Nash Bargaining in Coalitional Games:

Supplementary Notes

Rajiv Vohra ⓡ Debraj Ray

April 2025

These notes discuss: (a) the independence of the various axioms considered in the main text, (b) an extension of Theorem 2 in the main text to allow for feasible sets that satisfy log subconvexity rather than convexity, (c) details from the examples in Section 4, and (d) an extension of the concept of viable coalition structures to accommodate external threats. Numbered equations in these Notes are of the form (a.xx), and all other referenced equations are from the main text.

1. Independence of the Axioms Used For Theorem 1

We've shown in the main text that [Exp] is independent of the other axioms. We now consider the independence of each of the other axioms in turn. In what follows, we normalize $d_i = 0$ for all i, which means that [Inv] will only refer to scale invariance.

Independence of [UHC]. Say that a threat configuration $\mathbf{\Theta}$ is *fully effective* (FE) if for every i, there is some coalition $S \ni i$ and some $x \in \Theta(S)$ such that $x_i > 0$. In what follows we work with the space of all FE threat configurations.

Define two FE threat configurations $\mathbf{\Theta}$ and $\mathbf{\Theta}'$ to be *connected* if there is $\lambda \gg 0$ such that for every coalition S,

$$\Theta'(S) = \{x \in \mathbb{R}^S | x = \lambda \circ y \text{ for some } y \in \Theta(S)\}.$$

Connectedness is an equivalence relation that partitions the space of all FE threat configurations. For each element H of this partition, we can pick a representative $\mathbf{\Theta}_H$ such that for every i, the maximum payoff (over all coalitions containing i, and all allocations in those coalitions) equals 1. For by the compactness of every threat set, that maximum payoff is finite, and because the threat constellation is fully effective, we can always find a multiplicative transform of every agent's payoff such that that maximum equals 1. Notice also that this requirement uniquely pins down $\mathbf{\Theta}_H$ from every element or equivalence class H.

Now define for every i, every class H, and every coalition S within $\mathbf{\Theta}_H$ with $S \ni i$, $m(i, S, H) \equiv \max\{x_i | x \in \Theta(S)\}$, and and let $\alpha(i, H)$ be the average value of $m(i, S, H)$

over all such S. Now define for every game G, a new solution $\phi(G)$ by

$$\phi(G) = \begin{cases} \sigma^{\text{sym}}(G) & \text{if } \boldsymbol{\Theta} \text{ is not fully effective} \\ \arg\max_{x \in U(G)} \prod_{j \in N} x_j^{\alpha(j,H)} & \text{if } \boldsymbol{\Theta} \text{ is fully effective and belongs to class } H \end{cases}$$

It is easy to verify that ϕ satisfies all the axioms except for UHC. It fails UHC for sequences of fully effective threat constellations converging "down" to an ineffective constellation. Continuity is actually maintained for fully effective threat constellations that are converging to some limit threat constellation which is also fully effective.

Independence of [Par]. It is well known in the context of pure bargaining problems that a solution that assigns the zero vector to each problem satisfies all of Nash's axioms other than [Par]. In our setting, this may be precluded by the requirement that a solution must lie within the unblocked set, which could therefore rule out the disagreement point. So consider instead the solution

$$\phi(G) = \arg\min_{x \in U(G)} \prod_{j \in N} x_j,$$

which clearly satisfies all our axioms except for [Par].

Independence of [Inv]. In standard bargaining problems, one example of a solution that satisfies all of Nash's axioms except for [Inv] is the Pareto-optimal allocation that assigns equal utilities to all players. Of course, this doesn't apply if the unblocked set is nonconvex or not comprehensive, but the the following modification will suffice. Consider a solution

$$\phi(G) = \arg\max_{x \in U(G)} \{\min_{i \in N} x_i\},$$

that seeks to maximize the utility of the individual with the lowest utility. This violates [Inv] but clearly satisfies [Par], [Sym], [IIA] and [UHC]. Additionally, [Exp] is met under any "uniform expansion" λ with $\lambda_i = \lambda_j > 1$ for all $i, j \in N$.

Independence of [IIA]. Consider the following solution, inspired by the solution of Kalai and Smorodinsky (1975). For any compact F, let

$$b_i(F) = \max\{x_i | (x_i, x_{-i}) \in F \text{ for some } x_{-i}\},$$

and define a solution

(a.1)
$$\phi(G) = \arg\max_{x \in U(G)} \left\{ \min_{i \in N} \frac{x_i}{b_i(F)} \right\}.$$

This satisfies all our axioms except for [IIA]. In particular, it satisfies [Exp].[1]

Observe that (a.1) allows for the possibility that points in F that are *not* in the unblocked set can affect the solution. It is possible that G and G' are such that $U(G) = U(G')$ but $\phi(G) \neq \phi(G')$ because $F \neq F'$. In other words, the fact that the coalitional Nash bargaining solution turns out to depend only on the unblocked set cannot be taken for granted simply by virtue of [Exp], if our other axioms are not satisfied.[2]

Moreover, the fact that a version of the Kalai-Smordinsky solution in our context satisfies [Exp] lends additional credence to the view that [Exp] is a reasonable axiom, and not one that is only applicable to the Nash solution or some modification thereof.

It is also worth noting that all the solutions above satisfy [Sym], showing the independence of these axioms extends to Corollary 1 as well.

2. An Extension of Theorem 2

We first provide an example that shows that the assumed convexity of $\{F(S)\}$ cannot be dropped free of charge from the statement of Theorem 2.[3]

Example 1. Let $N = \{1234\}$ with $d_i = \zeta_i = 0$ for all i, and:

(i) $F(N) = \{x \in \mathbb{R}^4_+ | x_1 + x_2 + x_3 + x_4 \leq v\}$.

(ii) $F(\{123\})$ has just two payoff allocations $(1, 1, 1)$ and (a, a, b), where $(a, b) \gg 0$.

(iii) $F(\{12\})$ consists of the single payoff allocation (c, c), with $c > 0$.

(iv) $F(S) = \{0_S\}$ for every other coalition S.

(v) $a^2 b < 1 < c < a < v$, and $b > \frac{v-c}{3} > 0$.[4]

[1] Interestingly, the required expansion is given by $\lambda = (\alpha, \alpha)$ for some $\alpha > 1$. Other "non-homogeneous" expansions do not always work.

[2] For example, suppose G is such that $F = \{x, x', y\}$, where $x = (2, 1)$, $x' = (1, 2)$, $y = (10, 0)$ and y is the only one that is blocked, e.g., because $\zeta_2 > 0$. This means that $U(G) = \{x, x'\}$. Consider G' with $F' = \{x, x'\}$ and inessential subcoalitions, so $U(G') = \{x, x'\} = U(G)$. Clearly, $\phi(G) = \{x\}$ while $\phi(G') = \{x, x'\}$.

[3] We could have used a three-player game to make the same point, but then ζ would need to be distinct from d. We prefer to not rely on the possibility that $\zeta \neq d$.

[4] For instance, $v = 3.05$, $a = 1.18$, $b = 0.66$ and $c = 1.1$, satisfy all the restrictions of the example, and additionally the game can be seen to be superadditive under these specifications.

The "naive" coalitional solution for N presumes that coalitions will block with their unconstrained Nash allocations. For coalition $\{12\}$ this is just (c, c), and for coalition $\{123\}$ it is $(1, 1, 1)$, using the assumption that $a^2b < 1$. So, keeping in mind that $c > 1$, the unblocked set $U'(N)$ for N is the set of all allocations in $F(N)$ that are unblocked by the threat (c, c) from coalition $\{12\}$. The naive solution for N is then easily seen to consist of the two allocations $(c, \frac{v-c}{3}, \frac{v-c}{3}, \frac{v-c}{3})$ and $(\frac{v-c}{3}, c, \frac{v-c}{3}, \frac{v-c}{3})$.[5]

The coalitional Nash solution for coalition $\{12\}$ is also (c, c). The coalitional Nash solution for coalition $\{1, 2, 3\}$ is then (a, a, b), given that the allocation $(1, 1, 1)$ is blocked by $\{12\}$ with (c, c), and also because $a > c$ by (v). Returning now to the grand coalition, we see that $U^*(N)$ must consist of allocations unblocked by (a, a, b), and this set is clearly nonempty because $b < 1$. The naive solution $(c, \frac{v-c}{3}, \frac{v-c}{3}, \frac{v-c}{3})$ is blocked by the recursively defined threat from $\{123\}$; after all, $a > c$ and $b > \frac{v-c}{3}$ by assumption. *The coalitional Nash solution for N is therefore different from the naive solution.* For instance, if $v = 3.05$, $a = 1.18$, $b = 0.66$ and $c = 1.1$, it can be checked that $\sigma^*(N)$ consists of the two allocations $(1.1, 0.645, 0.66, 0.645)$ and $(0.645, 1.1, 0.66, 0.645)$. The naive solution consists of $(1.1, 0.65, 0.65, 0.65)$ and $(0.65, 1.1, 0.65, 0.65)$. $\diamond$

At the same time, in the main text we pointed out that it is possible to weaken the assumption that all feasible sets are convex. Say that a set $A \subseteq \mathbb{R}^m$ is *subconvex* if for every $x, y \in A$ and $t \in [0, 1]$, there is $z \in A$ such that $z \geq tx + (1 - t)y$, and a set $A \subseteq \mathbb{R}^m_+$ is *log subconvex* if $\ln A$ is subconvex.[6] Log subconvexity is a weak property that could apply to connected sets as well as to sets with isolated elements. (Of course, in Example 1, $F(\{1, 2, 3\})$ is not log subconvex.) It can be also be verified that a convex set in $\mathbb{R}^m_+$ must be log subconvex.[7]

Theorem 2 holds in the more general case where, for every coalition T, $F(T)$ is assumed to be log subconvex (rather than convex). In proving the Theorem, the convexity of $F(T)$ was used only in Claim 2, which can be generalized as follows.

Claim 2′. Assume that $F(T)$ is log subconvex for every coalition T. Suppose there exists $x \in \Theta(T)$ such that $x \notin \Psi(T)$. Then there exists $W \subset T$ and $y \in \Psi(W)$ such that $y \geq x_W$.

[5]Because $a > 1$ and $a^2b < 1$, we have $b < 1$. Therefore $c > 1 > b > \frac{v-c}{3}$, and this assures us that the naive solution is indeed as stated in the main text.

[6]Note that log subconvexity is only defined for subsets A of $\mathbb{R}^m_+$, so that $\ln x$ is well-defined in the extended reals for all $x \in A$. The vector ordering "$\geq$" is then applied to the extended reals in the obvious way.

[7]Let A be convex. Then for every $x, y \in A$ and $t \in (0, 1)$, $z \equiv tx + (1 - t)y \in A$. But we know that $\ln(z) \geq t \ln x + (1 - t) \ln y$, which proves that A is log subconvex.

Proof. We proceed by induction on coalition size. Clearly the assertion is trivially true for all coalitions T of size 2 or less: $\Theta(T) = \Psi(T)$. Now suppose that for some $k \geq 2$, the lemma is true for any coalition of size k or less. Let T have cardinality $k+1$. Suppose that $x \in \Theta(T)$ and $x \notin \Psi(T)$. Then the Nash product over $F(T)$ exceeds that under x, so

$$\sum_{i \in T} \ln \tilde{x}_i > \sum_{i \in T} \ln x_i. \tag{a.2}$$

for any $\tilde{x} \in \Psi(T)$. For any $t \in (0,1)$, consider the allocation $t \ln x + (1-t) \ln \tilde{x}$. By the log-subconvexity of $F(T)$, there is a payoff allocation $z(t) \in F(T)$ such that $\ln z(t) \geq t \ln x + (1-t) \ln \tilde{x}$. Combining this information with (a.2), we must conclude that

$$\sum_{i \in T} \ln z_i(t) \geq \sum_{i \in T} [t \ln x_i + (1-t) \ln \tilde{x}] > \sum_{i \in T} \ln x_i;$$

i.e., the Nash product under $z(t)$ exceeds that under x. So $z(t)$ is blocked by some subcoalition of T using some coalitional Nash solution for that subset. Notice that for any limit point z of $z(t)$ as $t \to 1$, we have $z \geq x$. So there exist $H \subset T$ and $y \in \Theta(H)$ such that

$$y \geq z_H \geq x_H. \tag{a.3}$$

If $y \in \Psi(H)$ we are done. If $y \notin \Psi(H)$, then because $|H| \leq k$, the induction hypothesis implies there is $W \subset H$ and $z \in \Psi(W)$ such that $z \geq y_W$. Moreover, W can be chosen so that $\Theta(W) = \Psi(W)$. Combining this with (a.3), we must conclude that $\psi(W) \geq x_W$. ∎

3. Details for the Examples in Section 4

3.1. **Example 3.** We take care of the following details.

1. *For any $r > 0$, there is $\bar{\alpha} > 1$ such that if $\alpha > \bar{\alpha}$, then the payoff $\Sigma(jk, \pi^{jk})$ exceeds 2/3. Indeed, as synergy levels become large ($\alpha \to \infty$), the payoff $\Sigma(jk, \pi^{jk})$ converges to 1, which is the maximum possible surplus.*

Proof. We first claim that as $\alpha \to \infty$, $\lambda_{jk} \to \infty$. Suppose not; then along some subsequence of $\alpha \to \infty$, λ_{jk} is bounded. Observe that λ_ℓ is bounded anyway as the cost of effort cannot exceed 1 and there are no synergies that apply to the singleton ℓ. Combining these two statements, it follows that the right hand side of the first equation in (19) converges to 0. But that is a contradiction, because $r > 0$ and $\ell \geq 0$.

With this claim in hand, we now claim that as $\alpha \to \infty$,

$$\frac{\lambda_{jk}}{r+\lambda_{jk}+\lambda_\ell} \to 1 \text{ and } \frac{1}{2\cdot 2^\alpha}\lambda_{jk}^2 \to 0. \tag{a.4}$$

The first of these claims immediately follows from $\lambda_{jk} \to \infty$ and λ_ℓ bounded. To prove the second claim, use the first equation in (19) again to see that

$$\frac{1}{2\cdot 2^\alpha}\lambda_{jk}^2 = \frac{\lambda_{jk}(r+\lambda_\ell)}{2[r+\lambda_\ell+\lambda_{jk}]^2} \to 0,$$

because the numerator is linear in λ_{jk} and the denominator is quadratic. Now use (a.4) in (20) to complete the proof. ■

2. *Computation of threshold α for low discounting.*

If the degree of impatience is small ($r \simeq 0$), it is easy to get a sense of the threshold $\bar{\alpha}$. Combining the two equations in (19) by putting $r \simeq 0$ and manipulating, we see that $\lambda_{jk} \simeq 2^{\alpha/2}\lambda_\ell$. The first equation in (19) then tells us that

$$\frac{1}{2^\alpha}\lambda_{jk}^2 \simeq \frac{\lambda_\ell\lambda_{jk}}{(r+\lambda_{jk}+\lambda_\ell)^2} \simeq \frac{2^{\alpha/2}}{(2^{\alpha/2}+1)^2}.$$

Using this information in equation (20) for the doubleton payoff, we see that

$$\Sigma(jk, \pi^{jk}) \simeq \frac{2^{\alpha/2}}{2^{\alpha/2}+1} - \frac{2^{\alpha/2}}{2(2^{\alpha/2}+1)^2},$$

which exceeds $2/3$ for all $\alpha \geq 3.3$.

3. *There is a unique equilibrium (which is symmetric) under the singleton coalition structure.*

Proof. Noting that $t = 1$ in (16), we have:

$$\frac{1}{(r+\Lambda)} - \frac{\lambda_i}{(r+\Lambda)^2} - \lambda_i = 0, \tag{a.5}$$

Given Λ, it is obvious that λ_i is pinned down uniquely by (a.5). It follows immediately that $\lambda_j = \lambda \equiv \Lambda/3$ for all j. Using this in (a.5) and moving terms around, we see that:

$$\lambda[(r+3\lambda)^2 - 2] = r, \tag{a.6}$$

which can be seen to yield a unique solution for λ, for any $r \geq 0$. That generates a symmetric equilibrium payoff to each firm under the singleton structure, which in turn generates threat points for the doubleton structure:

$$\Theta(i, \pi^{jk}) \equiv \frac{\lambda}{(r+3\lambda)} - (1/2)\lambda^2 \text{ for all } i, \text{ where } \lambda \text{ uniquely solves (a.6)}, \tag{a.7}$$

for each $i \in T$.

4. *Asymmetric bargaining weights and the continued viability of the grand coalition.*

Faced with (23), can the grand coalition assure its own viability? Without loss, suppose that 3 is a person who has the highest of these y-values, and that this occurs in the coalition $\{32\}$. Consider any allocation x for the grand coalition such that $x_1 = \max\{y_1^{13}, y_1^{12}\}$, $x_2 = y_2^{23}$, and x_3 is the remainder of the surplus. Note that x is feasible. That is because $x_1 + x_2 = \max\{y_1^{13}, y_1^{12}\} + y_2^{23} \leq y_3^{23} + y_2^{23} =$ doubleton payoff in any intermediate structure, and so by the efficiency of the grand coalition, $x_3 \geq \Sigma(3, \pi^{12}) \geq 0$. The allocation x is also unblocked: player 1 will not want to participate in a potential block by $\{12\}$ or $\{13\}$, player 2 will not want to participate in a potential block by $\{23\}$, and no player will want to stand alone, knowing that the other two will band together.

3.2. **Example 4.** We wish to show that condition (a.9), reproduced here as

$$A + C > (n+1)[\max_j c_j]. \tag{a.8}$$

also guarantees positive production by every coalition in coalitional equilibrium in any coalitional structure. Given linear costs and costless transfers, all production within a coalition will be carried by its lowest-cost members — the other members are effectively paid to stay out of the way. Therefore, given any structure π, if we define $c(T) \equiv \min_{i \in T} c_i$ for any $T \in \pi$ and $C(\pi) \equiv \sum_{T \in \pi} c(T)$, the necessary and sufficient condition for positive production by every coalition is given by:

$$A + C(\pi) > (|\pi| + 1)[\max_{S \in \pi} c(S)], \tag{a.9}$$

where $|\pi|$ is the cardinality of π. Arrange firms so that $c_{\leq} c_2 \leq \cdots \leq c_n$, and define $C_{|\pi|} \equiv \sum_{j=1}^{|\pi|} c_j$. Then we must conclude that

$$A + C(\pi) \geq A + C_{|\pi|} > (|\pi| + 1)c_n + (n - |\pi|) \sum_{j=|\pi|+1}^{n} [c_n - c_j] \geq (|\pi| + 1)c_n \geq (|\pi| + 1)[\max_{S \in \pi} c(S)],$$

where the first inequality follows from $C(\pi) \geq C_{|\pi|}$, the second from (a.8) and the fact that $C = C_{|\pi|} + \sum_{j=|\pi|+1}^{n} c_j$, and the third and fourth from $c_n \geq c_j$ for all j.

4. Some Remarks on Viable Coalition Structures

Recall the definition of Nash-in-Nash equilibrium in the main text: Given a threat constellation $\boldsymbol{\Theta}$ as described in the text, a is a *Nash-in-Nash equilibrium* for π relative to $\boldsymbol{\Theta}(\pi)$ if for every $S \in \pi$,

$$a_S \in \arg\max_{a'_S \in A(S)} \prod_{i \in S} f_i(a'_S, a_{-S})^{\gamma_i(S)}, \text{ s.t. } f_S(a'_S, a_{-S}) \text{ unblocked by any } T \subset S; \tag{a.10}$$

that is, a_S is a coalitional Nash solution relative to $\boldsymbol{\Theta}(\pi)$ and weights $\gamma(S) \gg 0$. Structure π is *viable* relative to $\boldsymbol{\Theta}(\pi)$ if it admits a Nash-in-Nash equilibrium relative to $\boldsymbol{\Theta}(\pi)$. Subsequently, $\boldsymbol{\Theta}$ is defined recursively via the procedure in Section 3.2, leading to equations (8) and (9) in the main text.

4.1. Interpretation.

Notice that in the interaction across coalitions in a coalition structure, coalitions choose their actions in ignorance of the actions of other coalitions. This is just the standard modern interpretation of Nash equilibrium, extended to coalitions. However, it is also the case in our definition that when a coalition chooses a best response in (**??**), it is itself constrained, in that its chosen response should be immune to blocking by subcoalitions. That implicitly means that *once* some subcoalition does break away from its parent coalition in the structure, that fact becomes common knowledge and coalitions re-optimize against the new coalition structure that results. Indeed, our recursive definition demands this interpretation, because $\Theta(T, \pi)$ is defined by coalitional payoffs that result once a *new* structure is formed after the departure of T; see (8) and (9).

These informational underpinnings serve as a shorthand for our fundamental emphasis. Our solution concept is *not* a refinement of Nash equilibrium, as it would be if these coalitional deviations were also to occur simultaneously, as in Bernheim, Peleg, and Whinston (1987). Rather, in line with the philosophy of our earlier work (Ray and Vohra 1997, 1999), we take the viewpoint that when binding agreements can be written within coalitions, a coalitional realignment must be "slow" enough relative to the flexibility of actions, so that the latter can be adjusted in response to the former once the former becomes known, and in our exercise we do presume that the existing coalition structure becomes commonly known.

4.2. On External Threats.

In the main paper we develop the Nash-in-Nash theory by assuming that every coalition S is only threatened by breakaways that are subsets of S. In these Notes, we consider an extension of that definition to the general case in which any coalition T can form, with S potentially threatened whenever $S \cap T$ is nonempty.

As in the main text, a coalition interacts non-cooperatively with other coalitions in a structure π but attempts to find a cooperative agreement within itself. As it does so, it contends with threats from potential *breakaways*, now defined not as subsets but as a new coalition T with $S \cap T = \emptyset$. Another way of putting it is that S has to guard against subsets T_1 who might link up with another subset T_2 (disjoint from S) to form the coalition $T \equiv T_1 \cup T_2$.

Fix some coalition structure π, and suppose provisionally that every coalition T not in π (but not necessarily just subsets of other coalitions in π) has a set of threat payoffs $\Theta(T, \pi)$, soon to be endogenized. (The reason π still indexes the threats is that T would have to work from the baseline π, and so its available threats could well be conditional on it.) Let $\boldsymbol{\Theta}(\pi) \equiv \{\Theta(T, \pi)\}$. In the main text we then defined a to be a Nash-in-Nash equilibrium for π (relative to $\boldsymbol{\Theta}(\pi)$) if for every $S \in \pi$, (a.10) is satisfied, *where T is a subset of S*. That brings us to:

Problem 1. *If we drop internal blocking, which sets T should S be concerned about now?*

Certainly, S should continue to be concerned about all $T \subset S$: whether or not they break away is entirely in the hands of S. But what if (for instance), $T \equiv T_1 \cup T_2$, where $T_1 \subseteq S$ and $T_2 \subseteq S'$ for some other $S' \in \pi$? S should only be cognizant of those threats $T \equiv T_1 \cup T_2$ for which T_2 has not already been appeased by S' (and vice versa). This creates an entanglement in the feasible payoff sets for S and S' in equation (a.10), which turns their non-cooperative interaction into something that is broader than a game; cf. Debreu (1952).

This is by no means an insuperable problem. But it would need to be developed and exposited. We could then define a Nash-in-Nash equilibrium relative to $\boldsymbol{\Theta}$.

Our next task is to close the circle so that $\boldsymbol{\Theta}$ is consistently defined, and along with $\boldsymbol{\Theta}$, the collection of viable coalition structures is fully identified. With internal blocking, we began by declaring the singleton structure π^s to be viable. (A Nash-in-Nash equilibrium is then just a Nash equilibrium; assume it exists.) We then proceeded recursively, going up the hierarchy to coarser and coarser coalition structures. Without internal blocking, that particular recursive avenue is closed off, which brings us to:

Problem 2. *When blocking is not internal, is a recursive definition possible? How should we proceed otherwise?*

We see two possible resolutions to Problem 2.

Option 1. Drop recursion altogether, and formulate an "equilibrium collection" of threats $\mathbf{\Theta}$ as a fixed point of a suitably defined mapping. To this end, start with $\mathbf{\Theta}$, which is a collection of coalition structures π and $\theta(T,\pi)$ for every coalition *not in* π. Say that a coalition structure π is viable relative to $\mathbf{\Theta}$ if it admits a Nash-in-Nash solution relative to $\mathbf{\Theta}$, exactly as in the main text or in (a.10). Collect all such solutions for each viable π, call them $\Sigma(\pi)$. If a coalition structure is not viable relative to $\mathbf{\Theta}$, then define $\Sigma(\pi) = \emptyset$.

To complete the fixed point, now consider any π and T not in π. Define $\pi^T \equiv \pi_{-S} \cup \{T, S_1 - T, \ldots, S_m - T\}$, where $S_1, \ldots S_m$ are all the coalitions that contain some member(s) of T. If π^T is viable relative to $\mathbf{\Theta}$, define:

$$\Theta'(T,\pi) \equiv \{f_T(a) | a \text{ is Nash-in-Nash for } \pi^T \text{ relative to } \mathbf{\Theta}\}.$$

If π^T is not viable, define $\mathcal{P}^T$ to be the collection of all the "closest" viable coalition structures π' to π^T, provided one exists.[8] Define

$$\Theta'(T,\pi) \equiv \{f_T(a) | a \text{ is Nash-in-Nash for some } \pi' \in \mathcal{P}^T \text{ relative to } \mathbf{\Theta}\}.$$

In words: when π^T is viable, then the threat raised by T is precisely is Nash-in-Nash payoffs once it precipitates π^T by its move. If π^T is not viable, our construction uses the (correct) prediction by T that the structure will evolve further. Just where it will evolve *to* is unclear. Our construction asserts that it will stop at one of the "closest" viable structures, where "closest" is defined in the sense of Footnote 8.

This defines $\mathbf{\Theta}'(\pi) \equiv \{\Theta'(T,\pi)\}$ for every π, *relative to the starting constellation* $\mathbf{\Theta}$. A fixed point $\mathbf{\Theta} = \mathbf{\Theta}'$ can be then taken to be the solution for the system as a whole.

Option 2. Allow any coalition (or more generally any coalition *structure*) to form an exogenously bounded number of times. Then a recursive definition is possible by tagging every structure that comes into being with a full history of which structures have formed in the past. It is evident that starting from some structures and some tagged histories, *no further coalition can form*. These are the terminal nodes for the recursion, and working back from these, it is possible to develop a theory of viability just as in the main text.

We do not pursue either of these options in the main paper. In our view, they detract from the main contribution, which has to do with the within-coalition solution concept that we put in place when studying interaction across coalitions.

[8] That is, π' must be viable, and there is a finite sequence $T_1, \ldots T_k$ of coalitions after T has already moved the system to π^T, such that $\pi' = (\cdots(\pi^T)^{T_1})\cdots)^{T_k}$, and *no* intermediate coalition structure along this path is viable.

References

BERNHEIM, B., B. PELEG, AND M. D. WHINSTON (1987): "Coalition-Proof Nash Equilibria I. Concepts," *Journal of Economic Theory*, 42, 1–12.

DEBREU, G. (1952): "A Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Sciences of the United States of America*, 38, 886–893.

KALAI, E. AND M. SMORODINSKY (1975): "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43, 513–518.

RAY, D. AND R. VOHRA (1997): "Equilibrium Binding Agreements,," *Journal of Economic Theory*, 73, 30–78.

——— (1999): "A Theory of Endogenous Coalition Structures," *Games and Economic Behavior*, 26, 286–336.