



**UNIVERSITY
OF WARWICK**

Economics

Nash Bargaining in Coalitional Games

Rajiv Vohra and Debraj Ray

October 2026

No: 1633

**Warwick
Economics
Research Papers**

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Nash Bargaining in Coalitional Games

Rajiv Vohra (✉) Debraj Ray[†]

August 2026

Abstract. We revisit Nash’s axiomatic bargaining solution when coalitions employ threats that must be consistent with *their* solutions. As in Nash, our solution maximizes a (possibly weighted) product of payoffs for each coalition, but subcoalitional threats appear as conventional constraints that are not netted out when maximizing the Nash product. We study different aspects of this coalitional solution, including its connections to a notion of “pragmatic egalitarianism”. We then embed the solution into a setting with externalities, and define viable coalitional structures, under which every coalition follows its coalitional solution but interacts noncooperatively with other coalitions. We discuss applications to public goods, R&D coalitions, oligopoly cartels, and hospital networks.

JEL classification: C70, D60, D63.

Keywords: Nash bargaining, coalitional threats, strategic externalities.

1. Introduction

Nash (1950) introduced and axiomatically derived his celebrated bargaining solution, which describes a procedure for choosing a payoff division from a feasible set of n -person payoff outcomes.¹ As is well known, the procedure involves maximizing the product of individual payoffs (net of individual disagreement points) over the feasible set.

The goal of this paper is to extend the Nash bargaining solution to situations in which *coalitions*, not just individuals, can walk away from the bargaining table. We do not have a widely applicable account of how a group of agents would divide a surplus, when subcoalitions of two or more members are in a position to thwart or “block” various agreements. Apart from providing that extension, we would like to place the resulting solution within

[†]Ray: New York University and University of Warwick; debraj.ray@nyu.edu. Vohra: Brown University; rajiv_vohra@brown.edu. We are grateful to Nemanja Antic, Piotr Dworczak, Sam Jindani, Marco Mariotti, David Pearce, Ariel Rubinstein, Mehmet Mars Seven, Roberto Serrano, and a Co-Editor and three anonymous referees for insightful comments. Claude (Fable 5 and Opus 5) provided invaluable help at several stages of this research. Ray acknowledges support under National Science Foundation Grant SES-2315720.

¹Nash established his solution for $n = 2$, but it is easily seen to extend to an arbitrary number of individuals. See Thomson (1994, 2022) for surveys of the extensive literature that followed Nash (1950).

different strategic settings and see what it yields. Currently, an entire applied literature — on which more below — is restricted to the study of agreements under the shadow of individual disagreements alone.

Specifically, our objectives are:

- (a) To characterize an analogue of Nash’s solution in a setting with coalitional threats, describing simultaneously how coalitional threats are derived and how they are dealt with;
- (b) To explore the resulting coalitional solution from different perspectives, including those of efficiency and fairness; and
- (c) To embed the solution into an environment with strategic externalities across coalitions.

The remainder of this Introduction expands on what we do under each of these headings.

We begin with characteristic function games $\mathcal{G} = (\mathcal{F}, d)$, where $\mathcal{F} = \{F(S)\}$ collects sets of feasible payoff vectors for every coalition S , including the grand coalition N of all players, and $\{d_i\}_{i \in N}$ is a vector of individual disagreement payoffs, to be interpreted as payoffs to complete inaction by any player i .² A solution σ maps every $\mathcal{G}$ to a collection of sets $\sigma(S, \mathcal{G}) \subseteq F(S)$, one for each coalition S . We ask that no element of any solution for S should be *blocked* by any subcoalition $T \subset S$, using some element from *its* solution $\sigma(T, \mathcal{G})$. That is, *threats* to a larger group are intimately connected to subgroup *solutions*. We axiomatize a solution of the form

$$(1) \quad \sigma(S, \mathcal{G}) = \arg \max \prod_{j \in S} [x_j - d_j]^{\gamma_j(S)}$$

for some collection $\gamma(S) = (\gamma_i(S))_{i \in S}$ of nonnegative weights, not all zero, where the argmax is taken over the set of all unblocked allocations in $F(S)$. See Theorem 1 for a precise statement, and Corollary 1 for the reduction to equal weights under a symmetry axiom. We refer to σ as the (*consistent*) *coalitional solution*.

Apart from the dual perspective that one coalition’s solution is another coalition’s threat, our characterization makes a central point about the handling of threats. Recall that “split-the-surplus” (net of disagreement payoffs) is a fundamental property of Nash’s solution. In related theory and in applications, split-the-surplus has assumed a distinctly moral significance — bargaining “should” set aside what we can get for ourselves, and “should” focus

²The reader is free to view $F(\{i\})$ as distinct from or equal to the singleton $\{d_i\}$, depending on the context.

on dividing what remains.³ But this property does not extend to coalitional threats. All such threats, including those from singleton coalitions, *bind in the conventional way*. They affect the solution only when the solution lies on the boundary of the unblocked set.

Similar outcomes emerge from noncooperative bargaining theory; see Chatterjee et al. (1993), Ray and Vohra (1999), and Compte and Jehiel (2010). Binmore, Shaked, and Sutton (1989) are to our knowledge the first to explicitly take note of this difference (though they make the point for individuals and not larger coalitions). They ask of outside options:

“When is such a threat credible? Only when dealing himself out gives the bargainer a bigger payoff than dealing himself in. It follows that the agreement that would be reached without outside options is immune to deal-me-out threats, unless the deal assigns one of the bargainers less than he can get elsewhere.”

Our Theorem 1 obtains the same finding from a fundamentally different perspective. It arises from axiomatic desiderata, and does not rely on specific game forms with particular bargaining protocols for offers and responses. In this sense, we are entirely in line with the well-known Nash program, where cooperative and noncooperative developments of the same concepts go hand in hand. For more discussion, see Section 2.4.

We should also note that these noncooperative solutions lie in the core of the underlying characteristic function. Our solution does not, because potential breakaway coalitions are themselves constrained by the use of the solution in their blocking. That expands the set of unblocked allocations relative to the core, and creates an entirely new, recursively defined set over which the maximization exercise occurs. The exact nature of this unblocked set appears complicated because of the recursive nature of the definition. But Theorem 2 shows that if feasible payoffs $F(S)$ for each coalition S form a convex set, our recursively defined coalitional solution has an easy non-recursive characterization. It is generated by a one-shot maximization of (1) over all allocations unblocked using the *unconstrained* Nash outcome for each coalition. That is, this artificial but easy-to-solve problem delivers the “correct” answer for the recursively defined solution.

Theorem 2 has strong consequences in transferable-utility settings. With side payments available, we are in the environment of the Coase Theorem: agents can freely engage in the writing of binding agreements with transfers, or just as freely abandon such negotiations.

³See, e.g., the analysis of a creditor problem from the Talmud in Aumann and Maschler (1985), or foundations of the pre-kernel in Davis and Maschler (1965).

The Coase Theorem asserts that the resulting agreements will be efficient. We pose this question here as follows: suppose that the union of any two disjoint coalitions has at least as much value as the sum of their values; that is, the game is superadditive. Is it the case that in all such games $\mathcal{G}$, $\sigma(N, \mathcal{G})$ is nonempty and Pareto optimal?

If a solution is required to be chosen from the core, then Bondareva (1963), Shapley (1967) and Scarf (1967) answer this question emphatically in the negative. Their work shows that superadditive games generally have empty cores, and that demanding conditions are required to restore nonemptiness. These results are often viewed as a rebuttal to the Coase Theorem. In contrast, Theorem 2 implies that in all superadditive transferable utility (TU) games satisfying a mild consistency condition on bargaining weights, $\sigma(N, \mathcal{G})$ is nonempty and Pareto optimal (Theorem 3). This finding deeply relies on the particular structure of the coalitional solution, restricting credible threats to a degree that permits efficiency with no further assumptions.

Theorem 2 can also be employed to examine a philosophical interpretation of our coalitional solution. We've already noted that coalitional threats are dealt with conventionally, rather than using some version of "split-the-surplus". That suggests a different "pragmatic ethic": move as close as possible to equality in payoffs, unless prevented from doing so by a potential breakaway threat. This allows us to explore a close connection between our bargaining solution and the egalitarian solution (Dutta and Ray 1989, 1991). The egalitarian solution combines a commitment to equity with the pragmatics of coalitional participation, in the sense of picking out Lorenz-maximal elements from the unblocked set.

Theorem 4, based yet again on the non-recursive characterization of Theorem 2, asserts that whenever every individual has positive standalone worth, our coalitional solution with symmetric bargaining weights is contained within the egalitarian solution.

Section 4 goes beyond characteristic functions. It embeds our solution into settings with cross-coalitional externalities that arise naturally from games in strategic form. Under some conditions, we can assign, for every partition π of the player set, a set of payoffs $F(S, \pi)$ to every coalition $S \in \pi$. That is, F is no longer defined "in isolation" but varies with the ambient environment (the partition, or coalition structure) in which S is located.

Sections 4.1–4.3 develop a definition and a recursive method — an extension of the recursive construction for characteristic functions — that allows us to describe "viable coalitional structures" in any particular application. The coalitional solution for characteristic function games now becomes one component of a larger recursive solution. Under any

coalition structure, every coalition internally follows its coalitional solution but interacts noncooperatively with other coalitions. The viability of the entire structure means that each such coalition has a nonempty coalition solution, and this naturally generates a theory of coalition formation.

Because our coalitional solution involves a variant of Nash bargaining, while coalitions interact in a noncooperative way, the notion of viability is also linked to the Nash-in-Nash literature; see, e.g., Horn and Wolinsky (1988), Collard-Wexler, Gowrisankaran, and Lee (2019), or Bagwell, Staiger, and Yurukoglu (2020). The literature typically allows only for bilateral Nash bargaining with individual outside options. In contrast, our theory accommodates multilateral outside options via coalitional threats. That said, we do restrict ourselves only to subcoalitional threats, which permits us to build the edifice of solutions recursively from the finest structure upwards.

There is no known equivalent of Theorem 2 for environments with externalities, so the analysis of specific cases can be nuanced and complex. In Section 4.4, we provide four applications: the provision of public goods, the formation of R&D coalitions, collusive behavior in Cournot competition, and the formation of hospital networks, and we study viable coalition structures in these applications. While the specifics vary, we attempt to illustrate two general themes.

First, because the feasible payoffs to a breakaway coalition depend on the coalition structure that emerges from its own defection, that structure needs to be predicted. For instance, suppose that four firms are deciding whether or not to form a cartel, and a “subcartel” of two firms threatens to abandon the negotiations. Their options are mediated by their interactions with the other two firms. Those interactions in turn will depend in part on whether those two other firms will remain together, or split up.

Second, these interactions could generate inefficient outcomes, unlike Theorem 3, which applies only to characteristic functions.⁴ And yet, it is important to appreciate that the standard intuition for inefficiency under noncooperative interaction does not apply. For instance, consider the Prisoner’s Dilemma. Then the grand coalition is indeed viable, and the outcome is efficient. Indeed, the greater the inefficiency of noncooperative interaction, the stronger is the impetus to write agreements that restore efficiency. So in this framework,

⁴Such inefficiencies are consistent with the findings of an earlier literature. For instance, Bloch (1996) and Ray and Vohra (1997, 1999, 2001) show that cooperation does not necessarily imply first-best efficiency in the presence of externalities. Maskin (2003) shows how this feature also extends to a coalitional solution concept that is based on axioms related to the Shapley value; see also de Clippel and Serrano (2008a,b).

inefficiency can happen but only when there isn't too much of it. The viability of an intermediate coalition structure with moderate inefficiencies is a double-edged sword: it avoids large degrees of inefficiency, but also threatens the viability of an efficient outcome.

We end the Introduction with two remarks. Theorem 1 relies on a new “expansion axiom” that we append to the otherwise standard arsenal of the Nash axioms. An axiomatic characterization is appealing if on the one hand, the axioms appear to be reasonable or intuitive, while at the same time their mathematically equivalent implication is “surprising” in that it delivers new intuition. The reader must judge whether the expansion axiom is satisfactory in this sense, while retaining the simplicity of Nash's original characterization.

Second, as a historical footnote, von Neumann was appreciative of Nash's foray into cooperative bargaining; see the acknowledgements in Nash (1950). The same reaction was not forthcoming for Nash's noncooperative equilibrium.⁵ As Bhattacharya (2022) notes in his biography of von Neumann, “The idea that people might not work together for mutual benefit was anathema to him . . . [and] ran counter to the spirit of his ‘coalitional’ conception of game theory.” So while “Nash equilibrium” did not appear to cut it for von Neumann, it seems likely that he would have welcomed an attempt to connect the Nash solution to his “coalitional conception of game theory”. Our paper is one such attempt.

2. Nash Bargaining with Coalitional Threats

We begin our analysis with characteristic functions. There is a finite set of players $N = \{1, \dots, n\}$. A coalition S is a nonempty subset of N , and a subcoalition T (of S) is a nonempty strict subset of S . For $A \subseteq N$, $\mathbb{R}^A$ is $|A|$ -dimensional Euclidean space with coordinates indexed by the elements of A . For each coalition S there is a set $F(S) \subset \mathbb{R}^S$, interpretable as the collection of feasible payoffs S can implement on its own. Finally, $d = (d_1, \dots, d_n)$ denotes individual disagreement payoffs resulting from complete inaction.

For $x, y \in \mathbb{R}^A$, say $x \geq y$ if $x_i \geq y_i$ for all $i \in A$, $x > y$ if $x \geq y$ and $x \neq y$, and $x \gg y$ if $x_i > y_i$ for all $i \in A$. For $x \in \mathbb{R}^A$ and $B \subset A$, x_B is the projection of x on B .

Consider the domain of all games $\mathcal{G} = (\mathcal{F}, d)$ on all finite N , where $\mathcal{F} = \{F(S)\}$ and:

⁵But a general sense of ambivalence is evident in von Neumann's letter to David Gale, dated November 5, 1949: “The idea of discussing n -person games ‘without cooperation’ is an attractive one that had occurred to me too, but I gave it up in favor of the coalition-type procedure because I didn't see how to exploit it.” (von Neumann 1949). We thank Dirk Bergemann and Philipp Strack for bringing this letter to our attention.

[Dom] For every coalition S , $F(S)$ is nonempty, compact, with $x \geq d_S$ for every $x \in F(S)$.

In particular $\zeta_i := \max F(\{i\}) \geq d_i$ for every i . There is no problem in reading the paper either by presuming that ζ_i always equals d_i (individual worth equals inaction payoff), or by allowing for $\zeta_i > d_i$ (individual worth involves decisions that dominate inaction).

Coalition S is *ineffective* if $F(S) = \{d_S\}$, and is *effective* otherwise. Any game with S ineffective for all $S \subset N$ represents a pure bargaining problem. The possibility that some coalition is effective represents our point of departure.

2.1. Solutions. A *solution* σ assigns to every coalition S in every game $\mathcal{G}$ a subset $\sigma(S, \mathcal{G})$ of $F(S)$. We maintain two conditions throughout. First, the solution is “pragmatic” in that it takes into account threats from potential breakaway subcoalitions, which come in turn from the solution applied to those breakaways. That is, for any coalition S , $\sigma(S, \mathcal{G})$ must be a subset of the *unblocked set* $U^\sigma(S, \mathcal{G})$ for S , or

$$(2) \quad \sigma(S, \mathcal{G}) \subseteq U^\sigma(S, \mathcal{G}) := \{x \in F(S) \mid \text{there is no } T \subset S \text{ and } y \in \sigma(T, \mathcal{G}) \text{ such that } y \gg x_T\},$$

where we additionally require that $\sigma(S, \mathcal{G})$ must be nonempty whenever $U^\sigma(S, \mathcal{G})$ is. Second, the solution is the same for S viewed as a subcoalition of N or as a grand coalition in its own right, provided all feasible sets for S and its subcoalitions are the same. Formally, for each $\mathcal{G}$ and $S \subseteq N$, define the “reduced game” $\mathcal{G}_S$ by $\{F(T)\}_{T \subseteq S}$ and d_S . Then

$$(3) \quad \sigma(S, \mathcal{G}) = \sigma(S, \mathcal{G}') \text{ whenever } \mathcal{G}_S = \mathcal{G}'_S.$$

Remark 1. The requirement that a threat from T must belong to the solution for T is a consistency restriction. Any other potential “threat” is considered not to be credible. Conditions (2) and (3) make this point explicit. Condition (2) defines an unblocked set within which S seeks its solution: that set is delineated by threats from subcoalitions that are restricted to follow their own solutions. Condition (3) clarifies that these latter solutions come precisely from applying σ to any potential breakaway coalition.

Remark 2. We only consider threats from subcoalitions.⁶ Our restriction removes the inherent circularity that arises from σ being present on both the left and right sides of (2), permitting the solution to be constructed recursively, from the singletons up. To allow for

⁶Our construction resembles equilibrium refinements such as coalition-proof Nash equilibrium (Bernheim, Peleg, and Whinston 1987). However, our solution concerns binding agreements and is *not* a refinement of any underlying noncooperative equilibrium. In this respect it resembles the equilibrium binding agreements (EBA) studied in Ray and Vohra (1997), except that our solution inserts a norm into every coalition that forms. EBAs have no within-coalition norm built in; the subset inclusion in (2) would hold with equality.

threats that include players from outside the parent coalition we would have to go beyond a recursive definition to resolve the circularity issue. As suggested in Ray and Vohra (2015a, 2015b), such a resolution can be found by adopting the notion of an abstract von Neumann-Morgenstern stable set. See the Supplementary Notes for a discussion of how that approach can be applied in the present context.

Remark 3. Coalitional threats lead naturally to nonconvex unblocked sets.⁷ That is why we must allow our solution to be multivalued.

Remark 4. While consistency is an important restriction, it is possible to consider unblocked sets that allow subcoalitions to use *any* feasible payoff as a threat:

$$(4) \quad U^{\text{feas}}(S, \mathcal{G}) := \{x \in F(S) \mid x \geq d_S, \text{ and there is no } T \subset S \text{ and } y \in F(T) \text{ such that } y \gg x_T\},$$

which leads to the standard definition of the core. In conjunction with Pareto optimality, (4) would imply that a solution for S is a subset of the core of the “reduced characteristic function” consisting of S and its subcoalitions. As we discuss in Section 2.4, that can yield a fundamentally different kind of coalitional solution. It isn’t what we do here.

2.2. Axioms. We begin with standard Nash axioms naturally adapted to $|N| \geq 2$ and to the possibility of multi-valued solutions. First, any solution must be Pareto optimal:

[PO] For any coalition S , whenever $\sigma(S, \mathcal{G})$ is nonempty, it is Pareto optimal: if $x \in \sigma(S, \mathcal{G})$, there does not exist $y \in F(S)$ such that $y \gg x$.

Next, any expected utility representation of preferences must be unique up to affine transformations. This induces an affine transformation of the entire game $\mathcal{G}$ in the obvious way.⁸ The *invariance axiom* asserts that:

[Inv] If $\mathcal{G}'$ is an affine transform of $\mathcal{G}$, then for every coalition S , $\sigma(S, \mathcal{G}')$ is the same affine transform of $\sigma(S, \mathcal{G})$.

Remark 5. Axiom [Inv] immediately permits us to normalize all disagreement points to zero, and indeed for some axioms to follow, we will. However, by explicitly carrying d for the time being, we can clarify an important aspect of our characterization below.

⁷Randomization does not help: $F(S, \mathcal{G})$ would be convex, but $U^\sigma(S, \mathcal{G})$ would, in general, remain nonconvex.

⁸ $\mathcal{G}' = (\mathcal{F}', d')$ is an affine transform of $\mathcal{G} = (\mathcal{F}, d)$ if $\mathcal{G}'$ is the result of converting every payoff x_i of every player i in $\mathcal{G}$ to y_i by some affine transform $y_i = \alpha_i + \beta_i x_i$, where $\alpha_i \in \mathbb{R}$ and $\beta_i > 0$.

Next, we apply Nash's *independence of irrelevant alternatives* to multi-valued solutions. We adapt Kaneko's (1980) version of this axiom written for pure bargaining problems. Our modification is minimal in that it considers shrinkages in the feasible set $F(S)$ for some coalition S but leaves all feasible sets $F(T)$ for subcoalitions $T \subset S$ unchanged.

[IIA] If $\mathcal{G} = (\mathcal{F}, d)$ and $\mathcal{G}' = (\mathcal{F}', d)$ have $F'(S) \subseteq F(S)$ for some S , but $F'(T) = F(T)$ for all $T \subset S$, then $\sigma(S, \mathcal{G}') = \sigma(S, \mathcal{G}) \cap F'(S)$ whenever this intersection is nonempty.

Remark 6. The intuition here is exactly as it is for pure bargaining problems. Because $F(T)$ is unaffected for $T \subset S$, all threats to coalition S are just as they were before; recall condition (3). If allocations from $F(S)$ that were part of the earlier solution $\sigma(S, \mathcal{G})$ still remain feasible, they must form the new solution; hence $\sigma(S, \mathcal{G}') = \sigma(S, \mathcal{G}) \cap F'(S)$. (If the intersection is empty, then of course the axiom is silent on what is to be done.)

Next, we ask that the solution satisfy a weak form of upper hemicontinuity. This restriction is automatically satisfied in the original Nash problem.

[UHC] Let $\mathcal{G}^k = (\mathcal{F}^k, d^k) \rightarrow \mathcal{G} = (\mathcal{F}, d)$ in the (product) Hausdorff metric. Suppose that for some coalition S , it is also the case that $U^\sigma(S, \mathcal{G}^k) \rightarrow U^\sigma(S, \mathcal{G})$ in the Hausdorff metric.⁹ Then, $x^k \in \sigma(S, \mathcal{G}^k)$ for all k and $x^k \rightarrow x$ imply $x \in \sigma(S, \mathcal{G})$.

These axioms are standard, just minimally modified for the coalitional setting. But our last condition is completely new. We invoke it to deal with non-trivial coalitional threats. Before formalizing this axiom, we normalize $d = 0$. For λ and x both in $\mathbb{R}^A$ for any index set A , write $\lambda \odot x := (\lambda_i x_i)_{i \in A}$, and let 1 denote the vector of ones. Say that λ is an *expansion* if $\lambda > 1$, and a *strict expansion* if $\lambda \gg 1$. By our normalization $d = 0$, λ can be viewed as a radial expansion net of inaction payoffs.

[Exp] Suppose that for some $\mathcal{G} = (\mathcal{F}, 0)$, $F(S) = \{x, y\}$ for $x \gg 0$ and $y \gg 0$, with both allocations unblocked so that $U^\sigma(S, \mathcal{G}) = F(S)$. Then, if $x \in \sigma(S, \mathcal{G})$, there is a strict expansion $\lambda \gg 1$, applied to each element of $F(S)$ to obtain $F'(S) = \{\lambda \odot x, \lambda \odot y\}$, with $F'(T) = F(T)$ for all other T , such that in the new game $\mathcal{G}' = (\mathcal{F}', 0)$, $\lambda \odot x \in \sigma(S, \mathcal{G}')$.

To understand [Exp], recall that d has been normalized to zero, by invariance. So an expansion is a multiplicative transform of payoffs relative to the disagreement point. Panel (a) of Figure 1 depicts a two-point feasible set $F(S) = \{x, y\}$ with no effective threats. Suppose

⁹We maintain the convention that the Hausdorff distance between the empty set and any nonempty set is $+\infty$, and zero between two empty sets. In particular, $U^\sigma(S, \mathcal{G}^k) \rightarrow U^\sigma(S, \mathcal{G})$ requires $U^\sigma(S, \mathcal{G})$ to be nonempty whenever the sets $U^\sigma(S, \mathcal{G}^k)$ are.

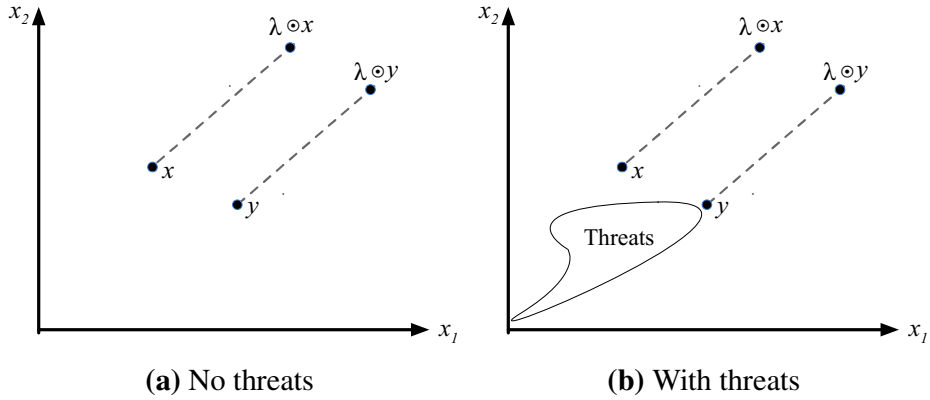


Figure 1. The expansion axiom applied to the feasible set $F(S) = \{x, y\}$.

that x is in the solution of this problem. The invariance axiom [Inv] tells us that *no matter* which expansion vector λ we use to scale x and y , $\lambda \odot x$ will be in the solution of the new bargaining problem with feasible payoff vectors $F'(S) = \{\lambda \odot x, \lambda \odot y\}$. (A stronger form of) [Exp] is therefore automatically implied by [Inv], when there are no effective threats.

Panel (b) illustrates the same two-point feasible set. This time there are nontrivial coalitional threats, and [Exp] is no longer implied by [Inv]. However, if x is the only outcome to start with, [UHC] guarantees that for some strict expansion λ , $\lambda \odot x$ will be a solution in the feasible set $\{\lambda \odot x, \lambda \odot y\}$, under the old threats. [Exp] is once again redundant. It is automatically implied, this time by [UHC].

[Exp] therefore has extra bite *only* when x and y are both in the solution to begin with. In that case, [Exp] is a weak tie-breaking requirement: if x and y are both acceptable resolutions of the bargaining problem, it should not be the case that *every* expansion of the feasible set strictly overturns x in favor of y . Some scaling retains x .

There is an equivalent way to read [Exp], one that also explains why the disagreement point (normalized to 0) serves as the origin of these expansions. With [Inv] in hand, expanding $F(S)$ while holding all subcoalitional feasible sets fixed is the same as holding $F(S)$ fixed while radially shrinking every subcoalitional feasible set toward inaction. That is:

Remark 7. Under [Inv], [Exp] is equivalent to the following statement. Suppose that for some $\mathcal{G} = (\mathcal{F}, 0)$, $F(S) = \{x, y\}$ for $x \gg 0$ and $y \gg 0$, with both x and y unblocked, and $x \in \sigma(S, \mathcal{G})$. Then there is μ with $1 \gg \mu \gg 0$ such that $x \in \sigma(S, \mathcal{G}'')$, where

$F''(S) = F(S)$ and $F''(T) = \mu_T \circ F(T)$ for all other T . That is, if x is a solution, it remains one under *some* strict “weakening” of every subcoalition toward inaction.¹⁰

“Toward inaction” is precisely the direction of powerlessness. A coalition is ineffective when $F(T) = \{d_T\}$, and radial contraction toward d is therefore the transformation that captures growing weakness.¹¹ Our axiom implies only that agreements survive *some* such weakening of the threats that surround them. That has significant implications.

2.3. Characterization.

Theorem 1. *A solution σ satisfies Axioms [Inv], [IIA], [UHC], [PO] and [Exp] if and only if, for every coalition S , it maximizes the (possibly asymmetric) product of payoffs in excess of disagreement values within the set of unblocked allocations $U^\sigma(S, \mathcal{G})$ for S :*

$$(5) \quad \sigma(S, \mathcal{G}) = \arg \max_{x \in U^\sigma(S, \mathcal{G})} \prod_{j \in S} [x_j - d_j]^{\gamma_j(S)},$$

for some vector $\gamma(S) > 0$ independent of $\mathcal{G}$, where $U^\sigma(S, \mathcal{G})$ is recursively defined in (2), and where it is understood that $\sigma(S, \mathcal{G})$ is empty if and only if $U^\sigma(S, \mathcal{G})$ is empty.

We will call σ the (*consistent*) *coalitional solution*. We discuss it extensively in Section 2.4. Note that the weights in $\gamma(S)$ can vary across players. Some of them might even be zero, though we do not take such cases seriously. In fact, if all players are symmetric, as in Nash’s original setting, the weights in $\gamma(S)$ must all be equal and *a fortiori* all strictly positive. To show this, it suffices to place our symmetry assumption only on bargaining problems in which all coalitions are ineffective:

[Sym] Suppose that for some game $\mathcal{G}$ and coalition S , every $T \subset S$ is ineffective and $d_i = d_j$ for all $i, j \in S$. If, for every permutation ν of S , $y(\nu) := (y_{\nu(i)})_{i \in S} \in F(S)$ for any $y \in F(S)$, then for every permutation ν , $x(\nu) := (x_{\nu(i)})_{i \in S} \in \sigma(S, \mathcal{G})$ for any $x \in \sigma(S, \mathcal{G})$.

¹⁰To see the equivalence, apply the affine transformation $z_i \mapsto \mu_i z_i$ with $\mu_i = 1/\lambda_i$ to the expanded game $\mathcal{G}'$ of [Exp]: it returns $F'(S) = \{\lambda \circ x, \lambda \circ y\}$ to $\{x, y\}$, carries $F(T)$ to $\mu_T \circ F(T)$, leaves $d = 0$ in place, and, by [Inv], carries $\lambda \circ x \in \sigma(S, \mathcal{G}')$ to $x \in \sigma(S, \mathcal{G}'')$.

¹¹Put another way, the anchoring of expansions at the disagreement point is not an *ad hoc* choice. An expansion anchored at any other point would either be defined by objects derived from the feasible set itself, or would presuppose that non-inaction coalitional threats act as reference points for increased “powerlessness”.

Corollary 1. *A solution σ^{sym} satisfies all the axioms of Theorem 1 and [Sym] if and only if for every game $\mathcal{G}$ and every coalition S ,*

$$(6) \quad \sigma^{\text{sym}}(S, \mathcal{G}) = \arg \max_{x \in U^{\sigma^{\text{sym}}}(S, \mathcal{G})} \prod_{j \in S} [x_j - d_j],$$

where it is understood that $\sigma^{\text{sym}}(S, \mathcal{G})$ is empty if and only if $U^{\sigma^{\text{sym}}}(S, \mathcal{G})$ is empty.

2.4. Discussion. The coalitional solutions in (5) and (6) are recursively defined and characterized under the restriction that all threats to a coalition come from subcoalitions of that coalition. One implication of the recursion is that the solution for a coalition S does double duty as a set of threats that coarser supersets of S must overcome in order to find *their* solutions. This is a consequence of conditions (2) and (3) on the solution, though we remark that Corollary 1 can be shown to stand independently of (3).

Given this somewhat schizophrenic nature of our solution, how then do we view the “solution to the game as a whole”? That requires additional perspective. We could adopt the point of view that an attempt to write an agreement begins with the grand coalition N . If $\sigma(N, \mathcal{G})$ is nonempty, N is *viable*. Otherwise, N must disintegrate as its attempts to reach an agreement from $F(N)$ are blocked, and the end result can be viewed as some partition π of N into coalitions $S \in \pi$, with each such coalition viable: $\sigma(S, \mathcal{G}) \neq \emptyset$. Section 4.3 develops this idea in more detail using farsighted objections.

This idea is easily extended to viewpoints where some $\bar{\pi}$ is the presumed parent structure, “above” which no coalition can form. Then our structures of interest would be the viable substructures that emanate from $\bar{\pi}$ (possibly $\bar{\pi}$ itself), again via farsighted objections.

Or we could ask of a coalition S of interest — often, the grand coalition — whether $\sigma(S, \mathcal{G})$ is empty or not. That is in the spirit of the core, though σ is far from being the core itself.

We turn now to the details of the solution. Given our characterization, a threat y , whether by an individual or a coalition, only affects the solution x when it is binding to begin with; i.e., $x_i = y_i$ for some i . All threats appear as standard constraints to the maximization of the Nash product. They are not somehow “subtracted” from the available payoff set, as one might imagine they would be in some direct generalization of Nash bargaining.

This observation shows that the subtraction of individual disagreement points — to be interpreted as the payoffs from complete inaction by all parties — can be a misleading

guide to generalization. That subtraction is an implication of the invariance axiom. Once applied to the disagreement points, there is no room to apply it again to any threat.

This alternative approach to constrained surplus division, with coalitional threats appearing as conventional constraints, has appeared before in the noncooperative bargaining literature. To our knowledge, Binmore, Shaked, and Sutton (1989) are the first to argue for individual outside options as conventional constraints, though they do not consider non-singleton coalitions. They argue that a threat can only be credible when a bargainer, by leaving the room, can obtain a higher payoff elsewhere; see the quotation in the Introduction. It is true that the disagreement payoff is subtracted, but that payoff is only derived from complete inaction (e.g., between bargaining rounds in the noncooperative model) and not from the explicit adoption of an outside option.

Compte and Jehiel (2010) (and before them Chatterjee, Dutta, Ray, and Sengupta 1993 for convex characteristic functions) arrive at a solution related to (6). Their solution, which Compte and Jehiel (2010) call the coalitional Nash solution, maximizes the Nash product over the unblocked set $U^{\text{feas}}(N)$ in (4), rather than the recursively defined unblocked set (2).¹² This can result in a markedly different set of outcomes.¹³ Example 1 illustrates this using a TU game: there is an affine transform of payoffs such that $d = 0$ and such that for each S , $F(S) = \{x \in \mathbb{R}_+^S \mid \sum_{i \in S} x_i \leq v(S)\}$ for some scalar $v(S)$.

Example 1. Consider an assignment game in the spirit of Shapley and Shubik (1971). One unit of housing is available from a seller (player 1). There are two buyers — players 2 and 3 — with valuations 5 and 4, transferable via money. Formally, $N = \{1, 2, 3\}$, $v(\{i\}) = v(\{2, 3\}) = 0$ for all i , $v(\{1, 2\}) = 5$, $v(\{1, 3\}) = 4$, while $v(\{1, 2, 3\})$ is the largest value obtainable via matching the agents, and so equals 5. The unblocked set for the grand coalition N is, therefore:

$$U^{\text{feas}}(N) = \{x \in \mathbb{R}_+^3 \mid x_1 + x_2 + x_3 \leq 5; x_1 + x_2 \geq 5; x_1 + x_3 \geq 4\},$$

where to ease the writing both here and later, we drop “ $\mathcal{G}$ ” in the notation for unblocked sets and solutions when there is no risk of confusion. The Pareto-frontier of $U^{\text{feas}}(N)$

¹²Binmore, Rubinstein, and Wolinsky (1986) show how the traditional Nash solution could also arise in a noncooperative setting, when there is some exogenous probability that the game ends, with individuals holding their outside options. While this is an interpretation that we choose not to pursue here, the approach has potentially interesting implications in a coalitional setting.

¹³Note that the imposition of consistency does not always result in an altered solution. That depends on the other axioms being applied. If we use the unblocked set $U^{\text{feas}}(S)$ from (4), and only apply the Pareto optimality axiom, the solution is the standard core for coalition S . As we know from Ray (1989), this standard solution is unchanged even if we restrict blocking to be consistent as in (2).

is precisely the core of the assignment game and it is “elongated”, in the terminology of Shapley and Shubik, with $x_1 \in [4, 5]$ and $x_2 = 5 - x_1$, with player 3 always cut out: $x_3 = 0$.

Maximizing the product of payoffs over $U^{\text{feas}}(N)$, as Compte and Jehiel (2010) do, makes no difference: the product of unblocked payoffs is always zero ($x_3 = 0$ in the core). Or perhaps one might adopt a lexicographic extension of their solution and declare $(4, 1, 0)$ to be the outcome. In contrast, if coalitions $\{1, 2\}$ and $\{1, 3\}$ employ Nash bargaining over *their* unblocked sets, then $\sigma^{\text{sym}}(\{1, 2\}) = \{(2.5, 2.5)\}$ and $\sigma^{\text{sym}}(\{1, 3\}) = \{(2, 2)\}$, and so the consistently defined unblocked set for N is given by:

$$U^{\sigma^{\text{sym}}}(N) = \{x \in F(N) \mid x_1 + x_2 + x_3 \leq 5; \max\{x_1, x_2\} \geq 2.5; \max\{x_1, x_3\} \geq 2\},$$

and maximizing the symmetric product of payoffs on this set — as we do — yields:

$$\sigma^{\text{sym}}(N) = \{(2.5, 1.25, 1.25)\}.$$

Observe that this solution is not in the core of the assignment game, but is nevertheless immune to consistently defined blocks. $\diamond$

But at a more fundamental level, Theorem 1 lives within a different tradition. It proceeds from an axiomatic perspective. Despite the emergence of the Nash product, the two sets of results are truly different. There is no tight game form with specific protocols which lay down the order of proposals, or insist that rejections occur sequentially, or stipulate that a unit of discounted time must pass between offers. Like Nash (1950), we do not restrict ourselves to finely specified formats for interaction, but rather use building blocks that are unrelated to specific protocols. Put another way, our characterization and that of Compte and Jehiel (2010) belong to the realm of the Nash program, where cooperative and non-cooperative developments of the same concepts are intended to reinforce and inform each other. We should add that a protocol whose stationary outcomes implement the *consistent* coalitional solution — with subcoalitional threats themselves generated by equilibrium bargaining — remains an open problem.

That said, there is also a literature extending the Nash bargaining solution in cooperative, often axiomatically derived settings. Kaneko (1980), Zhou (1997), Mariotti (1998), and Xu and Yoshihara (2020) extend the Nash bargaining solution to nonconvex feasible sets with n players, where $n \geq 2$. They do so for pure bargaining problems, i.e., when there are no coalitional threats. More relatedly, Burguet and Caminal (2016) extend the Nash solution to TU games, where every coalition S has a scalar “worth” of $v(S)$. Outside options ω_j^S for

$j \in S$ are subtracted and the rest divided equally, so that the payoff for any $i \in S$ is

$$(7) \quad u_i^S = \omega_i^S + \frac{1}{|S|} \left[v(S) - \sum_{j \in S} \omega_j^S \right].$$

In turn, ω_i^S is given by i 's expected payoff from joining other coalitions. An endogenous probability distribution governs the formation of such coalitions, with payoffs calculated just as in (7), so the vector $\{u_i^S\}$ is an equilibrium object that runs over all players and coalitions. Coalitional threats are fully subsumed in these individual payoffs and assumed *not to be* conventional. A related approach that collapses coalitional threats into individual values can be found in Isbell (1960), whose work draws in turn on Harsanyi (1959).

In a similar vein, Serrano and Shimomura (1998) study an efficient payoff profile x that precipitates a particular feasible set on any pair $\{i, j\}$: all allocations in F that respect the granting of x_k to any $k \neq i, j$. The outside option of i is defined to be the maximum that she can get by joining some coalition T that doesn't include j , again paying off x_k to each $k \in T$. Serrano and Shimomura (1998) impose a consistency condition that asks (x_i, x_j) to be the Nash bargaining solution for the induced two-player game, for every pair $\{i, j\}$. In TU games, this approach characterizes the pre-kernel of Davis and Maschler (1965).

2.5. More on the Expansion Axiom. Among our axioms, [Exp] is the one that determines how threats enter the bargain, and Theorem 1 is the measure of its content. As already noted, parts of it hold for free under the other axioms. It is implied by [Inv] when no subcoalitions are effective, and by [UHC] whenever the solution is unique. It is also neutral across bargaining norms: suitably adapted, the Nash (1950) bargaining solution, the solution of Kalai and Smorodinsky (1975), the egalitarian solution of Kalai (1977), or the coalitional solution in Compte and Jehiel (2010) all satisfy it (see the Supplementary Notes for more). That said, it is not innocuous. Along with the other axioms it is central to Theorem 1. To illustrate, we construct a solution that satisfies all our axioms including symmetry but barring [Exp], which differs from our solution. (This also establishes *en passant* the independence of [Exp].¹⁴)

For any solution σ and $T \subset S$, let $\tilde{\sigma}(T)$ be the Pareto frontier of $\sigma(T)$. For each $i \in T$, let

$$m_i(T) := \frac{\min_{x \in \tilde{\sigma}(T)} x_i + \max_{y \in \tilde{\sigma}(T)} y_i}{2}$$

¹⁴The Supplementary Notes show that the other axioms are likewise independent.

be i 's mean over her worst and best payoffs in $\tilde{\sigma}(T)$. The average of these mean payoffs over all subcoalitions of S containing i is $a_i(S) := [\sum_{T \ni i} m_i(T)]/[2^{|S|-1} - 1]$, where $2^{|S|-1} - 1$ is the number of such coalitions. Now, we will show by example that there is a game $\mathcal{G}$ such that the solution described by

$$\phi(S) := \arg \max_{x \in U^\phi(S, \mathcal{G}), x \geq \beta a(S) + (1-\beta)d} \prod_{j \in S} [x_j - \beta a_j(S) - (1-\beta)d_j],$$

where β is the largest number in $[0, 1]$ such that $x \geq \beta a(S) + (1-\beta)d$ for some $x \in U^\phi(S, \mathcal{G})$,¹⁵ fails [Exp]. (It does satisfy the other axioms.) Under ϕ , coalitional constraints are clearly not “conventional”. *However, [Exp] eliminates this solution.* This is shown in:

Example 2. Consider a three-person game with $d = (0, 0, 0)$, in which all coalitions except N and $\{1, 2\}$ are ineffective. Assume that:

$$F(\{1, 2\}) = \{(6, 0), (0, 6)\} \text{ and } F(N) = \{y, z\},$$

where $y = (3, 3, 1)$ and $z = (2, 2, 4)$. Observe that $a = (1, 1, 0)$ and $\beta = 1$. So, because y and z are both unblocked,

$$(8) \quad \phi(N, \mathcal{G}) = \arg \max_{x \in \{y, z\}} (x_1 - 1)(x_2 - 1)x_3 = \{y, z\},$$

where the second equality follows from the fact that $(y_1 - 1)(y_2 - 1)y_3 = 4 = (z_1 - 1)(z_2 - 1)z_3$. We claim that ϕ does not satisfy [Exp]. To this end, consider any expansion $\lambda = (\lambda_1, \lambda_2, \lambda_3) \gg (1, 1, 1)$. We shall show that for the game $\mathcal{G}'$ where the only change is $F'(N) = \{\lambda \odot y, \lambda \odot z\}$,

$$(9) \quad \phi(N, \mathcal{G}') = \arg \max_{x \in \{\lambda \odot y, \lambda \odot z\}} (x_1 - 1)(x_2 - 1)x_3 = \{\lambda \odot z\},$$

which, given $y \in \phi(N, \mathcal{G})$ by (8), violates [Exp]. To show (9), let $\delta_i := \lambda_i - 1$ for each i . Then:

$$\lambda \odot y = (3[1 + \delta_1], 3[1 + \delta_2], 1 + \delta_3) \text{ and } \lambda \odot z = (2[1 + \delta_1], 2[1 + \delta_2], 4[1 + \delta_3]).$$

Let the maximand in ϕ be denoted by $f(x) = (x_1 - 1)(x_2 - 1)x_3$. Then we have:

$$f(\lambda \odot y) = (2 + 3\delta_1)(2 + 3\delta_2)(1 + \delta_3) = (4 + 6\delta_1 + 6\delta_2 + 9\delta_1\delta_2)(1 + \delta_3),$$

and

$$f(\lambda \odot z) = (1 + 2\delta_1)(1 + 2\delta_2)[4(1 + \delta_3)] = (4 + 8\delta_1 + 8\delta_2 + 16\delta_1\delta_2)(1 + \delta_3).$$

¹⁵If $a(S)$ is itself in $U^\sigma(S, \mathcal{G})$, then $\beta = 1$. Otherwise, $\beta \in [0, 1]$ is a scaling factor, and whenever $U^\phi(S, \mathcal{G}) \neq \emptyset$, a unique β satisfying the condition in the text must exist, given our domain restrictions.

Clearly, $f(\lambda \circ z) > f(\lambda \circ y)$, which implies (9). $\diamond$

3. Some Properties of the Consistent Coalitional Solution

Our aim in this Section is threefold. First, we show that our recursive definition admits a simple non-recursive characterization. Second, we show that in superadditive TU games, there is always a nonempty coalitional solution for every coalition, including the grand coalition, *even when the core is empty*. Third, we uncover a philosophical connection between the symmetric coalitional solution and an ethic based on “pragmatic egalitarianism”.

As before, our primitives are summarized by a characteristic function: a collection of nonempty compact sets of feasible payoffs $F(S)$ for every S satisfying the domain restriction [Dom]. To rule out uninteresting cases, we presume that $x \gg d_S$ for some $x \in F(S)$ whenever $F(S) \neq \{d_S\}$. Normalize d to 0.

3.1. A Non-Recursive Representation. Untangling a recursive definition can be hard work. It is generally necessary to navigate inductively through the outcomes for every sub-coalition to arrive at the corresponding solution for any given coalition. However, when feasible payoff sets for each coalition are *convex*, that recursive computation becomes unnecessary. To this end, define the *unconstrained Nash bargaining solution* $\Psi(T)$ for any coalition T and weights $\gamma(T)$ by ignoring all coalitional constraints:

$$\Psi(T) = \arg \max_{x \in F(T)} \prod_{j \in T} x_j^{\gamma_j(T)}, \text{ remembering } d = 0.$$

Now define a “naive” unblocked set for S using these (generally non-credible) solutions $\{\Psi(T)\}_{T \subset S}$ as threats:

$$(10) \quad U^{\text{naive}}(S, \mathcal{G}) := \{x \in F(S) \mid \text{there is no } T \subset S \text{ and } y \in \Psi(T) \text{ such that } y \gg x_T\}.$$

This is a simple set. No recursive definition is needed for it. Moreover, when $F(T)$ is convex and $\gamma(T) \gg 0$ for $T \subset S$, each potential threat $\Psi(T)$ is just a singleton. And yet:

Theorem 2. *Consider any game $\mathcal{G}$ such that $F(S)$ is compact and convex for every coalition S . Then $U^\sigma(S, \mathcal{G}) = U^{\text{naive}}(S, \mathcal{G})$, and so*

$$\sigma(S, \mathcal{G}) = \arg \max_{x \in U^{\text{naive}}(S, \mathcal{G})} \prod_{j \in S} x_j^{\gamma_j(S)};$$

that is, the consistent coalitional solution for any coalition need only guard against the threats posed by the unconstrained Nash solutions of its subcoalitions.

Remark 8. Theorem 2 shows that recursive consistency is equivalent to one-shot optimization against unconstrained Nash “threats”.¹⁶ (And, as already noted, there will typically be just one such threat per subcoalition.) The Supplementary Notes establish a more general version of this result by weakening convexity of the feasible sets, but also provide a counterexample to show that the convexity of feasible sets cannot be dispensed with entirely.

3.2. Viability and Efficiency. We now come to a classically Coasian question. Say that a solution is *efficient* if the outcome for the game as a whole is Pareto optimal. (That overall perspective is obviously required because a solution is defined on *all* coalitions, and many of them, e.g., the singletons, will produce inefficient outcomes.) The central concept of the core returns a clear negative answer to this question, even when the underlying game is strictly superadditive. Indeed, in a TU game, the core is nonempty if and only if that game is balanced (Bondareva 1963, Shapley 1967), a notion that is significantly more demanding than superadditivity. When balancedness fails, every efficient solution is blocked.¹⁷

Our coalitional solution, while core-like in its construction, is different. Subcoalitions cannot object in an unrestricted way; they must block with their own coalitional solutions. It turns out that these constraints on blocking, coupled with the particular structure of the coalitional solution, guarantees efficiency for all TU games under two mild conditions. First, Nash bargaining weights are *consistent*: there is $\gamma = (\gamma_1, \dots, \gamma_n) \gg 0$ such that within any S , the weights are given by $\gamma(S) = \{\gamma_i\}_{i \in S}$. Second, the TU game is superadditive; that is, for every pair of disjoint coalitions S and T , $v(S \cup T) \geq v(S) + v(T)$. (This is, of course, significantly weaker than the Bondareva–Shapley requirement of balancedness.)

Theorem 3. *Consider any superadditive TU game with consistent bargaining weights. Then the grand coalition is viable, and its coalitional solution is Pareto optimal.*

¹⁶The careful reader will observe that the proof of Theorem 2 applies to a wider class of solutions: those that emerge from the maximization of any system of coalitional welfare functions that are quasiconcave and strictly increasing in individual payoffs, under the assumption that all feasible sets are convex.

¹⁷Here is a standard example: $N = \{1, 2, 3\}$, $F(\{i\}) = \{0\}$ for all i , $F(\{i, j\}) = \{x \in \mathbb{R}_+^2 \mid x_1 + x_2 \leq 1\}$, and $F(N) = \{x \in \mathbb{R}_+^3 \mid x_1 + x_2 + x_3 \leq v(N)\}$, where $v(N) > 0$. Then if $v(N) \in (1, 3/2)$, the game is superadditive but has an empty core, and every efficient allocation is blocked. See also Example 3.

Theorem 3 stands in stark contrast to the possibility of an empty core in superadditive characteristic function games. As our proof shows, the contrast depends intimately on the use of the coalitional solution at every level, both in the solution to be achieved and in the blocks or threats that potentially impede the achievement of that solution. Our argument is constructive (see the Appendix), and makes crucial use of Theorem 2. We first build a particular structure and allocation, and then show that that allocation cannot be blocked by subcoalitions using *their* unconstrained Nash bargaining solution.

The following example illustrates how Theorem 3 relates to the core.

Example 3. Consider a TU game in which $N = \{1, 2, 3\}$, $v(\{i\}) = 0$ for all i , $v(S) = 0.8$ for all two-player coalitions S , and $v(N) = 1$. This is a superadditive game but it is not balanced, and its core is empty. Now consider the symmetric solution with weight $\gamma_i = 1$ for all i . The coalitional solution for every two-player coalition involves equal division, so each player gets 0.4. The unblocked set for N relative to these threats is not empty. It includes, for example, the allocation $(0.4, 0.4, 0.2)$ as well as all its permutations. Indeed, the coalitional solution consists of precisely these allocations. $\diamond$

Remark 9. Whether or not Theorem 3 generalizes to *all* TU games, and not just superadditive TU games, is an open question. That is, is there always a partition of the grand coalition with an efficient outcome? One can easily relax superadditivity to some degree if one is only interested in the viability of the grand coalition.¹⁸ The question is whether superadditivity can be dispensed with altogether. What we do know is that Theorem 3 does not generalize free of charge to non-TU games, even when they are superadditive. In such games, the feasible set can be arbitrary in ways that make it impossible to exploit the deeper structure of the coalitional solution, as we do in Theorem 3.¹⁹

3.3. Constrained Egalitarianism. Having established that the solution possesses strong efficiency properties, we conclude this Section by asking what distributive principle it embodies. To do so, we continue to restrict ourselves to TU games but additionally assume

¹⁸Then it will suffice to assume that the game is *grand coalition superadditive*, in the sense that $v(N) \geq \sum_{i=1}^m v(S_i)$ for any partition $\{S_1, \dots, S_m\}$ of N .

¹⁹The familiar roommate problem can be used to establish this claim. Suppose $N = \{1, 2, 3\}$, $F(\{i\}) = \{0\}$ for all $i \in N$, $F(\{1, 2\}) = \{(3, 2)\}$, $F(\{2, 3\}) = \{(3, 2)\}$, $F(\{1, 3\}) = \{(2, 3)\}$ and $F(N) = \{(3, 2, 1), (1, 3, 2), (2, 1, 3)\}$. Recalling that we only consider blocking from subcoalitions, every structure of the form $\{\{j, k\}, \{\ell\}\}$ is viable: no player in a two-player coalition will break away to precipitate the coalition structure of singletons. But this makes the grand coalition non-viable; every efficient allocation is blocked by some two-player coalition.

symmetric bargaining weights, so that payoffs can be compared and transferred across individuals, and are anonymized.

Recall that the “split-the-surplus” property of the Nash bargaining solution is a technical consequence of the invariance axiom.²⁰ And yet, given its undertone of “I set aside what’s mine, you set aside yours, and we divide the rest”, this property has assumed a moral undertone for “fairness”; see, e.g., the disputed garment principle in Aumann and Maschler (1985), or the analysis of the pre-kernel in Davis and Maschler (1965).

But our analysis has shown how fragile this interpretation is. Apart from the disagreement points, *all* coalitional threats enter as conventional constraints in our characterization. They affect the outcome only when it lies on the edge of the unblocked set. To us, that reopens the question of whether split-the-surplus can be viewed as an easy moral imperative. Our solution suggests a connection to a different ethic, one that we now explore.

The constrained egalitarian solution (Dutta and Ray 1989, 1991) combines a commitment to equality with the pragmatics of coalitional participation. A social ethic (egalitarianism, in the case at hand) is applied to the greatest degree possible, while remaining limited by the need to secure individual and coalitional buy-ins. The following example illustrates the notion of “pragmatic egalitarianism” used in Dutta and Ray:

Example 4. Consider a three-player TU game, with $v(N) = 1$, $v(\{1, 2\}) = 0.8$, $v(S) = 0$ for all other S , and $d = 0$. Coalitions want distributional equality but not at the cost of breaking up (this is formalized below). Because players 1 and 2 each earn zero on their own, coalition $\{1, 2\}$ can fully enjoy its egalitarian goals, implementing $(0.4, 0.4)$ if it were to form. But N is then constrained by the possibility that $\{1, 2\}$ could exit with its credible threat $(0.4, 0.4)$ if pushed too far. A “constrained egalitarian solution” is therefore given by the two allocations $(0.4, 0.3, 0.3)$ and $(0.3, 0.4, 0.3)$, which are the Lorenz-maximal elements of the unblocked set for N . This also happens to be the consistent coalitional solution. In contrast, the core is the set $\{x \in \mathbb{R}_+^3 \mid x_1 + x_2 + x_3 = 1, x_1 + x_2 \geq 0.8\}$ and the Compte-Jehiel (2010) solution is $\{(0.4, 0.4, 0.2)\}$. $\diamond$

Dutta and Ray (1989, 1991) generalize this idea of constrained egalitarianism to societies of all sizes. Consider two allocations x and y in $\mathbb{R}^k$ that add to the same total, arranged so that $x_i \leq x_{i+1}$ and $y_i \leq y_{i+1}$ for all $i = 1, \dots, k-1$. Say that x *Lorenz-dominates* (or is majorized by) y if $\sum_{i=1}^j x_i \geq \sum_{i=1}^j y_i$ for every $j = 1, \dots, k$, with strict inequality for

²⁰If payoffs are expected utilities, as Nash takes them to be, the subtraction of disagreement points is viewable as an additive shift, and any solution must be invariant to such shifts.

some j . This partial ordering agrees with the ethics of egalitarianism (see, e.g., Kolm 1969, Dasgupta, Sen, and Starrett 1973 and Fields and Fei 1978). For any set of allocations A adding to the same total, let $L(A)$ be its set of *Lorenz-maximal* elements.

Now consider any TU game $\mathcal{G}$ with worths $\{v(S)\}$, so that with d normalized to 0, $F(S) = \{x \in \mathbb{R}_+^S \mid \sum_{i \in S} x_i \leq v(S)\}$. We define the constrained egalitarian solution $E(S)$ for every coalition S . (We drop $\mathcal{G}$ in the arguments for economy of notation.) For any singleton coalition $\{i\}$, define $E(\{i\}) = \{v(\{i\})\}$. Recursively, consider any coalition S and suppose that we've defined $E(T)$ for every strict subset T of S . Then define

$$U^e(S) := \{x \in F(S) \mid x \text{ is unblocked by any } (T, y) \text{ with } T \subset S \text{ and } y \in E(T)\},$$

and, if $U^e(S) \neq \emptyset$, define the *constrained egalitarian solution* for coalition S by

$$E(S) = L(U^e(S)^+),$$

where $U^e(S)^+ := \{x \in U^e(S) : \sum_{i \in S} x_i = v(S)\}$ picks up all payoff vectors in $U^e(S)$ that sum to precisely $v(S)$. Proceed in this manner until all coalitions are covered. We can now connect these constrained egalitarian solutions to the coalitional solution, on the domain of TU games in which *every singleton coalition is effective*: $v(\{i\}) > 0$ for each player i or, equivalently, $\zeta_i > d_i$; recall Section 2.²¹

Theorem 4. *In a TU game in which every singleton coalition is effective, the symmetric coalitional solution σ^{sym} is a subset of the constrained egalitarian solution for every coalition S :*

$$\sigma^{\text{sym}}(S) \subseteq E(S).$$

Suppose additionally that a TU game is superadditive. Then for all S , $\sigma^{\text{sym}}(S)$ is nonempty, and is found by maximizing the Nash product over the set of allocations that are unblocked by any subcoalition using equal division.

Remark 10. Theorem 4 states that the coalitional solution with equal bargaining weights can be interpreted as constrained egalitarianism. The (symmetric) Nash product seeks equality of outcomes, but is restricted by coalitional threats. By Theorem 2, we need only consider naive threats imposed by unconstrained Nash bargaining. But in a TU game, that's just equal division; hence the result.

²¹Some restriction of this kind is necessary: the Supplementary Notes describe a five-player TU game, with zero worth for two of its singleton coalitions, in which $\sigma^{\text{sym}}(N)$ is *not* contained in $E(N)$. At the same time, the restriction is not necessary for the conclusion to hold in particular cases: all singleton worths are zero in Example 4, and yet the conclusion of Theorem 4 applies there.

As a final observation of separate interest, one can use Theorem 4 to ask when the coalitional solution is unique, at least up to a renaming of players.²² This is a deep question to which we do not have a complete answer. But Dutta and Ray (1991) provide some leads. They invoke the following partial ordering from Maschler and Peleg (1966): for two players i and j , say that $i \succsim j$ if for all $S \subseteq N - \{i, j\}$ (possibly empty), $v(S \cup \{i\}) \geq v(S \cup \{j\})$. This ordering is transitive (cf. Maschler and Peleg 1966), but additionally, there is a wide class of games for which $\succsim$ is complete. In such cases, we have

Corollary 2 (to Theorem 4 and Dutta and Ray 1991, Theorem 3). *If every singleton coalition is effective and $\succsim$ is complete, then every allocation in $\sigma^{\text{sym}}(N)$ is identical up to a permutation of the players.*

We omit the proof, which is an immediate consequence of combining Theorem 4 and Dutta and Ray (1991), Theorem 3. It remains to be seen whether we can do better. Our completeness condition is sufficient but far from necessary. For instance, it can be shown that the coalitional solution must contain a single allocation (again up to possible permutations) for all three-player games, though $\succsim$ is not always complete for such games.²³

4. Embedding $\mathcal{G}$ into a Strategic Environment

Thus far we have studied how a given coalition bargains internally. We now ask a different question: which structures emerge when coalitions interact strategically in an environment with externalities? Our coalitional bargaining solution becomes the internal decision rule for each coalition and it continues to play its dual role both as a set of outcomes and as a set of potential threats to larger coalitions. But in addition, given cross-coalitional externalities, the viability of the entire coalition *structure* now becomes of interest.

Our embedding exercise is broadly related to the “Nash-in-Nash” literature (Horn and Wolinsky 1988, Ho and Lee 2017, 2019, Collard-Wexler, Gowrisankaran, and Lee 2019, or Bagwell, Staiger, and Yurukoglu 2020). That literature studies bilateral (Nash) bargaining with individual outside options determined in noncooperative (Nash) equilibrium.²⁴

²²In such cases, it is easy to see that $\sigma^{\text{sym}}(N)$ must coincide with $E(N)$. Otherwise it is possible that $\sigma^{\text{sym}}(N)$ is a strict subset of $E(N)$; Dutta and Ray (1991), Example 4, can be used to demonstrate this.

²³The coalitional solution also exhibits a single allocation when the grand coalition has the highest average worth; that is, $v(N)/|N| \geq v(S)/|S|$ for every coalition S , but $\succsim$ might be incomplete.

²⁴Outside options are sometimes netted out; see, e.g., Lee and Fong (2013). Sometimes, they are treated as conventional constraints, as in Ho and Lee (2019) or Collard-Wexler, Gowrisankaran, and Lee (2019). This

Our embedding allows us to think about groupings or cartels that go beyond pairs. The applications we have chosen are specifically meant to illustrate such situations.

One theme that will emerge from the applications is that coalitional outcomes are often inefficient. And yet, the intuition behind this inefficiency is not the usual one that underlies noncooperative interaction. For instance, in any two-person strategic setting with noncooperative inefficiencies, such as the Prisoner's Dilemma, the grand coalition is viable and the outcome is efficient. Indeed, the greater the inefficiency of noncooperative interaction, the stronger is the impetus to write agreements that restore efficiency.

Nor is it the case that an efficiency failure can be traced to core-like constraints alone; see Theorem 3 for characteristic functions. In short, the efficiency failure cannot be traced to cross-coalitional externalities *alone*, nor can it be traced to core-like constraints *alone*. It is the interaction of these two features that drives such a failure (when it happens).

4.1. Partitions and Payoffs. A *coalition structure* π partitions N into coalitions. They interact noncooperatively, while attempting to internally cooperate, subject to subcoalitional threats, just as before. We are interested in identifying “viable” coalition structures.

The central change from our baseline exercise is the replacement of the set $F(S)$ by the set $F(S, \pi)$, defined for every coalition structure π — a partition of the grand coalition — and for every coalition $S \in \pi$. That is, the available payoff profiles to any coalition could depend on the entire coalition structure within which that coalition is embedded. This object replaces the characteristic function, and builds in inter-coalitional externalities as a reduced form of an underlying equilibrium. We maintain the same conditions on $\{F(S, \pi)\}$ as we did for the collection $\mathcal{F} = \{F(S)\}$ under [Dom].

To see how $F(S, \pi)$ might be derived from a game in strategic form, suppose that for any coalition structure π each coalition $S \in \pi$ has a set $A(S)$ of joint actions, possibly payoff-relevant to both itself and other coalitions. Assume that each coalition receives some aggregative payoff $v_S(a)$ from every action profile $a \in \prod_{S \in \pi} A(S)$. A *coalitional Nash equilibrium* is an action profile a^* such that for every $S \in \pi$,

$$(11) \quad v_S(a^*) \geq v_S(a_S, a_{-S}^*) \text{ for all } a_S \in A(S),$$

where $a_{-S} = (a_T)_{T \in \pi, T \neq S}$. That is, every coalition best-responds to the actions of other coalitions. If every structure π has a unique coalitional Nash equilibrium, that effectively

literature also permits an individual to enter into several (bilateral) agreements, a feature that our coalition structures do not accommodate. The embedding of $\mathcal{G}$ into a network setting is a task for future research.

generates a *partition function* $v(S, \pi)$ for each π and every $S \in \pi$ (Lucas 1963, Thrall and Lucas 1963, Myerson 1977, Ray and Vohra 1999). Note that $\{v(\{i\}, \pi^s)\}$ is the usual Nash equilibrium payoff profile, where π^s is the coalition structure of singletons.

We assume this uniqueness throughout in what follows: multiplicity at the strategic stage is deliberately set aside. All four applications below satisfy this assumption. If called upon to study some other underlying game with multiple equilibria for some structure, we would presume a selection has been made before our theory applies — just as an equilibrium is selected before outside options are computed in the Nash-in-Nash literature.

With access to a within-coalition technology to transfer payoffs — possibly at some cost — coalition S allocates $v(S, \pi)$ to trace out the feasible set $F(S, \pi)$ posited above. (It is presumed that the use of these transfers has no relevance to players not in S .) When every coalition in π does this, we obtain a *coalitional equilibrium payoff profile under π* : x with $x_S \in F(S, \pi)$ for every $S \in \pi$.

4.2. Viable Structures. As before, S 's choices will serve both as its solution set given π , *and* as its threat set for yet coarser coalitions (including the grand coalition) that might attempt to form. The resulting recursion is a bit more involved than for characteristic functions, because we need to keep track of the noncooperative interaction at every stage.

The recursion will set up a threat set $\Theta(T, \pi)$ for each potential breakaway $T \subset S \in \pi$, describing what T can “credibly get” if it departs from S starting from π .²⁵ Define

$$U(S, \pi) = \{x \in F(S, \pi) \mid x \text{ is not blocked by any } T \subset S \text{ with } y \in \Theta(T, \pi)\},$$

and recalling that d_S has been normalized to 0, let

$$(12) \quad \sigma(S, \pi) := \arg \max_{x \in U(S, \pi)} \prod_{i \in S} x_i^{\gamma_i(S)}.$$

This is S 's coalitional solution under π (given the threats). The solution for π as a whole is the set $\Sigma(\pi) := \{x \in \mathbb{R}_+^N \mid x_S \in \sigma(S, \pi) \text{ for every } S \in \pi\}$. Say that π is *viable* if $\Sigma(\pi)$ is nonempty (equivalently, if $U(S, \pi)$ is nonempty for every $S \in \pi$).

It remains to recursively build $\Theta(T, \pi)$ for $T \subset S \in \pi$, applying (12) as we go.

²⁵Our definition does not allow for breakaways to combine with other players who are not in their parent coalition; threats can only serve to refine a coalition structure. The Supplementary Notes contain a discussion of the issues that could arise with “external” blocking. With the grand coalition at the apex of our recursion, any conceivable coalition structure can potentially form with or without our restriction.

4.3. Recursion. For the singleton structure π^s , $\sigma(\{i\}, \pi^s)$ is just the Nash equilibrium payoff to player i . Declare π^s to be viable. Next comes the recursive step. Pick any π of cardinality m , where $m \in \{1, \dots, n-1\}$. For any refinement π' of cardinality $m+1$ or greater, suppose that we have all the threats $\{\Theta(T, \pi')\}_{T \subset S \in \pi'}$ to π' , and in particular that we know whether or not π' is viable. We will now use this information to recursively define the threats to π .

For any $S \in \pi$ and potential breakaway $T \subset S$, write $\pi^T := \pi_{-S} \cup \{T, S-T\}$. By the recursive step just described, we know all possible threats to π^T . If π^T is viable, then the *terminal structure* after T 's move is π^T , and T 's coalitional solution represents T 's threats:

$$\Theta(T, \pi) = \sigma(T, \pi^T) \text{ if } \pi^T \text{ is viable.}$$

If π^T is not viable, every coalitional equilibrium payoff profile in π^T is blocked, causing π^T to disintegrate until we arrive at some viable terminal structure.²⁶ Formally, setting $\pi_0 := \pi^T$, we have for every coalitional equilibrium payoff profile x^0 under π_0 , a sequence $(\pi_0, x^0) \xrightarrow{T_1} (\pi_1, x^1) \dots \xrightarrow{T_k} (\pi_k, x^k)$ such that

- (i) $T_j \subset S$ for some $S \in \pi_{j-1}$ and $\pi_j = \pi_{j-1}^{T_j}$ for all $j = 1, \dots, k$;
- (ii) For each j , x^j is a coalitional equilibrium payoff profile under π_j ;
- (iii) π_j is not viable for each and every $j = 1, \dots, k-1$, but π_k is viable; and
- (iv) $x^k \in \Sigma(\pi_k)$, and for every $j = 1, \dots, k$, $x_{T_j}^k \gg x_{T_j}^{j-1}$.

Following Harsanyi (1974), Chwe (1994) and Ray and Vohra (2015b), we can view the above sequence $(T_1, \pi_1), \dots, (T_k, \pi_k)$ as a *farsighted objection* to the coalitional equilibrium payoff x^0 under $\pi_0 = \pi^T$. Each step j is “justified” as follows: coalition T_j precipitates x^j from x^{j-1} , knowing that at the final point x^k in the solution for π_k , each of its members will be better off relative to where they are at x^{j-1} .²⁷ Let $\mathcal{P}(\pi^T)$ refer to all terminal structures of the form π_k that result from farsighted objections to π^T . The solutions at these viable terminal structures determine the threats from coalition T . In summary,

$$(13) \quad \Theta(T, \pi) = \begin{cases} \sigma(T, \pi^T) & \text{if } \pi^T \text{ is viable} \\ \{x_T | x \in \Sigma(\pi') \text{ for some } \pi' \in \mathcal{P}(\pi^T)\} & \text{if } \pi^T \text{ is not viable.} \end{cases}$$

²⁶The process cannot continue indefinitely because π^s serves as a limit to the refinements.

²⁷One can refine farsighted objections even further, by formalizing additional restrictions as various notions of maximality, studied, for example, in Dutta and Vohra (2017) and Ray and Vohra (2019). While it would be possible to add such refinements, in the interest of simplicity, we shall refrain from doing so here.

Remark 11. Implicit in (13) is the idea that if π^T is viable, then T can choose *any* coalitional solution in $\sigma(T, \pi^T)$ as a threat. That's its prerogative; it has sovereignty over those choices. On the other hand, if π^T is not viable, we allow T to threaten with *any* farsighted objection emanating from π^T .²⁸ In particular, T presumes that all players will adopt common expectations about the ensuing farsighted path, after T has brought about the structure π^T . That presumption is based on T 's optimism about the path that will follow.

Remark 12. As before, we could suppose that all attempts at agreement begin with the grand coalition, N . If $\{N\}$ is viable, the solution to the game is $\Sigma(\{N\})$. Otherwise, If N is not viable, it must disintegrate into some viable partition serving as a farsighted objection to some payoff profile in N . The collection of solutions in all such viable partitions then constitutes the set of solutions for the game. This idea is easily extended to cases in which $\bar{\pi}$ is some presumed parent structure, “above” which no coalition can form. Then our structures of interest would be $\bar{\pi}$ itself, if viable, or otherwise the viable substructures that are reached as farsighted objections to equilibrium payoff profiles in $\bar{\pi}$.

Remark 13. Our definition does not immediately reduce to the definition for characteristic functions. Farsightedness was not an issue there, because coalitions had sovereignty over their allowable set of payoffs following a block. It was therefore more intuitive to use the term “viable coalition” rather than “viable structure”. Nevertheless, the two definitions are easily seen to be completely equivalent for characteristic functions: if a coalition blocks an allocation, then there is a farsighted block to that same allocation, and vice versa.

A final observation: the possible lack of viability is a feature, not a bug. That is, the viability of a structure could fail for two reasons. One is that breakaway threats overpower what some coalitions in the structure can achieve. The other is a technical failure of existence, such as openness or the lack of continuity. We are not really interested in these latter failures.²⁹ The real test of viability comes from the threats posed by breakaways, and this leads to sharp predictions regarding the structures we expect to see in applications.

4.4. Applications. The following applications are designed to highlight the issues that arise when studying viable coalition structures. As in Section 3.2, we place particular focus on the Coasian question: is there some viable *and* efficient coalition structure? With cross-coalitional interactions, the answer is typically negative.

²⁸Note that T need not be a coalition of the terminal structure π' in the second line of (13).

²⁹Indeed, under our assumption that there is a unique coalitional equilibrium for every coalition structure, the existence of a solution to a game is assured by the fact that the singleton coalition structure is viable.

As discussed in Section 4, the failure of efficiency cannot be traced to the usual externality argument under noncooperative interaction, because our solution concept *allows those externalities to be deliberately internalized*. Indeed, in any two-person strategic setting our solution is indeed efficient. However, as the applications below will show, efficiency is not inevitable with three or more players, though it could happen; see, for instance, Observations 1 and 2 below and the sharp contrast between them.

The applications that follow highlight distinct features of the analysis so far.

Example 5, a “Cournot cartel”, studies a Cournot oligopoly in which firms can endogenously form coalitions. We study the viability of the grand cartel of all firms. With three symmetric firms, the departure of a single firm from the cartel causes a disintegration down to singletons, nullifying the threat posed by the departing firm. In contrast, with more firms or in non-symmetric settings, a single firm can pose a bigger threat because the other firms so find it in their interest to stick together, and only inefficient intermediate structures are then viable. Farsightedness plays a central role, as they do in the other examples as well.

Example 6, an “R&D race”, emphasizes coalitional outside options, and how to forestall them. In the context of organizing R&D consortiums, it shows that the departure of a larger consortium becomes a binding threat to the sustainability of the grand consortium of all firms. In reaction, the grand consortium divides its payoff unequally, even when its members are symmetric *ex ante*, but retains its viability in this manner. This example contrasts sharply with Example 5, in which the three-firm cartel can get away with equal division of its monopoly profit.

Example 7, a “global environmental problem” across countries, is designed to show how *non-singleton* groupings of countries might break away from cooperation when it (correctly) predicts that the others will cling to cooperation. We show that the grand coalition is unable to appease every such breakaway. The example is chosen to demonstrate how this can happen even when singleton countries wish to stay within the grand coalition, and in this sense it highlights a new feature relative to Examples 5 and 6.

Finally, Example 8, on “hospital networks”, contributes to a recent literature by viewing a hospital network as a coalition of providers that takes membership premiums up front from a population of heterogeneous patients. This example explicitly models the demand side and permits a welfare-based analysis of network formation, by distinguishing between hospital profits on the one hand and consumer surplus for patients on the other. Both efficient and inefficient network structures are studied.

These applications show that a plethora of outcomes is possible even in relatively simple settings. Together, they highlight a rich and varied set of features, involving coalitional viability in some cases but not in others. But that does not mean that the theory is devoid of power: any *particular* setting can be fully solved for, and there is a clear prediction regarding the viability of every coalition structure in each specific problem.

There is an important reason why equilibrium coalition structures vary so richly with the underlying strategic game. It is that inefficiency can happen (or viability lost) when there isn't too much of inefficiency to begin with. With the possibility of "excessive" inefficiency following a deviation, there would be too much to lose — the edifice would crumble very far — and that eradicates the gains for a renegade subcoalition, creating a safe, relatively efficient ledge. For instance, in Example 5, a single firm refrains from deviating, knowing that it would lead the entire structure to collapse. In contrast, in Example 8, the inefficiency occurs because there isn't "too much" of it when a single country breaks off from an agreement — the others stay together. The viability of an intermediate coalition structure is a double-edged sword: it avoids large degrees of inefficiency, but also threatens the viability of an efficient outcome. It is for this reason that safe ledges are rarely spaced next to each other in the recursive ladder: each ledge is shored up by the possibility of a precipitous fall.

Finally, these applications exemplify the full integration of Nash bargaining into coalition formation, as opposed to other allocation rules such as unconditional equal division (Farrell and Scotchmer 1988, Bloch 1996), or no rule at all (Ray and Vohra 1997). Alternatives such as equal division often bypass the need to hold coalitions together, thereby generating widespread inefficiency. Our coalitional solution is constrained by the pragmatics of coalitional integrity. Nevertheless, that does not entirely prevent inefficiency, as we saw on more than one occasion in the examples.

Notation. As before, π^s is the structure of singletons. Write π^g for the grand structure $\{N\}$, and π^{jk} for any three-player intermediate structure $\{\{j, k\}, \{i\}\}$, in which j and k form a coalition while i stands alone. In the notation of Section 4.3, π^{jk} is the structure induced from π^g when $\{j, k\}$ departs — equivalently, when $\{i\}$ does.

Example 5. [Cartels in Oligopoly] n firms sell a homogeneous product. They can form coalitions. Coalitions in a structure interact noncooperatively in the sense of Cournot. Within a coalition, transfers are used freely to implement coalitional solutions. We study viable coalition structures; in particular, whether or not full cartelization occurs.

There is a linear inverse demand curve, with units normalized so that $p = A - x$, where p is price, x is aggregate output, and $A > 0$. Firm i has constant unit cost c_i . Define $C := \sum_j c_j$. Under the singleton structure, it is easy to show that each firm i produces

$$(14) \quad x_i = \frac{A + C}{n + 1} - c_i,$$

provided the following necessary and sufficient condition for interior output holds:

$$(15) \quad A + C > (n + 1)[\max_j c_j],$$

a condition that we assume in what follows. Using (14) to calculate aggregate quantity and then price p , we obtain firm i 's equilibrium profit:

$$(16) \quad (p - c_i)x_i = \left(\frac{A + C}{n + 1} - c_i \right)^2.$$

Now consider any coalition structure π . Given linear costs and costless transfers, all production within a coalition will be carried out by its lowest-cost members — the other members are effectively paid to stay out of the way. The Supplementary Notes show that condition (15) guarantees positive production by every coalition in any structure in coalitional Nash equilibrium. That equilibrium is governed by the same equations that we had for the singletons, and in particular, the aggregate profit for any coalition $S \in \pi$ is given by

$$(17) \quad v(S, \pi) = \left[\frac{A + C(\pi)}{|\pi| + 1} - c(S) \right]^2,$$

where $c(S) := \min_{i \in S} c_i$, $C(\pi) := \sum_{T \in \pi} c(T)$ and $|\pi|$ is the cardinality of π . We can use this formula to characterize viable structures in any given setting. In our example, we use just three firms. Then the singleton structure π^s yields a profit of

$$(18) \quad v(\{i\}, \pi^s) := \left[\frac{A + C}{4} - c_i \right]^2 = \frac{(A - c)^2}{16},$$

the last equality holding in the special case when all firms are symmetric with $c_1 = c_2 = c_3 = c$. Likewise, in an intermediate structure of the form π^{jk} , we see that

$$(19) \quad \begin{aligned} v(\{\ell\}, \pi^{jk}) &:= \left[\frac{A + \min\{c_j, c_k\} - 2c_\ell}{3} \right]^2 = \frac{(A - c)^2}{9} \text{ and} \\ v(\{j, k\}, \pi^{jk}) &:= \left[\frac{A + c_\ell - 2\min\{c_j, c_k\}}{3} \right]^2 = \frac{(A - c)^2}{9}, \end{aligned}$$

while the aggregate payoff to the grand coalition is

$$(20) \quad v(N, \pi^g) := \left[\frac{A - \min\{c_1, c_2, c_3\}}{2} \right]^2 = \frac{(A - c)^2}{4},$$

where again, the last equalities are for the case in which all firms have symmetric cost c . In that special case, note that the intermediate coalition structure *cannot* be viable, because (18) and (19) jointly imply that

$$v(\{j, k\}, \pi^{jk}) = \frac{(A - c)^2}{9} < 2 \frac{(A - c)^2}{16} = v(\{j\}, \pi^s) + v(\{k\}, \pi^s).$$

This peculiar property of intermediate structures was first noted by Salant, Switzer, and Reynolds (1983), who observe that an exogenous merger of firms (here, j and k) can earn lower joint profit in the presence of competitive outsiders (here, i). In our framework, this is one of the steps that permits us to study viable structures; e.g., the viability of the grand coalition. It tells us that any breakaway from the grand coalition must precipitate the singleton structure, so that the grand coalition is indeed viable. For by (18) and (20),

$$v(N, \pi^g) = \frac{(A - c)^2}{4} > 3 \frac{(A - c)^2}{16} = v(\{1\}, \pi^s) + v(\{2\}, \pi^s) + v(\{3\}, \pi^s).$$

These observations are robust to asymmetric bargaining weights, as long as costs are identical. But intermediate structures may become viable with heterogeneous costs. Let

$$A = 10, c_1 = 2, c_2 = 3, \text{ and } c_3 = 4.$$

Then (18) implies that $v(\{1\}, \pi^s) \simeq 7.56$, $v(\{2\}, \pi^s) \simeq 3.06$, and $v(\{3\}, \pi^s) \simeq 0.56$. Likewise, (19) tells us that $v(\{1\}, \pi^{23}) = 9$ and $v(\{2, 3\}, \pi^{23}) = 4$, that $v(\{2\}, \pi^{13}) = 4$ and $v(\{1, 3\}, \pi^{13}) = 9$, and that $v(\{3\}, \pi^{12}) \simeq 1.78$ and $v(\{1, 2\}, \pi^{12}) \simeq 11.11$. Finally, (20) tells us that $v(N, \pi^g) = 16$. These computations show that in contrast to the symmetric case, *every* intermediate structure may now be viable. For instance, $v(\{2, 3\}, \pi^{23}) = 4 > 3.06 + 0.56 = v(\{2\}, \pi^s) + v(\{3\}, \pi^s)$, so $\{2, 3\}$ will not split up under the structure π^{23} . The precise division of the surplus within 23 will, of course, depend on the Nash bargaining weights, but firm 2 will get at least 3.06 and firm 3 at least 0.56. To summarize, using the definition of threats in (13),

$$\Theta(\{1\}, \pi^g) = \{9\} \text{ and } \Theta(\{2, 3\}, \pi^g) \subseteq \{(x_2, x_3) | x_2 \geq 3.06, x_3 \geq 0.56, \text{ and } x_2 + x_3 = 4\}.$$

Similar computations for the other intermediate structures show that

$$\Theta(\{2\}, \pi^g) = \{4\} \text{ and } \Theta(\{1, 3\}, \pi^g) \subseteq \{(x_1, x_3) | x_1 \geq 7.56, x_3 \geq 0.56, \text{ and } x_1 + x_3 = 9\},$$

and

$$\Theta(\{3\}, \pi^g) \simeq \{1.78\} \text{ and } \Theta(\{1, 2\}, \pi^g) \subseteq \{(x_1, x_2) | x_1 \geq 7.56, x_2 \geq 3.06, \text{ and } x_1 + x_2 \simeq 11.11\}.$$

As it so happens, the grand coalition can withstand all these threats, regardless of the bargaining weights. The unblocked set for N consists of any split x of the monopoly profits with $x_1 \geq 9$, $x_2 \geq 4$ and $x_3 \geq 1.78$. So the coarsest viable structure is one in which the grand coalition forms, and deploys its coalitional solution, the exact division depending on bargaining weights. For instance, with equal weights, the unique solution is:

$$\sigma^{\text{sym}}(N) = \{(9, 4, 3)\}.$$

Had the intermediate structures not been viable (as in the equal-cost case), firms 1 and 2 would have received only their Cournot payoffs following any unilateral departure from the grand coalition. It is due to the viability of intermediate coalition structures that they can guarantee themselves more. While the grand coalition might still withstand the threats (as in this example), that is not generally the case, pointing to another important potential difference when costs are heterogeneous. Consider, for instance:

$$A = 23, c_1 = 1, c_2 = 3, c_3 = 5.$$

Again using (18), we have $v(\{1\}, \pi^s) = 49$, $v(\{2\}, \pi^s) = 25$, $v(\{3\}, \pi^s) = 9$. Meanwhile, (19) tells us that $v(\{1\}, \pi^{23}) = 64$ and $v(\{2, 3\}, \pi^{23}) = 36$, that $v(\{2\}, \pi^{13}) = 36$ and $v(\{1, 3\}, \pi^{13}) = 64$, and that $v(\{3\}, \pi^{12}) \simeq 21.78$ and $v(\{1, 2\}, \pi^{12}) \simeq 75.11$. Finally, by (20), $v(N, \pi^g) = 121$. Observe that π^{23} is viable because firms 2 and 3 together earn 36, which can be divided to ensure that 2 gets at least 25 and 3 gets at least 9. It can similarly be shown that the other two intermediate coalition structures are also viable, the precise allocations depending on the Nash weights. So the viability of the grand coalition hinges on its ability to generate some $x = (x_1, x_2, x_3)$ such that $x_1 \geq 64$, $x_2 \geq 36$, and $x_3 \geq 21.78$. These add up to $121.78 > v(N, \pi^g)$, so the grand coalition is not viable. $\diamond$

Example 6. [An R&D Race] We extend Loury (1979) and Reinganum (1982) to incorporate coalition formation in R&D races. Consider a set N of *symmetric* firms with equal bargaining weights, engaged in an R&D race. A formed coalition $S \subseteq N$ works towards a Poisson prize with parameter λ_S given by

$$(21) \quad \lambda_S = \sqrt{2e_S},$$

where e_S is the overall research effort undertaken by S . This effort is derived from synergistic research budgets $\{x_j\}$ of firms in S , as follows:

$$(22) \quad e_S = \left[\sum_{j \in S} x_j^{1/\alpha} \right]^\alpha,$$

where $\alpha \geq 1$ is a measure of synergy. (There is no coordination and so no synergy across firms in different coalitions, and the Poisson arrival rates across such firms are additive.)

The prize has value 1, and has present value $\exp(-r\tau)$ to whoever wins it at random arrival time τ , where r is the common discount rate. But coalition S obtains the prize only if its Poisson arrival happens first. Define λ_{-S} to be the aggregate arrival rate over coalitions other than S (i.e., $\lambda_{-S} := \sum_{W \in \pi - S} \lambda_W$). Then the instantaneous probability of first arrival at date τ in favor of S is $\lambda_S \exp(-\lambda_S \tau) \exp(-\lambda_{-S} \tau) d\tau$, so that the expected value to coalition S is $\int \exp(-r\tau) \lambda_S \exp(-\lambda_S \tau) \exp(-\lambda_{-S} \tau) d\tau = \lambda_S / (r + \lambda_S + \lambda_{-S})$.

If transfers are freely available within a coalition, then by symmetry and the convexity of costs, it is optimal to divide the overall effort equally among firms in S . Let s be the cardinality of S . From (21) and (22), the per-firm budget needed to generate an arrival rate of λ_S is given by $(1/2s^\alpha) \lambda_S^2$, and so the expected aggregate payoff to S is given by:

$$(23) \quad v_S := \frac{\lambda_S}{r + \lambda_S + \lambda_{-S}} - \frac{1}{2s^{\alpha-1}} \lambda_S^2.$$

This is a strictly concave problem in λ_S with an interior solution, so first-order equality conditions are necessary and sufficient for characterizing best responses:

$$(24) \quad \frac{1}{r + \Lambda} - \frac{\lambda_S}{(r + \Lambda)^2} - \frac{1}{s^{\alpha-1}} \lambda_S = 0,$$

where $\Lambda := \lambda_S + \lambda_{-S}$. These conditions easily generate a partition function. Let's specialize now to the case of just three firms.

The grand coalition N of all firms has no rivals. Set $S = N$, $s = 3$ and $\lambda_{-N} := 0$ in (23). Then the first-order condition in (24) reduces to:

$$(25) \quad \frac{1}{3^{\alpha-1}} \lambda_N (r + \lambda_N)^2 = r,$$

which pins down λ_N . Using (23) and (25), the equilibrium payoff to N is given by:

$$(26) \quad v(N, \pi^g) := \frac{\lambda_N}{r + \lambda_N} - \frac{1}{2 \cdot 3^{\alpha-1}} \lambda_N^2, \text{ where } \lambda_N \text{ uniquely solves (25).}$$

Now consider a structure π^{jk} , where $\{j, k\}$ is a coalition and $i \neq j, k$ the residual standalone firm. With $\{j, k\}$ and i maximizing payoff as in (23), the first-order conditions reduce to:

$$(27) \quad r + \lambda_i = \frac{1}{2^{\alpha-1}} \lambda_{jk} (r + \lambda_{jk} + \lambda_i)^2 \text{ (for } \{j, k\}), \text{ and } r + \lambda_{jk} = \lambda_i (r + \lambda_{jk} + \lambda_i)^2 \text{ (for } i).$$

Using (23), the aggregate payoff to coalition $\{j, k\}$ can therefore be written as

$$(28) \quad v(\{j, k\}, \pi^{jk}) = \frac{\lambda_{jk}}{r + \lambda_{jk} + \lambda_i} - \frac{1}{2^\alpha} \lambda_{jk}^2.$$

The heart of this example lies in the following observation: for any $r > 0$, there is $\bar{\alpha} > 1$ such that if $\alpha > \bar{\alpha}$, then the payoff $v(\{j, k\}, \pi^{jk})$ exceeds $2/3$, while $v(\{i\}, \pi^{jk}) < 1/3$.³⁰ If the degree of impatience is small ($r \simeq 0$), it is easy to get a sense of the threshold $\bar{\alpha}$ via some direct calculations. When $r \simeq 0$,

$$(29) \quad v(\{j, k\}, \pi^{jk}) \simeq \frac{2^{(\alpha-1)/2}}{2^{(\alpha-1)/2} + 1} - \frac{2^{(\alpha-1)/2}}{2(2^{(\alpha-1)/2} + 1)^2} > 2/3 \text{ for all } \alpha \geq 4.3.$$

The details of these arguments are in the Supplementary Notes, but intuitively, high synergies enable the doubleton coalition to easily overpower its single opponent. Contrast this with the intermediate structure in Example 5, where the singleton free-rider does well relative to its doubleton opponent. To complete the three-firm example, we note that the singleton structure has a unique equilibrium that's symmetric (see Supplementary Notes). The exact details are unimportant, except to note that for all i, j, k :

$$(30) \quad v(\{i\}, \{i, j, k\}) = v(\{j\}, \{i, j, k\}) = v(\{k\}, \{i, j, k\}) = b, \text{ where } b \in (0, 1/3).$$

So for large enough α , the main threat to the grand consortium therefore comes from a potential doubleton breakaway, yielding the joint payoff as described in (28). In fact, there are *three* such outcomes as we move over all $\{j, k\}$ that can break away, each yielding a total payoff exceeding $2/3$ for α large, by (29).

Note that the grand coalition has no more than a surplus of 1 to play with, and for finite synergy and $r > 0$, it will have strictly less. With equal bargaining weights, the grand coalition attempts to maximize the Nash product $x_1 x_2 x_3$, subject to the constraint that $\max\{x_j, x_k\} \geq v(\{j, k\}, \pi^{jk})/2$ for every j and k . Because this number exceeds $1/3$ for large α , the solution cannot be symmetric, *even if* all the firms are symmetric. The grand coalition survives, though, because superadditivity guarantees that it can pay out

³⁰Indeed, as synergy levels become large ($\alpha \rightarrow \infty$), $v(\{j, k\}, \pi^{jk})$ converges to 1, the maximum possible surplus, while $v(\{i\}, \pi^{jk}) \rightarrow 0$.

$v(\{j, k\}, \pi^{jk})/2$ to at least two firms, thereby guaranteeing that the constraint just described holds. But it must divide its payoff unequally.

The argument above can be shown to extend to any number of firms, because the main threats will all arise from minimal majority coalitions. The argument also survives asymmetric bargaining weights. The doubleton structures are still viable as they were before. However, they divide the value $v(\{j, k\}, \pi^{jk})$ in (28) unequally using their Nash bargaining weights, under the constraint that no individual firm falls below the singleton payoff b in (30). The Supplementary Notes show that an unblocked allocation can still be found for the grand coalition. Unequal treatment of member firms continues to hold, at least for all consistent bargaining weights that maintain their relative values across all coalitions.³¹

The example highlights how the viability of the grand coalition may need to be predicated on treating its symmetric members asymmetrically. Note the difference from Example 7, to which we now turn, where the most powerful threats come from minority coalitions, and cannot be appeased by the grand coalition. $\diamond$

Example 7. [An Environmental Problem] We adapt Ray and Vohra (2001) to the current framework. Consider n symmetric countries engaged in pollution control $(z_1, \dots, z_n)$, with aggregate benefit $\sum_i z_i$ accruing to each of them. For each i , the cost of z_i is borne by country i alone, and is given by $z_i^3/3$. The payoff to country i is therefore

$$(31) \quad \sum_j z_j - \frac{z_i^3}{3}.$$

A coalition can make unlimited transfers among its members. Because externalities are additive, production is uniquely pinned down in every structure: a coalition T of size t maximizes $t \sum_j z_j - \sum_{i \in T} z_i^3/3$, so each of its members produces $z_t = \sqrt{t}$, no matter what the other countries do. As every country is alike, a coalition structure is fully described by its *numerical* form $\mathbf{t} \equiv \langle t_1, \dots, t_m \rangle$; angle brackets record only coalition *sizes*. Aggregate *global* control is then $\sum_k t_k z_{t_k} = \sum_k t_k \sqrt{t_k} = \sum_k t_k^{3/2}$, and (abusing notation slightly) the aggregate payoff $v(t, \mathbf{t})$ to a coalition of size t is

$$(32) \quad v(t, \mathbf{t}) := t \left[\sum_k t_k^{3/2} - \frac{1}{3} t^{3/2} \right] = t \sum_{k=1}^m t_k^{3/2} - \frac{1}{3} t^{5/2}.$$

³¹If bargaining weights are entirely arbitrary across coalitions, then it is possible to have equal division in the grand coalition, though we do not think of this as a realistic possibility.

Two forces are visible in (32). A larger coalition raises aggregate control, which benefits everyone; but it also raises the producer's own burden, at rate $\frac{1}{3}t^{5/2}$. The first force limits free-riding. The second encourages it.

Setting $t = n$ and $m = 1$ in (32), we see that the grand structure $\pi^g = \langle n \rangle$ delivers $2n^{5/2}/3$ in global payoff. It is the unique efficient structure. Whether it is *viable* is another matter. When $n = 2$, it clearly is. For π^g generates a total payoff of $2^{7/2}/3$, whereas the sum of payoffs to the only other structure, π^s , is given by setting $t = 1$ and $m = 2$ in (32), so that

$$v(1, \pi^s) + v(1, \pi^s) = 2 \times (5/3) < 2^{7/2}/3 = v(N, \pi^g).$$

The only point of this simple observation is to reiterate that the usual noncooperative inefficiency does not necessarily translate into inefficient outcomes for the coalitional solution.

Suppose a group of t countries departs. If the remaining $n - t$ hold together, the departing group lands in $\langle n - t, t \rangle$. But that structure need not itself be viable, in which case the deviants must look further ahead.

When $n = 3$, observe that the intermediate structure $\langle 2, 1 \rangle$ is viable. For setting $t = 2$ and $\mathbf{t} = (2, 1)$ in (32) yields a total payoff to the doubleton of $2 + (2^{7/2}/3)$. If they were to separate, then each of their payoffs would be given by $2^3/3$ (set $t = 1$ and $\mathbf{t} = \langle 1, 1, 1 \rangle$ in (32)), for a total of $2^4/3$. Because this is smaller than $2 + (2^{7/2}/3)$, the pair holds together.

That means, in turn, that the lone country in the structure $\langle 2, 1 \rangle$ can count on a payoff of $2\sqrt{2} + (2/3)$ (set $t = 1$ and $\mathbf{t} = (2, 1)$ in (32)). This is easily seen to exceed a third of the grand coalition's payoff of $2 \cdot 3^{3/2}$, found by setting $t = 3$ and $m = 1$ in (32). It follows that no division of the grand coalition surplus can *simultaneously* prevent all three countries from rushing to the exit. The grand structure fails, and the coarsest viable structures are the three doubletons. Notice that the threat that destroys cooperation is an individual one.

Now consider $n = 6$, where matters are different in an instructive way.

Observation 1. *With six symmetric countries, the grand coalition would be viable if doubleton coalitions were artificially prevented to block any allocation. With no such artificial constraint, the grand structure is not viable. All outcomes are Pareto inefficient. Thus doubleton coalitions are instrumental in destroying viability, unlike in the case of $n = 3$.*

Proof. A complete proof, provided in the Supplementary Notes, checks the viability of all numerical structures. Here we shall sketch the main elements of the argument. Given (32), it is easy to compute the per-capita payoff within each coalition of a coalition structure.

Numerical Structure	Per-capita Payoff per Coalition	Viability
$\langle 6 \rangle$	(9.80)	Viable if and only if $\langle 4, 2 \rangle$ is not permitted
$\langle 5, 1 \rangle$	(8.45, 11.85)	Nonviable
$\langle 4, 2 \rangle$	(8.16, 9.89)	Viable
$\langle 4, 1, 1 \rangle$	(7.33, 9.67, 9.67)	Viable

Table 1. Per capita coalitional worths for some structures when $n = 6$.

Table 1 provides the per capita worths for some of the coalition structures (the rest are in the Supplementary Notes).

It can be shown that $\langle 4, 1, 1 \rangle$ and $\langle 4, 2 \rangle$ are both viable, and in each case the unique coalitional solution involves equal division. It is also the case that the maximum per-capita payoff to any coalition in a structure other than the four listed in this table, is less than 9. Given these facts, we can now assess the viability of the grand coalition.

First observe that $\langle 5, 1 \rangle$ is not viable because the five-member block receives only 8.45 per-capita, so no matter how it divides its spoils, at least one of its members gets no more than this amount. By exiting, that country precipitates $\langle 4, 1, 1 \rangle$ and thereby obtains 9.67. Because $\langle 4, 1, 1 \rangle$ is viable, this means that $\langle 5, 1 \rangle$ is not.

Suppose a doubleton coalition is artificially prevented from blocking the grand coalition. Then we claim that the grand coalition is viable (giving each player 9.80). Since no 3, 4 or 5 player coalition can provide each of its members more than 9, they cannot threaten the grand coalition. While a singleton may be tempted to deviate to $\langle 5, 1 \rangle$, we've already seen that this is not viable and will disintegrate further, e.g. to $\langle 4, 1, 1 \rangle$. The best that the original deviator can get by abandoning the grand coalition is 9.67 – in $\langle 4, 1, 1 \rangle$ – which is less than 9.80. Thus, the grand coalition is viable.

Now allow a doubleton to form. Any such coalition can deviate from the grand coalition to the viable structure $\langle 4, 2 \rangle$, with each member assigned 9.89. Can the grand coalition protect itself against this threat? To do so, it must allocate at least 9.89 to at least 5 of its members. Since the per-capita worth of the grand coalition is 9.80, its remaining member must then get no more than 9.35. But then this player can initiate a farsighted objection to $\langle 4, 1, 1 \rangle$. Thus, the grand coalition is viable if and only if $\langle 4, 2 \rangle$ is not permitted to form. $\square$

The contrast with $n = 3$ is worth drawing out. There, cooperation collapses because a single country prefers to stand aside and enjoy the control efforts of the others. Here that

motive is entirely absent — a solitary deviant is *worse off* relative to equal division of grand-coalition payoffs. (Three-, four-, and five-member coalitions fare no better.) But cooperation collapses anyway, because *two* countries acting together can do what neither can do alone. The reason is not that a pair free-rides more effectively; it is that a pair leaves behind a residue of four, which coheres, whereas a lone deviant leaves behind a residue of five, which does not, and is thereby punished by the very fragmentation it sets off. Thus *non-singleton* coalitional threats can play a pivotal role in precipitating inefficiency. $\diamond$

Example 8. [Hospital Networks] A unit mass of patients is differentiated by their probability q of good health, where $q \sim \text{Unif}[0, 1]$. A patient enjoys value 1 in good health, and if he does fall ill, a stochastic value $x \in [0, 1]$ depending on the medical care he receives.

There is a set N of hospitals. The stochastic value that an ill patient receives from getting care at hospital h is given by distribution $F(x) = x^{a(h)}$, where $a(h)$ can be viewed as hospital *quality*. Membership in a hospital *network* (or coalition) $S \subseteq N$ gives the patient access to the best value draw across S : $x(S) := \max_{h \in S} x(h)$, where, if these draws are independent, $x(S)$ has distribution

$$F_S(x) = \prod_{h \in S} x^{a(h)} = x^{a(S)}, \text{ where } a(S) \equiv \sum_{h \in S} a(h).$$

So the expected recovery from joining network S is given by:

$$(33) \quad b(S) := \mathbb{E}[x(S)] = \int_0^1 (1 - x^{a(S)}) dx = \frac{a(S)}{a(S) + 1} \in (0, 1),$$

which we also call “quality”. Of course, $b(S) \leq 1$, the value of perfect health.

All healthcare is produced at zero cost to the hospitals.

Under a network structure π , each $S \in \pi$ posts a price or “premium” $p(S)$. If a patient enrolls in S and falls ill, she is treated at her best in-network match at no further charge, so the expected payoff to a patient of type q enrolled in S (net of the good health outcome) is $(1 - q)b(S) - p(S)$. The patient therefore enrolls in any one network maximizing $(1 - q)b(S) - p(S)$, provided that maximum is nonnegative; otherwise she stays out and receives no care.

Arrange networks in any partition in descending order of quality: $b_1 \geq b_2 \geq \dots \geq b_m$. Let K be the smallest integer such that $b_{K+1} = b_{K+2}$ (with $K = m$ if $b_i \neq b_j$ for all $i \neq j$, or $K = 0$ if $b_1 = b_2$). The Supplementary Notes show that in any price-setting equilibrium, every network k from 1 to K enjoys positive profits at premium p_k^* , while from $j = K + 1$

onward, the premium drops to zero and no profits are made. The premium p_k^* for network k is pinned by the adjacent prices p_{k-1}^* and p_{k+1}^* (and qualities b_{k-1}, b_k, b_{k+1}). Initial and final conditions for networks 1 and K then determine the pricing chain. Patient sorting is assortative, with network k serving types in an interval $[q_{k-1}^*, q_k^*]$, where $q_0^* := 0$.

In this way, a unique value is generated for each network in the partition, so that a partition function is well-defined. For any $\pi = (S_1, \dots, S_m)$, arranged so that $b(S_k) \geq b(S_{k+1})$ for all k , the *partition function* is the resulting profit of each network:

$$(34) \quad v(S_k, \pi) := p_k^*[q_k^* - q_{k-1}^*]$$

if $k \leq K$, and $v(S_k, \pi) = 0$ for $k = K + 1, \dots, m$ (if any such index exists). The argument is constructive, so a computer can easily find accurate solutions for any partition.

With just the “grand” network, the solution is that of a standard monopoly with linear demand, with half the market served at a price $b(N)/2$, so $v(N, \pi^g) = b(N)/4$ (to be divided among the hospitals in the network). With $\pi = \{S_1, S_2\}$ and $b_1 > b_2$, equilibrium prices (p_1^*, p_2^*) and allocations $[0, q_1^*]$ and $[q_1^*, q_2^*]$ pin down the partition function:

$$(35) \quad v(S_1, \pi) = p_1^* q_1^* = \frac{4b_1^2(b_1 - b_2)}{(4b_1 - b_2)^2} \quad \text{and} \quad v(S_2, \pi) = p_2^*(q_2^* - q_1^*) = \frac{b_1 b_2(b_1 - b_2)}{(4b_1 - b_2)^2}.$$

Again, these values must be allocated to network members. With three or more networks, closed-form solutions are complicated but can be written down; see Supplementary Notes.

Of all possible network structures, it can be shown that the grand network generates the largest total profit; see the Supplementary Notes for detailed discussion. For *patients*, though, the story is more nuanced. Network fragmentation breeds *competition*: Bertrand-driven lower premia and wider coverage, which raises consumer surplus. Consolidation enhances *quality*, which also raises consumer surplus. Therefore network structures that maximize consumer surplus are typically of intermediate size. If it is at all possible, a “balanced” structure of just two equal-quality blocks $b_1 = b_2$ would aggregate to the highest quality possible — and then, being tied, generate fierce price competition. In short, patients want hospitals pooled enough to be good, and split enough to compete.

But hospitals that form networks have their own goals. Consolidation is not just access to a widespread array of in-network providers, but often an economic transfer from patients to networks conducted under the guise of “insurance”. We’ve already noted that the grand network π^g is the “efficient solution” from the perspective of the hospitals. Is it viable? It turns out that in a world with three hospitals, it always is.

Observation 2. When $|N| = 3$, the grand network is viable for any profile of qualities.

We relegate the (computation-intensive) proof of this Observation to the Supplementary Notes. But some salient features of the argument are worth remarking on. The coalitional solution, designed to handle a variety of threats, is important for this result. If, for instance, network worth were to be equally divided among the participating hospitals, the grand coalition would generally fail to be viable. Under the coalitional solution, within-network outcomes remain egalitarian, *but only as far as the threats permit*. When the threats are too strong, equal treatment itself must give way.

What that solution looks like depends on which threat binds. There are two possibilities. There is, first, a unilateral threat from a dominant hospital. For instance, when $a(1)$ is large relative to $a(2) \approx a(3)$, the equal-quality weak hospitals earn zero in the singleton structure, but positive profits if they stay together, so the binding deviation is hospital 1's exit to π^{23} . The solution is unique — hospital 1 is pinned at its outside value $v(\{1\}, \pi^{23})$, with the weak pair splitting the rest equally.

The second kind of potential threat is a bilateral one from two comparable leaders. When $a(1) \approx a(2) > a(3)$, the singleton structure is again worthless. The threat that binds is secession by the pair $\{1, 2\}$, leading to the structure π^{12} . The resulting value to the two leaders, given by $v(\{1, 2\}, \pi^{12})$, is split equally. Deterring this threat “in the most egalitarian way possible” requires only *one* leader to be bought off, since the pair blocks only when *both* gain. In fact, the maximization of the Nash product over the resulting nonconvex set appeases one leader and drives the other down to the weak hospital's payoff. The solution is two-valued, and which leader is favored is indeterminate.

With three hospitals, then, the solution always selects the grand network, which is the structure best for the hospitals and far from best for patients. The grand network that forms gives patients a consumer surplus of $b(N)/8$ — at $a = 1$, a net value of $\simeq 0.094$, against the 0.250 of the singleton tie they would prefer. The grip of the grand network is not unbreakable, though: with enough symmetric hospitals it fails. What follows explores this possibility and, returning to the welfare theme, asks what that failure hands back to the patients.

Consider the case of n symmetric hospitals with $a(h) = 1$ for each of them. Each partition can now be identified with a numerical coalition structure of the form $\langle n_1, \dots, n_m \rangle$. A block of k hospitals has quality $b(k) = k/(1 + k)$, and division within a network follows the coalitional solution. Viable structures arise by examining the threats associated with

Structures with one singleton		Structures with two singletons	
Structure	Max	Structure	Max
$\langle 10, 2, 1 \rangle$	7.73	$\langle 11, 1, 1 \rangle$	9.47
$\langle 9, 3, 1 \rangle$	6.00	$\langle 9, 2, 1, 1 \rangle$	8.08
$\langle 8, 4, 1 \rangle$	4.34	$\langle 8, 3, 1, 1 \rangle$	6.16
$\langle 7, 5, 1 \rangle$	2.48	$\langle 7, 4, 1, 1 \rangle$	4.19
$\langle 7, 3, 2, 1 \rangle$	5.66	$\langle 6, 5, 1, 1 \rangle$	1.69
$\langle 6, 4, 2, 1 \rangle$	3.57	$\langle 6, 3, 2, 1, 1 \rangle$	5.77
$\langle 5, 4, 3, 1 \rangle$	2.39	$\langle 5, 4, 2, 1, 1 \rangle$	2.64

Table 2. Potential partitions (and their per-capita payoffs $\times 1000$) that can arise from $\langle 12, 1 \rangle$.

farsighted disintegration, a threat being credible if every member of the breakaway gains strictly at the farsightedly anticipated structure.

Observation 3. *If $n \leq 12$ the grand network is viable. At $n = 13$, it is not.*

While this Observation can be established by the brute-force expedient of applying our solution to all conceivable partitions, the number of cases is reduced by a simple claim. From any partition, a hospital earning positive profits can never *deliberately* deviate to create a tie among two networks, while remaining in one of them. (One-line proof: each member of a deliberate breakaway must gain strictly, but a tie generates zero profits.) That implies that a relatively small number of partitions need to be checked for threats. For instance, if the starting partition is $\langle 10 \rangle$, the only credible threats can come from at most nine numerical partitions: $\langle 9, 1 \rangle$, $\langle 8, 2 \rangle$, $\langle 7, 3 \rangle$, $\langle 7, 2, 1 \rangle$, $\langle 6, 4 \rangle$, $\langle 6, 3, 1 \rangle$, $\langle 5, 4, 1 \rangle$, $\langle 5, 3, 2 \rangle$, and $\langle 4, 3, 2, 1 \rangle$. And if $\pi = \langle 12, 1 \rangle$ (a case of particular interest; see below), an anticipated partition from here cannot have any numerical ties apart from just one two-way tie of singletons. The only credible threats can come from at most 14 numerical partitions; see Table 2.

The proof of Observation 3, then, first evaluates every “credible” partition that can arise from the grand network $\langle n \rangle$, for $2 \leq n \leq 13$; see Supplementary Notes. For every $n \leq 12$, no per-capita payoff in these partitions exceeds the per-capita payoff to the grand coalition. At $n = 13$, these calculations fail for the first time. A lone hospital leaving $\langle 13 \rangle$ induces $\langle 12, 1 \rangle$, and at that partition, it collects a value that exceeds its equal-share payoff from the grand network. Assume for the moment that $\langle 12, 1 \rangle$ is viable, and let us complete the

argument. By symmetry, *every* hospital has the same deviation. No division of the grand network payoff can deter every hospital, so the grand network is not viable.

The proof is completed by showing that $\langle 12, 1 \rangle$ is indeed viable. As already noted, there are precisely 14 numerical sub-partitions of $\langle 12, 1 \rangle$ that can be credibly reached by farsighted deviations from $\langle 12, 1 \rangle$. Table 2 collects the per-capita payoffs ($\times 1000$ for easier reading) from each of them. It shows that a credible coalitional deviation from $\langle 12, 1 \rangle$ can generate *at most* a per-capita payoff of 9.47, while the per-capita payoff to the hospital block of 12 is given by 11.79, using the duopoly formula in equation (35).³² Therefore $\langle 12, 1 \rangle$ is viable, and the argument is complete.

Observe how the break at $n = 13$ is set off by a smidgen of self-interest. Under $\langle 12, 1 \rangle$, the lone hospital lifts its own average payoff from 17.86 to 19.16, a gain of around 7.3%. That relatively small gain suffices to detonate the grand coalition. Each of the twelve left behind experiences a fall in payoffs from 17.86 to 11.79, a more substantial loss of around 34%, with overall industry profit dropping by approximately 31%. Yet *consumers*, bystanders throughout, are the windfall winners: patient-consumer surplus leaps from 116.1 to 258.9, a gain of 123%. In absolute terms, this gain is larger than the entire profit forfeited by the hospitals, and additive social welfare goes up by more than 20%.

The lesson is not that the coalitional solution shields patients — it does nothing of the kind; see the case of three hospitals. Relief comes only as a *byproduct* of intra-cartel free-riding, and only once that free-riding clears the threshold as it does at $n = 13$; not for instance at $n = 12$, where no defector exists and patients stay pinned at their consumer surplus floor of $b(N)/8$. At $n = 13$, the deviating hospital serendipitously does the antitrust authority's work — though only halfway to be fair, for $\langle 12, 1 \rangle$ still leaves consumers well below their own optimum, which is the symmetric tie $\langle 6, 6, 1 \rangle$, at $1000 \cdot b(6)/2 \simeq 428.6$. Consolidation is policed here not by any guardian, but by the occasional greed of a subset of hospitals.

It may be worth recording where this example stands relative to the literature on hospital networks; in particular Ho and Lee (2017, 2019). As in Ho–Lee, patients enroll *ex ante* at posted premiums; point-of-care prices are (near) zero, so a sick patient's choice of hospital turns on match value rather than price. The value of a hospital network (or coalition) is the best realized match that its member-hospitals can provide to a patient who falls sick. And consumers are heterogeneous in health risk. But there are distinct points of difference.

³²Under $\{S_1, S_2\}$ with $|S_1| = 12$ and $|S_2| = 1$, $b(S_1) = 12/13$ and $b(S_2) = 1/2$, so that per-capita payoff ($\times 1000$) to the 12-block is $\frac{1000v(S_1, \{S_1, S_2\})}{12} = \frac{4000b_1^2(b_1 - b_2)}{12(4b_1 - b_2)^2} \simeq 11.79$.

Ho and Lee (2017) study competition among insurers, *taking provider networks as given*, and ask how that competition shapes negotiated prices and consumer welfare. Ho and Lee (2019) endogenize the network, studying the inclusion decision of a strategic insurer — one that can threaten to replace a hospital — holding rival networks fixed. Contracts are bilateral, offered by the network owner or insurer to hospitals. So a hospital network is really a collection of bilateral agreements between a single insurer and one or more hospitals, with the insurer collecting premia and reimbursing the hospitals in the network.

In our application, the *entire* network structure is the object of analysis. Hospitals (or “insurers” who play the role of coordinator or facilitator) form networks with possibly multi-lateral in-network agreements, and receive premia directly from patients. A viable structure must survive all credible subcoalitional deviations that threaten to form new networks, with threats evaluated consistently and farsightedly. Whether the grand network forms, which fragmented structures are stable, and who earns what within a network are the questions the theory is built to answer. Ho and Lee (2019) pose the first of these, but with bilateral contracting and a single insurer’s inclusion decision; what lies beyond that framework is stability against *coalitional* deviations, in which several hospitals move together and evaluate their prospects farsightedly. ◇

5. Summary and Research Directions

We study a *consistent coalitional solution* in the spirit of Nash’s seminal work on two-person bargaining. We begin with characteristic functions. Our solution is recursively defined for all coalitions and subcoalitions. Throughout, the solution has a dual personality: it serves both as an end-goal *for* the coalition in question, and as a threat *by* the coalition that applies to and constrains the outcome of coarser coalitions — all the way “up” to the grand coalition.

As in Nash bargaining, our solution maximizes the product of payoffs net of individual disagreements, with the difference that coalitional threats appear as conventional constraints, and are not netted out from payoffs. This leads to a philosophical interpretation of our coalitional solution based on a “pragmatic ethic” that’s quite distinct from “split-the-surplus”: move as close as possible to equality in payoffs, unless prevented from doing so by a potential breakaway threat.

Under our coalitional solution, the grand coalition is viable in every TU superadditive game with consistent bargaining weights, and therefore has an efficient outcome. This finding stands in sharp contrast to the well-known Bondareva–Shapley observation that the core of a superadditive game is generally empty, unless significantly stronger assumptions are made.

We then go beyond characteristic functions, by embedding our solution into a broader setting where each coalition obtains payoffs that could depend on the entire coalition structure. This dependence arises from noncooperative interaction across coalitions in the structure, and we extend our recursive definition to take account of such interaction. We illustrate the extended solution, and the associated concept of the *viability of a coalition structure*, with four different applications: the production of public goods with global externalities, competition in R&D, Cournot oligopoly, and hospital networks.

It should be emphasized that we have worked throughout with the case of “internal blocking”, in which every coalition anticipates further deviations by its *subcoalitions*, but does not entertain other deviations engineered by coalitions that intersect the coalition in question but are not necessarily subsets of it. This extension will require that we replace our recursive definition of internal consistency by a consistency notion based on a fixed point argument, or by a set of rules that prevents the same coalition structure from forming twice in any allowable blocking chain. This is the subject of research to come, and the analysis here will serve as a base for such research. The Supplementary Notes contain further remarks on this topic.

Appendix A. Proofs and Some Technical Remarks

By [Inv], there is no loss of generality in assuming that for every $\mathcal{G} = (\mathcal{F}, d)$, $d = 0$. This normalization will be in force throughout this Appendix. Our first step combines [Exp], [Inv] and [UHC] to show that a solution to any two-allocation problem for coalition S is also a solution to the same two-allocation problem when all subcoalitions of S are ineffective. For some given S and game $\mathcal{G}$, denote by $\mathcal{G}^0$ the game in which all subcoalitions of S are ineffective, with no other changes.

It will ease the exposition to suppose (with no further explicit qualification) that every game $\mathcal{G}$ we consider has S as its grand coalition. By virtue of the consistency condition (3), all our results will trivially extend to games in which S is embedded as a subcoalition.

Lemma 1. *Let σ satisfy [Inv], [UHC] and [Exp]. Suppose that $F(S) = \{x, y\} = U^\sigma(S, \mathcal{G})$ for $x \gg 0$ and $y \gg 0$, where $U^\sigma(S, \mathcal{G})$ is recursively defined as in (2). Then:*

$$(36) \quad \text{if } x \in \sigma(S, \mathcal{G}), \text{ then } x \in \sigma(S, \mathcal{G}^0);$$

that is, x is also a solution to the pure bargaining problem with feasible set $F = \{x, y\}$.

Proof. Consider any game $\mathcal{G}$, with $F(S) = \{x, y\} = U^\sigma(S, \mathcal{G})$ for some $x \gg 0$ and $y \gg 0$. Suppose that $x \in \sigma(S, \mathcal{G})$. For $\mu \in \mathbb{R}^S$ with $1 \gg \mu \geq 0$, say that $\mathcal{G}''$ is a *shrinking* of $\mathcal{G}$ (by factor μ) if $F''(S) = F(S)$, and $F''(T) = \mu_T \odot F(T)$ for every $T \subset S$. Say that $\mathcal{G}'' \prec \mathcal{G}$ if $\mathcal{G}''$ is a shrinking of $\mathcal{G}$ with the added property that $x \in \sigma(S, \mathcal{G}'')$.

Fix $\mathcal{G}$ as in the statement of the Lemma. Let $\mathcal{T}$ be the collection of all games $\mathcal{G}''$ such that $\mathcal{G}'' \prec \mathcal{G}$. We claim that $\mathcal{G}^0 \in \mathcal{T}$, and in particular that $x \in \sigma(S, \mathcal{G}^0)$.

Given [Exp] and [Inv], it follows from Remark 7 that $\mathcal{T}$ is nonempty. Note that $\mathcal{T}$ is partially ordered by $\prec$, which is obviously transitive. For any totally ordered subset $\mathcal{T}^*$ of $\mathcal{T}$, define:

$$m^* := \inf\{\|\mu\| \mid \text{there is some } \mathcal{G}'' \in \mathcal{T}^* \text{ which is a shrinking of } \mathcal{G} \text{ by factor } \mu\}.$$

If there is some $\mathcal{G}^* \in \mathcal{T}^*$ with associated μ satisfying $\|\mu\| = m^*$, then clearly $\mathcal{G}^* \prec \mathcal{G}''$ for every $\mathcal{G}'' \in \mathcal{T}^*$, and so $\mathcal{G}^*$ is a lower bound for $\mathcal{T}^*$. Otherwise, by the definition of m^* , there is a sequence $\{\mathcal{G}^k\}$ in $\mathcal{T}^*$, where $\mathcal{G}^k$ is a shrinking of $\mathcal{G}$ for some factor μ^k with $\|\mu^k\| \rightarrow m^*$. Let $(\mathcal{G}^*, \mu^*)$ be any limit point of $\{(\mathcal{G}^k, \mu^k)\}$, the first component a limit under the product Hausdorff metric and the second in the standard sense.³³ Because $x \in \sigma(S, \mathcal{G}^k)$ for every k , and $U^\sigma(S, \mathcal{G}^k) = \{x, y\} = U^\sigma(S, \mathcal{G}^*)$ throughout (so that $U^\sigma(S, \mathcal{G}^k) \rightarrow U^\sigma(S, \mathcal{G}^*)$ trivially),³⁴ it follows from [UHC] that $x \in \sigma(S, \mathcal{G}^*)$. Therefore $\mathcal{G}^* \in \mathcal{T}$, and just as in the previous case, $\mathcal{G}^* \prec \mathcal{G}''$ for every $\mathcal{G}'' \in \mathcal{T}^*$ and so serves as a lower bound for $\mathcal{T}^*$ in $\mathcal{T}$.

So we can apply Zorn's Lemma to assert that $\mathcal{T}$ admits a minimal element $\underline{\mathcal{G}}$. We claim that $\underline{\mathcal{G}} = \mathcal{G}^0$; that is, all subcoalitions of S are ineffective under $\underline{\mathcal{G}}$. Otherwise, if some such subcoalition is effective, by Remark 7, $\underline{\mathcal{G}}$ could be shrunk further, with both x and y remaining unblocked, and with x remaining a solution, which violates minimality. We

³³These are well-defined given that $0 \leq \mu^k \leq 1$ for every k . Moreover, because $\mathcal{T}^*$ is totally ordered, the sequence $\{\mathcal{G}^k\}$ can be chosen so that each $\mathcal{G}^{k+1}$ is a shrinking of $\mathcal{G}^k$; the factors μ^k are then coordinatewise non-increasing, so that $\mu^* \leq \mu^1 \leq 1$, and $\mathcal{G}^*$ is a bona fide shrinking of $\mathcal{G}$.

³⁴In any shrinking $\mathcal{G}''$ of $\mathcal{G}$, both x and y remain unblocked, so that $U^\sigma(S, \mathcal{G}') = F(S) = \{x, y\}$. If $\mu_T \gg 0$, (3) and [Inv] imply $\sigma(T, \mathcal{G}'') = \mu_T \odot \sigma(T, \mathcal{G})$, and no scaled-down threat can strictly dominate x_T or y_T when the original did not; and if $\mu_i = 0$ for some $i \in T$, every threat available to T pays i zero, while $x, y \gg 0$.

have therefore proved the claim. And as an implication of the claim, we have established (36). $\square$

Fix a coalition S , our solution σ , and the collection $\mathcal{K}$ of feasible sets $F(T)$ for every $T \subset S$. We will be moving $F(S)$ around for given $\mathcal{K}$, so it will be useful to define $U^\mathcal{K}$ as the set of all allocations in $\mathbb{R}_+^S$ unblocked by $\mathcal{K}$ (using σ). That is, $F(S) \cap U^\mathcal{K}$ is the usual unblocked set for S once $F(S)$ is specified. We abuse notation a bit and write $\sigma(F(S), \mathcal{K}) = \sigma(S, \mathcal{G})$ for the solution applied to S . Our next step follows Kaneko (1980).

Lemma 2. *Define a binary relation $\succsim^\mathcal{K}$ on $U^\mathcal{K}$ by $x \succsim^\mathcal{K} y$ if and only if $x \in \sigma(\{x, y\}, \mathcal{K})$. Then, if σ satisfies [PO], [IIA], and [UHC],*

$$(37) \quad x \gg y \text{ implies } \sigma(\{x, y\}, \mathcal{K}) = \{x\},$$

and $\succsim^\mathcal{K}$ is complete, transitive, and continuous on $U^\mathcal{K}$.

Proof. [PO] immediately implies (37). Completeness is a consequence of the fact that $\sigma(\{x, y\}, \mathcal{K}) \neq \emptyset$. For transitivity, we redo the steps in Kaneko (1980).³⁵ To this end, consider x_1, x_2, x_3 , all in $U^\mathcal{K}$, with $x_1 \in \sigma(\{x_1, x_2\}, \mathcal{K})$ and $x_2 \in \sigma(\{x_2, x_3\}, \mathcal{K})$. We must prove that $x_1 \in \sigma(\{x_1, x_3\}, \mathcal{K})$.

Because $\sigma(\{x_1, x_2, x_3\}, \mathcal{K}) \neq \emptyset$, $x_i \in \sigma(\{x_1, x_2, x_3\}, \mathcal{K})$ for some i . Suppose that $i \neq 1$. By [IIA], $x_i \in \sigma(\{x_{i-1}, x_i\}, \mathcal{K})$ and because we already have $x_{i-1} \in \sigma(\{x_{i-1}, x_i\}, \mathcal{K})$, it must be that $\sigma(\{x_{i-1}, x_i\}, \mathcal{K}) = \{x_{i-1}, x_i\}$. At the same time, by [IIA], $\sigma(\{x_{i-1}, x_i\}, \mathcal{K}) = \{x_{i-1}, x_i\} \cap \sigma(\{x_1, x_2, x_3\}, \mathcal{K})$, which implies that $x_{i-1} \in \sigma(\{x_1, x_2, x_3\}, \mathcal{K})$. Recursing back, we must conclude that $x_1 \in \sigma(\{x_1, x_2, x_3\}, \mathcal{K})$.

But then by [IIA] once more, $x_1 \in \sigma(\{x_1, x_3\}, \mathcal{K})$ and we are done.

To establish continuity, suppose that for some pair of sequences of allocations $x^m \rightarrow x$ and $y^m \rightarrow y$, all in $U^\mathcal{K}$, $x^m \succsim^\mathcal{K} y^m$ for all m . Then $x^m \in \sigma(\{x^m, y^m\}, \mathcal{K})$ for every m , and by [UHC], it follows that $x \in \sigma(\{x, y\}, \mathcal{K})$, or $x \succsim y$. $\square$

For the special case of a pure bargaining game, remove the parameter $\mathcal{K}$, and write $\sigma(\{x, y\})$ and $\succsim$ instead. Note that in this case, $U^\mathcal{K}$ is all of $\mathbb{R}_+^S$.

³⁵The reason this needs to be redone is that the arguments in Kaneko (1980), Lemmas 2 and 3, are made for pure bargaining games with strictly positive allocations.

Lemma 3. *If σ satisfies [PO], [Inv], [IIA], and [UHC], then for every pure bargaining game $\mathcal{G}$ with $F(S) = \{x, y\}$ for $x, y \gg 0$,*

$$(38) \quad \sigma(\{x, y\}) = \arg \max_{z \in \{x, y\}} \prod_{j \in S} z_j^{\gamma_j(S)}$$

for some $\gamma(S) > 0$ that is independent of $\mathcal{G}$.

Proof. We follow the proof of Theorem 2 in Kaneko (1980) with some adaptations. By Lemma 2, $\succsim$ is a continuous weak order and (37) holds, so Proposition 3.C.1 in Mas-Colell, Whinston, and Green (1995) implies that $\succsim$ can be represented by a continuous function $H : \mathbb{R}_{++}^S \rightarrow \mathbb{R}$; that is, $H(x) \geq H(y)$ if and only if $x \succsim y$. (Kaneko establishes the same representation under a stronger Pareto optimality restriction that we do not assume here.) Next, we follow Kaneko (1980) again, who uses [Inv] to apply a lemma of Osborne (1976), and shows that H can be taken to be of the form

$$(39) \quad H(x) = \prod_{j \in S} x_j^{\gamma_j(S)},$$

for some $\gamma(S) \in \mathbb{R}_+^S$. [PO] additionally (and immediately) implies that $\gamma(S) > 0$. (38) now follows because $x \in \sigma(\{x, y\})$ if and only if $x \succsim y$, if and only if $H(x) \geq H(y)$. $\square$

Proof of Theorem 1. It is straightforward that the coalitional solution satisfies [PO], [Inv], [IIA], and [Exp]. It follows from the argument used in the proof of Theorem 2 in Kaneko (1980) that our solution also satisfies [UHC]. We now prove the converse.

Lemma 3 implies that there exists $\gamma(S) > 0$ such that for any pure bargaining game $\mathcal{G}$ with S as its grand coalition and with $F(S) = \{x, y\}$:

$$\text{If } x, y \gg 0 \text{ and } x \in \sigma(\{x, y\}), \text{ then } \prod_{j \in S} x_j^{\gamma_j(S)} \geq \prod_{j \in S} y_j^{\gamma_j(S)}.$$

Recalling (36) from Lemma 1, this implies that even if $\mathcal{G}$ is not a pure bargaining game, but $U^\sigma(S, \mathcal{G})$ has been recursively defined as in (2) and $F(S) = U^\sigma(S, \mathcal{G}) = \{x, y\}$:

$$(40) \quad \text{If } x, y \gg 0 \text{ and } x \in \sigma(S, \mathcal{G}), \text{ then } H(x) \geq H(y).$$

We now extend (40) to allow for allocations on the boundary $\partial \mathbb{R}_+^S$ of $\mathbb{R}_+^S$. To this end, note that $H(x) = \prod_{j \in S} x_j^{\gamma_j(S)}$ is well-defined and continuous on all of $\mathbb{R}_+^S$. We claim that for any $x, y \in \mathbb{R}_+^S$,

$$(41) \quad \text{if } x, y \text{ are each unblocked and } x \in \sigma(\{x, y\}, \mathcal{K}), \text{ then } H(x) \geq H(y).$$

Suppose, on the contrary, that $x \in \sigma(\{x, y\}, \mathcal{K})$ but $H(y) > H(x)$. By the continuity of H , there is $z \gg x$ such that $H(z) < H(y)$. If the feasible set is $\{x, z\}$, then x is unblocked and so is $z \gg x$, and by [PO], $\sigma(\{x, z\}, \mathcal{K}) = \{z\}$.

If $y \in \sigma(\{y, z\}, \mathcal{K})$, we would then have $x \succsim^{\mathcal{K}} y$ and $y \succsim^{\mathcal{K}} z$, so that the transitivity of $\succsim^{\mathcal{K}}$ (Lemma 2) yields $x \in \sigma(\{x, z\}, \mathcal{K})$, a contradiction. Therefore it must be that

$$(42) \quad y \notin \sigma(\{y, z\}, \mathcal{K}).$$

At the same time, for $\epsilon > 0$, define $y^\epsilon := y + \epsilon \mathbf{1} \gg 0$. Each y^ϵ is unblocked (because y is), and $H(y^\epsilon) \geq H(y) > H(z)$. Because $z \gg 0$ and $y^\epsilon \gg 0$, (40) and the nonemptiness of solutions together imply that $\sigma(\{z, y^\epsilon\}, \mathcal{K}) = \{y^\epsilon\}$ for every $\epsilon > 0$. As $\epsilon \rightarrow 0$, the corresponding games converge, their unblocked sets converge as well, and $y^\epsilon \rightarrow y$. By [UHC], $y \in \sigma(\{y, z\}, \mathcal{K})$, which contradicts (42). That establishes (41).

To see that this characterization extends beyond two-point sets, consider a game $\mathcal{G}$ with S as its grand coalition, and with $x \in \sigma(S, \mathcal{G})$. Take any $y \in U^\sigma(S, \mathcal{G})$. [IIA] implies that if we shrink $F(S)$ to just the two-point set $\{x, y\}$, then $x \in \sigma(\{x, y\}, \mathcal{K})$, where $\mathcal{K}$ stands for the collection of feasible sets of $\mathcal{G}$ restricted to subcoalitions of S . As this argument holds for any $y \in U^\sigma(S, \mathcal{G})$, it follows from (41) that

$$(43) \quad \sigma(S, \mathcal{G}) \subseteq \arg \max_{z \in U^\sigma(S, \mathcal{G})} \prod_{j \in S} z_j^{\gamma_j(S)} \equiv D.$$

To show that (43) holds with equality, assume that $U^\sigma(S, \mathcal{G}) \neq \emptyset$, so that $\sigma(S, \mathcal{G}) \neq \emptyset$ as well. Pick any $x' \in D$ and $x \in \sigma(S, \mathcal{G})$; because $x \in \sigma(S, \mathcal{G}) \subseteq D$, we have $H(x') = H(x)$. If $x' = x$ there is nothing to prove, so suppose $x' \neq x$. We claim that $x' \in \sigma(\{x, x'\}, \mathcal{K})$. For any positive integer m , consider the game $\mathcal{G}^m$ with $F^m(S) = \{x, x' + (1/m)\mathbf{1}\}$ (with $\mathcal{K}$ unchanged). Since both allocations are unblocked, from (43) applied to $\mathcal{G}^m$, it follows that $\sigma(S, \mathcal{G}^m) = \{x' + (1/m)\mathbf{1}\}$. Now let $m \rightarrow \infty$ and note that $\mathcal{G}^m$ converges to $\mathcal{G}^\infty$ with feasible set $\{x, x'\}$, and by [UHC], $x' \in \sigma(\{x, x'\}, \mathcal{K})$, proving the claim. Finally, by [IIA], $\sigma(\{x, x'\}, \mathcal{K}) = \sigma(S, \mathcal{G}) \cap \{x, x'\}$, and therefore $x' \in \sigma(S, \mathcal{G})$.

As noted at the start of this Appendix, all the arguments above apply to games for which S is a subcoalition and not only the grand coalition. By condition (3), the vector of bargaining weights $\gamma(S)$ used to find $\sigma(S, \mathcal{G})$ is the same over all such games. $\square$

Proof of Corollary 1. It is straightforward to check that if σ^{sym} is given by (6), then it satisfies [PO], [Inv], [IIA], [UHC], [Exp] and [Sym].

Conversely, suppose that σ^{sym} satisfies all the axioms. By Theorem 1, there is $\gamma(S) > 0$, independent of $\mathcal{G}$, such that

$$\sigma^{\text{sym}}(S, \mathcal{G}) = \arg \max_{x \in U^\sigma(S, \mathcal{G})} \prod_{j \in S} x_j^{\gamma_j(S)}$$

Fix S and consider the game in which every $T \subset S$ is ineffective and $F(S) = \{e^k\}_{k \in S}$, where $e_k^k = 2$ and $e_j^k = 1$ for all $j \neq k$. The Nash product at e^k equals $2^{\gamma_k(S)}$, so the characterization implies

$$\sigma^{\text{sym}}(S, \mathcal{G}) = \{e^k \mid \gamma_k(S) \geq \gamma_m(S) \text{ for every } m \in S\}.$$

By [Sym], $\sigma^{\text{sym}}(S, \mathcal{G}) = F(S)$, and so $\gamma_k(S) = \gamma_m(S) > 0$, for all $k, m \in S$ (also using $\gamma(S) > 0$). Without loss of generality, we can set $\gamma(S) = 1$. $\square$

Remark 14. Although we have relied on Kaneko (1980) in proving Lemma 3, the literature on Nash bargaining with nonconvex feasible sets³⁶ allows for other options, provided some form of symmetry is assumed. Particularly relevant in our context is the approach in Mariotti (1998), which sharpens Kaneko's characterization to drop [UHC]. We could use Mariotti's approach to directly prove Corollary 1 as an independent theorem. However, unlike Mariotti, there is no getting away from [UHC] for the coalitional case. [UHC] cannot be dropped from Theorem 1, as we show in the Supplementary Notes.

Proof of Theorem 2. Our argument combines Claims 1–3 below.

Claim 1. For any coalition S , $U^\sigma(S, \mathcal{G}) \subseteq U^{\text{naive}}(S, \mathcal{G})$.

Proof. Suppose $x \in U^\sigma(S, \mathcal{G})$ but $x \notin U^{\text{naive}}(S, \mathcal{G})$. Then there is $T \subset S$ and $y \in \Psi(T)$ such that $y \gg x_T$. If $y \in \sigma(T, \mathcal{G})$, that would contradict $x \in U^\sigma(S, \mathcal{G})$. If $y \notin \sigma(T, \mathcal{G})$, then there is $W \subset T$ and $w \in \sigma(W, \mathcal{G})$ with $w \gg y_W$. But then (W, w) also blocks x , and we again have a contradiction to the supposition that $x \in U^\sigma(S, \mathcal{G})$. $\square$

Claim 2. Suppose $T \subset S$ and there exists $x \in \sigma(T, \mathcal{G})$ such that $x \notin \Psi(T)$. Then there exist $W \subset T$ and $w \in \Psi(W)$ such that $w \geq x_W$.

Proof. We proceed by induction on coalition size. The assertion is trivially true for all singleton coalitions T , for then $\sigma(T, \mathcal{G}) = \Psi(T)$, so the antecedent does not hold.

Now suppose that the assertion is true for any coalition T of size k or less, where $k \geq 1$. Let T have cardinality $k + 1$. Suppose that $x \in \sigma(T, \mathcal{G})$ and $x \notin \Psi(T)$. Observe that the Nash

³⁶See Xu and Yoshihara (2020) and Thomson (2022).

product value of $\Psi(T)$ strictly exceeds that of x . By the strict quasiconcavity of the Nash product on the dimensions where player weights are strictly positive, the Nash product value of any allocation of the form $tx + (1 - t)y$, where $y \in \Psi(T)$ and $t \in (0, 1)$, must in turn strictly exceed that of x . These allocations all lie in $F(T)$ because $F(T)$ is convex, so each of them is blocked by some subcoalition W^t of T using some element $z^t \in \sigma(W^t, \mathcal{G})$. In particular, by taking t to 1, we see that there exist $W \subset T$ and $z \in \sigma(W, \mathcal{G})$ such that

$$(44) \quad z \geq x_W.$$

If $z \in \Psi(W)$, we are done. If $z \notin \Psi(W)$, then noting that $|W| \leq k$, the induction hypothesis implies that there is $W' \subset W$ and $z' \in \Psi(W')$ with $z' \geq z_W$. Combining this information with (44), we must conclude that $z' \geq x_W$. $\square$

Claim 3. For any coalition S , $U^{\text{naive}}(S, \mathcal{G}) \subseteq U^\sigma(S, \mathcal{G})$.

Proof. Suppose there is $x \in U^{\text{naive}}(S, \mathcal{G})$ but with $x \notin U^\sigma(S, \mathcal{G})$. Then there is some subcoalition T and $y \in \sigma(T, \mathcal{G})$ such that $y \gg x_T$. Because $x \in U^{\text{naive}}(S, \mathcal{G})$, we have $y \notin \Psi(T)$. But then by Claim 2, there exists $W \subset T$ and $w \in \Psi(W)$ such that $w \geq y_W \gg x_W$, which contradicts the hypothesis that $x \in U^{\text{naive}}(S, \mathcal{G})$. $\square$

The Theorem is an immediate consequence of Claims 1 and 3. $\square$

Proof of Theorem 3. For any coalition S , let $\Gamma(S) := \sum_{j \in S} \gamma_j$ and $\alpha(S) := v(S)/\Gamma(S)$. Let $S_1 \in \arg \max_{S \subseteq N} \alpha(S)$. Recursively, with $S_1, \dots, S_{k-1}$ already defined, let $S_k \in \arg \max \alpha(S)$ over $S \subseteq N - \cup_{j=1}^{k-1} S_j$. This yields a structure $\pi = \{S_1, \dots, S_m\}$, where $\alpha(S_j) \geq \alpha(S_{j+1})$ for all $j = 1, \dots, m-1$. Consider the allocation $y = (y_1, \dots, y_n)$ in π obtained by horizontally stacking the unconstrained Nash payoffs for each coalition in π . That is, if $i \in S_j$, then (remembering that $d = 0$) $y_i = \gamma_i \alpha(S_j)$. Finally, recalling superadditivity, we note that there is $x \geq y$ that is feasible for the grand coalition, and Pareto-efficient.

We claim that there is no $T \subseteq N$ and $z \in \mathbb{R}^T$ such that z is the unconstrained Nash bargaining solution for T and $z \gg x_T$. For if not, then there is T and z feasible for T such that $z_i = \gamma_i \alpha(T) > y_i$ for all $i \in T$. Let S_k be the lowest-indexed coalition in π that has a nonempty intersection with T . Then for any $i \in T \cap S_k$,

$$\gamma_i \alpha(T) > \gamma_i \alpha(S_k).$$

But this means that $\alpha(T) > \alpha(S_k)$, which contradicts our construction of the partition π .

So the allocation y — and *a fortiori* x — is unblocked by any coalition T using its unconstrained Nash bargaining solution. Thus, $U^{\text{naive}}(N, \mathcal{G}) \neq \emptyset$. Because $F(S)$ is compact and convex for all S , Theorem 2 applies, and $U^\sigma(N, \mathcal{G}) \neq \emptyset$. We can now maximize $\prod_{j \in N} x_j^{\gamma_j}$ over this nonempty, compact set to obtain the coalitional solution. Because the grand coalition is viable and superadditivity holds, any such solution is efficient.

By superadditivity, the same argument applies to show the viability of every coalition S . $\square$

Proof of Theorem 4. In a TU game with symmetric weights, the unconstrained Nash solution for every T contains only the equal division solution:

$$\Psi(T) = \{e(T)\}, \text{ where } e(T) := \left(\frac{v(T)}{|T|}, \dots, \frac{v(T)}{|T|} \right).$$

Define $U := \{x \in F(S) \mid x \text{ is unblocked by any } (T, y) \text{ with } T \subset S \text{ and } y = e(T)\}$. Because $F(S)$ is compact and convex for all S , Theorem 2 applies, and therefore:

$$(45) \quad \sigma(S, \mathcal{G}) = \arg \max_{x \in U} \prod_{j \in S} x_j.$$

Define $U^+ := \{x \in U \mid \sum_{i \in S} x_i = v(S)\}$. We claim that

$$(46) \quad \sigma(S, \mathcal{G}) \subseteq L(U^+).$$

It is obvious that $\sigma(S, \mathcal{G}) \subseteq U^+$, because every $x \in U \setminus U^+$ is *strictly* Pareto-dominated by some $y \in U^+$. So if the claim is false, there exists $x \in \sigma(S, \mathcal{G})$ and $y \in U^+$ such that y Lorenz-dominates x . Additionally, $x \gg 0$, because $x \in U$ requires $x_i \geq v(\{i\}) > 0$ for all i . It therefore follows from symmetry and strict quasiconcavity of the product operator on the strictly positive orthant (see Dasgupta, Sen, and Starrett 1973) that

$$\prod_{q \in S} y_q > \prod_{q \in S} x_q,$$

which contradicts (45). That proves the claim and establishes (46).

To complete the proof of the first part, note that by Theorem 1 of Dutta and Ray (1991) establishes that the recursively defined egalitarian solution coincides with the Lorenz-maximal allocations of the set constrained only by the equal-division threats $e(T)$, which is precisely our U^+ , so:

$$(47) \quad E(S) = L(U^+).$$

Combining (46) and (47) completes the first part of the proof.

The second part follows from Theorem 3 specialized to the case of equal weights. $\square$

References

- Aumann, Robert J and Michael Maschler. 1985. “Game Theoretic Analysis of a Bankruptcy Problem from the Talmud.” *Journal of Economic Theory* 36 (2):195–213.
- Bagwell, Kyle, Robert W. Staiger, and Ali Yurukoglu. 2020. ““Nash-in-Nash” Tariff Bargaining.” *Journal of International Economics* 122:103263.
- Bernheim, B. Douglas, Bezalel Peleg, and Michael D Whinston. 1987. “Coalition-Proof Nash Equilibria I. Concepts.” *Journal of Economic Theory* 42 (1):1–12.
- Bhattacharya, Ananyo. 2022. *The Man from the Future: The Visionary Life of John von Neumann*. W. W. Norton & Company.
- Binmore, K, A Shaked, and J Sutton. 1989. “An Outside Option Experiment.” *Quarterly Journal of Economics* 104 (4):753–770.
- Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky. 1986. “The Nash Bargaining Solution in Economic Modelling.” *The RAND Journal of Economics* :176–188.
- Bloch, Francis. 1996. “Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division.” *Games and Economic Behavior* 14:90–123.
- Bondareva, Olga N. 1963. “Some Applications of Linear Programming Methods to the Theory of Cooperative Games.” *Problemy Kybernetiki* 10:119–139.
- Burguet, Roberto and Ramon Caminal. 2016. “Coalitional Bargaining with Consistent Counterfactuals.” Mimeo., Instituto de An’alisis Económico (CSIC), Barcelona.
- Chatterjee, Kalyan, Bhaskar Dutta, Debraj Ray, and Kunal Sengupta. 1993. “A Noncooperative Theory of Coalitional Bargaining.” *The Review of Economic Studies* 60 (2):463.
- Chwe, Michael Suk-Young. 1994. “Farsighted Coalitional Stability.” *Journal of Economic Theory* 63 (2):299–325.
- Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S. Lee. 2019. ““Nash-in-Nash” Bargaining: A Microfoundation for Applied Work.” *Journal of Political Economy* 127 (1):163–195.
- Compte, Olivier and Philippe Jehiel. 2010. “The Coalitional Nash Bargaining Solution.” *Econometrica* 78 (5):1593–1623.
- Dasgupta, Partha, Amartya Sen, and David Starrett. 1973. “Notes on the Measurement of Inequality.” *Journal of Economic Theory* 6 (2):180–187.
- Davis, Morton and Michael Maschler. 1965. “The Kernel of a Cooperative Game.” *Naval Research Logistics Quarterly* 12 (3):223–259.

- de Clippel, Geoffroy and Roberto Serrano. 2008a. *Bargaining, Coalitions and Externalities: A Comment on Maskin*. Brown University, Working Paper 2008-16.
- . 2008b. “Marginal Contributions and Externalities in the Value.” *Econometrica* 76 (6):1413–1436.
- Dutta, Bhaskar and Debraj Ray. 1989. “A Concept of Egalitarianism Under Participation Constraints.” *Econometrica* 57 (3):615–635.
- . 1991. “Constrained Egalitarian Allocations.” *Games and Economic Behavior* 3:403–422.
- Dutta, Bhaskar and Rajiv Vohra. 2017. “Rational Expectations and Farsighted Stability.” *Theoretical Economics* 12:1191–1227.
- Farrell, Joseph and Suzanne Scotchmer. 1988. “Partnerships.” *The Quarterly Journal of Economics* 103 (2):279–297.
- Fields, Gary S and John CH Fei. 1978. “On Inequality Comparisons.” *Econometrica* :303–316.
- Harsanyi, John C. 1959. “A Bargaining Model for the Cooperative n -Person Game.” In *Contributions to the Theory of Games IV*, edited by R. D. Luce and A. W. Tucker. Princeton: Princeton University Press, 325–355.
- . 1974. “An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition.” *Management Science* 20 (11):1472–1495.
- Ho, Kate and Robin S. Lee. 2017. “Insurer Competition in Health Care Markets.” *Econometrica* 85 (2):379–417.
- . 2019. “Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets.” *American Economic Review* 109 (2):473–522.
- Horn, Henrick and Asher Wolinsky. 1988. “Bilateral Monopolies and Incentives for Merger.” *The RAND Journal of Economics* 19 (3):408–419.
- Isbell, J. R. 1960. “A modification of Harsanyi’s Bargaining Model.” *Bulletin of the American Mathematical Society* 66 (2):70–73.
- Kalai, Ehud. 1977. “Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons.” *Econometrica* 45:1623–1630.
- Kalai, Ehud and Meir Smorodinsky. 1975. “Other Solutions to Nash’s Bargaining Problem.” *Econometrica* 43 (3):513–518.
- Kaneko, Mamoru. 1980. “An Extension of the Nash Bargaining Problem and the Nash Social Welfare Function.” *Theory and Decision* 12:135–148.
- Kolm, S. Ch. 1969. “The Optimal Production of Social Justice.” In *Public Economics*, edited by Margolis J. and H. Guitton. Macmillan UK, 145–200.

- Lee, Robin S. and Kyna Fong. 2013. "Markov-Perfect Network Formation: An Applied Framework for Bilateral Oligopoly and Bargaining in Buyer-Seller Networks." Unpublished.
- Loury, Glenn C. 1979. "Market Structure and Innovation." *The Quarterly Journal of Economics* 93 (3):395–410.
- Lucas, William F. 1963. *On Solutions to n -Person Games in Partition Function Form*. Ph.d. dissertation, University of Michigan, Ann Arbor, MI.
- Mariotti, Marco. 1998. "Nash bargaining theory when the number of alternatives can be finite." *Social Choice and Welfare* 15 (3):413–421.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.
- Maschler, Michael and Bezalel Peleg. 1966. "A Characterization, Existence Proof and Dimension Bounds for the Kernel of a Game." *Pacific Journal of Mathematics* 18 (2):289–328.
- Maskin, Eric. 2003. *Bargaining, Coalitions and Externalities*. Presidential Address to the Econometric Society.
- Myerson, Roger B. 1977. "Values of Games in Partition Function Form." *International Journal of Game Theory* 6:23–31.
- Nash, John F. 1950. "The Bargaining Problem." *Econometrica* 18 (2):155–162.
- Osborne, D.K. 1976. "Irrelevant Alternatives and Social Welfare." *Econometrica* 44 (5):1001–1015.
- Ray, Debraj. 1989. "Credible Coalitions and the Core." *International Journal of Game Theory* 18 (2):185–187.
- Ray, Debraj and Rajiv Vohra. 1997. "Equilibrium Binding Agreements." *Journal of Economic Theory* 73 (1):30–78.
- . 1999. "A Theory of Endogenous Coalition Structures." *Games and Economic Behavior* 26 (2):286–336.
- . 2001. "Coalitional Power and Public Goods." *Journal of Political Economy* 109 (6):1355–1385.
- . 2015a. "Chapter 5 - Coalition Formation." In *Handbook of Game Theory with Economic Applications*, vol. 4, edited by H. Peyton Young and Shmuel Zamir. Elsevier, 239–326.
- . 2015b. "The Farsighted Stable Set." *Econometrica* 83 (3):977–1011.
- . 2019. "Maximality in the Farsighted Stable Set." *Econometrica* 87 (5):1763–1779.

- Reinganum, Jennifer F. 1982. "A Dynamic Game of R and D: Patent Protection and Competitive Behavior." *Econometrica* 50 (3):671–688.
- Salant, Stephen W., Sheldon Switzer, and Robert J. Reynolds. 1983. "Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium." *The Quarterly Journal of Economics* 98 (2):185–199.
- Scarf, Herbert E. 1967. "The Core of an N Person Game." *Econometrica* 35 (1):50–69.
- Serrano, Roberto and Ken Ichi Shimomura. 1998. "Beyond Nash Bargaining Theory: The Nash Set." *Journal of Economic Theory* 83 (2):286–307.
- Shapley, Lloyd S. 1967. "On Balanced Sets and Cores." *Naval Research Logistics Quarterly* 14:453–460.
- Shapley, Lloyd S and Martin Shubik. 1971. "The Assignment Game I: The Core." *International Journal of Game Theory* 1 (1):111–130.
- Thomson, William. 1994. "Chapter 35, Cooperative Models of Bargaining." In *Handbook of Game Theory with Economic Applications*, vol. 2, edited by Robert J Aumann and Sergiu Hart. Elsevier, 1237–1284.
- . 2022. "On the Axiomatic Theory of Bargaining: A Survey of Recent Results." *Review of Economic Design* 26 (4):491–542.
- Thrall, Robert M. and William F. Lucas. 1963. "*N*-Person Games in Partition Function Form." *Naval Research Logistics Quarterly* 10 (1):281–298.
- von Neumann, John. 1949. *Private Communication to Professor David Gale*. Container 24, John von Neumann and Klára Dán von Neumann Papers, Manuscript Division, Library of Congress, Washington, D.C.
- Xu, Yongsheng and Naoki Yoshihara. 2020. "Nonconvex Bargaining Problems: Some Recent Developments." *Homo Oeconomicus* 37 (1-2):7–41.
- Zhou, Lin. 1997. "The Nash Bargaining Theory with Non-Convex Problems." *Econometrica* 65 (3):681–685.