It does not exclude the possibility that there may be other equilibria yielding entirely different coalition structures. To see this, consider the following modification of the example in Section 1.2. Modify that partition function so that $\mathbf{v}(4,1) = (6,1)$. Again, $t(\phi) = 4$. But now $a(4) = 1 < a(\phi)$ and it is easy to see that the weak nonincreasing average worth condition holds. So there exists a standard equilibrium with the coalition structure $(4,1)$. However, the non-symmetric equilibrium with the coalition structure $(3,2)$ continues to be an equilibrium here as well.

This discussion makes it clear that uniqueness needs more than the condition that average algorithmic work is weakly nonincreasing. To state this additional requirement, we need to extend a bit the definition of $t(\mathbf{n})$.

At $\mathbf{n}$, say t is a *restricted maximizer* if it maxes

$$\frac{v(t, c(\mathbf{n}.t))}{t}$$

subject to some upper bound on t. Now say that algorithmic average worth is *strongly nonincreasing* if

$$a(\mathbf{n}) \geq a(\mathbf{n}.t)$$

for every substructure $\mathbf{n}$ and restricted maximizer t such that $\mathbf{n}.t$ is also a substructure.

Notice that the *unrestricted* maximizer $t(\mathbf{n})$ is also a restricted maximizer for some large bound. Therefore the new condition implies the old, weak version of nonincreasingness.

Theorem **3.** *If algorithmic average worth is strongly nonincreasing, then for discount factors close to one every equilibrium is standard, and* $\mathbf{n}^*$ *is the unique numerical coalition structure.*

Proof. First we fix a threshold δ. For any substructure $\mathbf{n}$ and for any t that is *not* a restricted maximizer, there is a restricted maximizer (using bound t), call it t', such that

$$\frac{v(t', c(\mathbf{n}.t'))}{t'} > \frac{v(t, c(\mathbf{n}.t))}{t}.$$

As there are only a finite number of these situations, we can pick δ_1 close enough to 1 such that for all of these situations,

$$\delta_1 \frac{v(t', c(\mathbf{n}.t'))}{t'} > \frac{v(t, c(\mathbf{n}.t))}{t} + (1 - \delta_1)M \tag{7}$$

where M is the maximum possible payoff to anyone in the game.

Next, recall $\hat{\delta}$ as given in Lemma 3. Finally, choose our threshold to be the larger of the two numbers $\hat{\delta}$ adn δ_1.

Fix any equilibrium. We will show that it must be standard. The proof is by induction on the cardinality of the set of active players. At every stage when there is only one active player left, the subgame equilibrium is trivially standard. Now suppose that for any π such that $K(\mathbf{n}(\pi)) \geq m+1, \ldots, n-1$, for some $m \geq 0$, the subgame equilibrium is standard. Consider a stage described by a structure of departed players, π, with the property that $K(\mathbf{n}(\pi)) = m$. Let $\mathbf{n} \equiv \mathbf{n}(\pi)$. Let S be the set of active players. Let $\{x_i\}_{i \in S}$ denote the vector of equilibrium payoffs to each player, if he is the proposer at this stage.

If all these players make acceptable offers to coalitions that include themselves (with positive probability), then we are done. So there must be some player who makes an acceptable offer $\mathbf{y}$ to a coalition T, of size t, that does not include himself.[2] There are two possibilities.

Case 1. t is not a restricted maximizer. Let t' be some restricted maximizer in a restricted problem with bound t. Now imagine that one of the worst-treated members of the acceptable proposal — call him i — for T were to reject that proposal and make a proposal to the set T' of himself and his $t' - 1$ worst-treated compatriots. Say this proposal offers each of his compatriots the amount

$$y_j + \epsilon,$$

where y_j is what j was getting under the original proposal, and $\epsilon > 0$ is a small number which we shall pin down more carefully below. If the previous proposal was accepted by the compatriots, *this one must be too.* By the induction hypothesis, our deviant's payoff, discounted for the cost of one period of rejection of the T-proposal, *minus* what he was

[2]It is not possible that a player makes an unacceptable offer. He could simply make the acceptable offer that would be made after him and speed the process up.

getting under the original proposal to T, is

$$\begin{aligned}
& \delta\left[v(t', c(\mathbf{n}.t')) - \sum_{j\in T', j\neq i}(y_j+\epsilon)\right] - y_i \\
= \ & \delta\left[v(t', c(\mathbf{n}.t')) - \sum_{j\in T', j\neq i}(y_j+\epsilon) - y_i\right] - (1-\delta)y_i \\
= \ & \delta\left[v(t', c(\mathbf{n}.t')) - \sum_{j\in T'} y_j - (t'-1)\right] - (1-\delta)y_i \\
\geq \ & \delta\left[v(t', c(\mathbf{n}.t')) - t'\frac{v(t, c(\mathbf{n}.t))}{t} - (t'-1)\epsilon\right] - (1-\delta)y_i
\end{aligned}$$

where this last inequality comes from the fact that our members of T' were the t' *worst-treated* members of T. And now, if ϵ is chosen positive but tiny, we can continue the chain of inequalities ...

$$\begin{aligned}
\geq \ & \delta v(t', c(\mathbf{n}.t')) - t'\frac{v(t, c(\mathbf{n}.t))}{t} - (1-\delta)y_i \\
> \ & t'\frac{v(t, c(\mathbf{n}.t))}{t} - t'\frac{v(t, c(\mathbf{n}.t))}{t} + (1-\delta)(M - y_i) \\
> \ & 0,
\end{aligned}$$

where the penultimate inequality comes from (7). But now we have shown that there is a profitable deviation from the equilibrium, so that Case 1 cannot hold. What remains is

Case 2. t is a restricted maximizer. First notice that the proposer s can, without loss of generality, be presumed to have the largest equilibrium payoff among all proposers in S (this follows from Lemma 2). However, because the proposer is making a proposal to a coalition that does not include him, his equilibrium payoff x_i must satisfy

$$x_s \leq \delta a(\mathbf{n}.t \ldots t_{k-1}, \delta) < a(\mathbf{n}, \delta), \tag{8}$$

using the fact that t is a restricted maximizer, and then applying strong nonincreasingness, Lemma 3 and the induction hypothesis. Because we took s to have the highest equilibrium payoff, it follows that

$$x_j < a(\mathbf{n}, \delta)$$

for everyone else as well. But now we have a contradiction, for this proves that s can make an acceptable proposal to a coalition of size $t(\mathbf{n})$, and — using induction one more time — can pick up a larger payoff. This completes the proof.

1.4. **A Cournot Oligopoly.** We apply our results to an example of a symmetric Cournot oligopoly. Suppose that n oligopolists produce a quantity x of a homogeneous product, the price P of which is determined by a linear demand curve: $P = A - bx$. Assume that there is a fixed unit cost of production, given by c.

Normalize the parameters so that $\frac{(A-c)^2}{b} = 1$. Using the formula for Cournot-Nash equilibrium, recall that the partition function for this symmetric game is

$$v(s, \mathbf{n}) = \frac{1}{(q+1)^2}.$$

where q is the number of coalitions in $\mathbf{n}$.

THEOREM **4.** *All equilibria in a Cournot oligopoly with n firms are standard equilibria. So there is a unique numerical equilibrium coalition structure. It consists of L singleton firms and a single cartel of size $n-L$, where L is the smallest nonnegative integer such that*

$$n - L < (L+2)^2 + 1.$$

Thus our results predict full cartelization in this example whenever there are 4 firms or less, and imperfect cartelization thereafter.

For some intuition, we invoke an important observation due originally to Salant, Switzer and Reynolds (1983): If several firms are already out of a potential cartel, and the number of firms left is "small enough", then the remaining firms will not find it advantageous to form a cartel. Intuitively, the gain in market concentration does not justify the profit-sharing that will be needed. Applying this idea recursively to the remaining number of players, we can find a threshold at which the average payoff to the remaining players, if they stay together, is approximately the same as when a player quits, sparking off a cartel collapse.

Summarizing so far, we see that at this threshold, firms would rather stay together than break up. But *knowing this is so*, those firms in excess of this threshold will disagree to form a cartel as well, predicting correctly that the remaining firms will stay together. This creates an equilibrium outcome with one large cartel and several singleton firms.

The next lecture will haev the proof of this proposition and another example.