Review of Economic Studies (2007) **74**, 1125–1147 0034-6527/07/00391125$02.00

Coalition Formation with Binding Agreements

KYLE HYNDMAN
Southern Methodist University

and

DEBRAJ RAY
New York University

First version received November 2005; *final version accepted January* 2007 (*Eds.*)

We study coalition formation in "real time", a situation in which coalition formation is intertwined with the ongoing receipt of pay-offs. Agreements are assumed to be permanently binding: They can only be altered with the full consent of existing signatories. For characteristic function games we prove that equilibrium processes—whether or not these are history dependent—must converge to efficient absorbing states. For three-player games with externalities each player has enough veto power that a general efficiency result can be established. However, there exist four-player games in which all Markov equilibria are inefficient from every initial condition, despite the ability to write permanently binding agreements.

1. INTRODUCTION

In this paper, we study a model of binding agreements in which pay-offs are received concurrently with the formation of coalitions and agreements. The model is extremely general, in that it accommodates a wide variety of proposer protocols, superadditive and non-superadditive pay-off structures, the ability to constantly make new proposals, and arbitrary history dependence in strategies. The approach is non-cooperative, but all agreements are *binding*, in the sense that their renegotiation requires the consent of all affected parties. This model of "coalition formation in real time" follows recent work by Konishi and Ray (2003) and Gomes and Jehiel (2005). It is potentially applicable to a variety of economic situations in which group formation is important, agreements can be written, and the process is dynamic and open ended: oligopolies, public goods provision, lobbies, conflict models, customs unions and free trade agreements, political party formation, and voting are just a few of many examples.[1]

The model, with its extremely flexible class of protocols, ongoing negotiation, and no restrictions on strategies, is well placed to serve as a formal examination of the Coase Theorem. This is not to say that that the so-called "Theorem" should be the only motivation for our study, but it does place our main question in context: Does an equilibrium path of agreements necessarily converge to efficient outcomes, at least in the long run? Existing literature throws mixed light on this question.

First, as Chatterjee, Dutta, Ray and Sengupta (1993), Bloch (1996), and Ray and Vohra (1997, 1999) have argued, the inability to renegotiate agreements can generate inefficiency:

1. For oligopolies, see, Bloch (1995) and Ray and Vohra (1999). For public goods, see Ray and Vohra (2001). For group-based models of conflict, see Esteban and Ray (1999), Esteban and Sákovics (2004), and Bloch, Sánchez-Pagés and Raphael (2006). For a model of political party formation, see Levy (2004). For models of free trade agreements or customs unions, see Krishna (1998), Aghion, Antras and Helpman (2004), and Seidmann (2005).

even if binding agreements can be freely negotiated to begin with, coalitional considerations fundamentally impede the attainment of efficiency if the ability to *re*write those agreements is missing.

Seidmann and Winter (1998) and Okada (2000) show that (long-run) efficiency may be reinstated if renegotiation is ongoing and coalitions can only expand. But these papers—and most others on coalition formation—study Markovian equilibria, which leads to a second issue.

With history dependence, additional complications appear. For instance, even simple bargaining games with three or more players are known to exhibit multiple equilibria, many of which are inefficient. The first result along these lines is due to Herrero and Shaked (see Herrero, 1985), in the context of the Rubinstein (1982) bargaining model. The problem is more widespread though and applies (with three or more players) to every *characteristic function* game, in which not only the "grand coalition" of all agents has a well-defined worth, but so does every other coalition. Chatterjee *et al.* (1993) state a folk-theorem-like proposition for all such games, not just n-person bargaining situations. This literature suggests—along the unsatisfactory lines of the folk theorem—that a plethora of outcomes (including inefficient ones) is possible.

With the exception of Gomes and Jehiel (2005)—to be discussed in more detail—the literature on ongoing renegotiation restricts itself to characteristic functions.[2] In many situations, externalities across players are fundamental, but the characteristic function effectively rules such externalities out.[3]

In this paper, we approach these questions at a fairly general level. We first study environments with no externalities, but allow for ongoing negotiations using arbitrary history-dependent strategies.[4] Our main result, summarized in Propositions 1 and 2, is that for all games describable by characteristic functions, as long as the number of possible states is finite, pure strategies are used, and players satisfy a mild lexicographic "benignness" condition, *every equilibrium path of agreements must converge to an efficient outcome*.

This proposition is particularly remarkable in the light of the folk-theorem-like results obtained in Herrero (1985) and Chatterjee *et al.* (1993). Under repeated negotiation, we assert that no amount of history dependence in strategies can hold players away from an (ultimately) efficient outcome. Mainly because we allow for such history dependence, but also because we include both superadditive and non-superadditive cases, this proposition represents a substantial extension of Okada (2000) and Seidmann and Winter (1998), who showed that renegotiation achieves efficiency in superadditive characteristic functions when equilibria are restricted to be Markovian. We attempt to evaluate the strength and generality of this result in Section 3.4.

The assumption of benignness, in conjunction with the scenario of ongoing negotiations, plays a fundamental role in the efficiency result. The condition, which may be interpreted either as a restriction on preferences or as a refinement of equilibrium, states that each player prefers higher pay-offs for other agents provided she is not made worse off thereby. This lexicographic restriction entails no sacrifice of player pay-offs and so may be regarded as mild. Indeed, it is

2. Bloch (1996) and Ray and Vohra (1997, 1999) consider more general scenarios with intercoalitional externalities, but do not consider ongoing negotiation. Seidmann and Winter (1998) and Okada (2000) do study ongoing negotiation, but restrict themselves to characteristic functions.

3. The implications of ongoing negotiation for environments with externalities are largely unknown. Do externalities automatically imply the failure of efficiency? After all, *binding* agreements are possible: If there is a Pareto-superior outcome it can always be proposed and players are free to accept it, in which case the outcome is binding thereafter. On the other hand, such a proposal may open unwanted doors in the case of rejection.

4. It must be noted that there is interesting middle ground between the assumption of irreversible commitments and ongoing negotiation. For instance, a coalition might be able to *decide* whether or not to permanently exit the game or entertain further offers. On this issue, see Bloch and Gomes (2006).

mild enough that it has no cutting power in the standard folk-theorem-like results. But coupled with the assumption of ongoing negotiations, it eliminates inefficiency.

We then turn to a study of externalities using the device of *partition functions*. These are natural extensions of characteristic functions that allow pay-offs for a coalition to depend on the entire coalition structure in which it is embedded.[5] We provide an example of a three-player partnership game in which there is an inefficient absorbing state *in every equilibrium* (other related examples are presented in a set of Supplementary Notes available on the journal's website). This example shows that there is no hope of replicating Proposition 2 for games with externalities.

However, we do manage to establish a weaker result for all three-player games that satisfy a mild transferability property: there is always a *Markovian* equilibrium with at least one efficient absorbing state (Proposition 4).

For four-person games, however, the preceding result is false. We show that there exists a four-person game in which *every* Markovian equilibrium must exhibit inefficiency from *every* starting point. This suggests an open question that we do not address: for every game, is there some history-dependent equilibrium that guarantees efficiency from some starting point?

Our approach is particularly motivated by Konishi and Ray (2003), who introduced a model of coalition formation in "real time". In this model, coalitions block proposed allocations with allocations of their own, but use as pay-off functions the entire value emanating from the intertemporal allocation process. This approach studies coalition formation and the generation of pay-offs as one intertwined process. The great advantage of this viewpoint is that one need not seek new definitions of farsightedness for games of coalition formation.[6] Players have discount factors and discount factors that approach one approximate farsightedness.

The main difference in this paper is that we adopt an explicitly non-cooperative approach, in which cooperative game-theoretic constructs such as "blocking coalitions" are eschewed. In our model, proposals are made, accepted, or rejected by individuals in an explicit strategic setting. This brings us closer to the bargaining literature, though the setting is more general. (We also drop the ubiquitous Markov assumption.)

Our results are also related to recent work by Gomes and Jehiel (2005), who study coalitional games with proposers and responders (as we do) and allow for coalition formation and pay-offs to occur together. However, they too restrict attention to Markovian equilibrium. But the main difference between our paper and theirs—which makes the two exercises complementary—is that they assume that proposers are able to make arbitrarily large *upfront* transfers to get a coalition to form. This has very different implications, which we discuss below.

2. A PROPOSAL-BASED MODEL OF COALITION FORMATION

In this section, we introduce a simple, yet general model. First, an informal description: Following each history, date t begins under the shadow of a "going state", in place from the previous date. Using a proposer protocol given to the model, a player is chosen to make a proposal. The player proposes a *state*, possibly different from the one already in place. The proposal must be made to an "approval committee"—a particular coalition that can implement the proposed state. (The formalities below ensure that "no change" needs no approval.) If the proposal is unanimously approved by the approval committee the state is implemented; otherwise it stays where it was. This process continues *ad infinitum*. Each player receives expected pay-offs as described

5. Partition functions were first used by Thrall and Lucas (1963). Ray and Vohra (1997, 1999) discuss why characteristic functions are inadequate to capture many economic situations, and they motivate the use of partition functions.

6. See, among others, Aumann and Myerson (1988), Ray (1989), and Chwe (1994).

below, where expectations are taken not just over proposer choices, but possibly over the stochastic choice of proposal as well.[7]

Now for a more formal account.

2.1. *General specification*

A proposal-based model of coalition formation in real time consists of the following objects:

(1) A finite set N of *players*.
(2) A compact set X of *states*, and an infinite set $t = 0, 1, 2\ldots$ of time periods.
(3) An initial state x_{-1} given at the start of date 0.
(4) A *protocol* describing the probabilistic choice of proposers—and order of respondents—at each date t, possibly depending on the history of events up to that date.
(5) For each state x and proposed new state y, a collection of subsets $\mathcal{S}(x, y)$ that can "approve" the move, with $\mathcal{S}(x, x)$ the collection of *all* subsets of N.[8]
(6) For each player i, a continuous one-period pay-off function u_i defined on X, and a (common) discount factor $\delta \in (0, 1)$.

2.2. *Partition functions*

While the above specification is extremely general, our interest in permanently binding agreements calls for more structure.

Let π be a partition of N into coalitions. A *partition function* assigns to each partition π and coalition $S \in \pi$ a set of pay-off allocations $U(S, \pi)$, each efficient for S given π.

Partition functions have been freely used in the literature on coalition formation.[9] They may be interpreted as the outcome of a coalitional game, in which individuals within each coalition attempt to cooperate, perhaps in a limited way, while interaction across coalitions is non-cooperative (Ray and Vohra, 1997, 1999). Unlike the device of characteristic functions so familiar in cooperative game theory, partition functions allow for a rich array of externalities across coalitions. Of course, a characteristic function—the case of no externalities—is a special case in which U is independent of π.

It is only fair to note, though, that the conversion of a game into partition function form is not automatic. The very definition presumes a "product structure" in pay-offs across coalitions, with each coalition having access to its efficient pay-offs (efficient relative to the structure π, of course). This strains the conversion if the underlying stage game has multiple equilibria.[10] The conversion works best when the underlying stage game across coalitions has a unique equilibrium for each coalition structure.[11]

We suppose that every state x can be expressed as a pair $(\pi, \mathbf{u})$, where π is a partition or coalition structure, and $\mathbf{u}_S \in U(S, \pi)$ for every coalition $S \in \pi$.[12]

7. While we have claimed generality, it is only to be noted that our formulation is not the only possible one. For instance, the very description of a protocol rules out features, such as simultaneous offers, or proposals that can stay on the table for (endogenously) varying lengths of time. Indeed, the very assumption that proposals are made may be restrictive. For instance, Konishi and Ray (2003) adopt a "blocking approach", which has no explicit bargaining structure.

8. This effectively ensures that "no change" needs no approval.

9. Most of the references in the introduction use the concept.

10. In the transferable utility case, for example, each coalition will have to expect its best equilibrium outcome to prevail.

11. One can think of a variety of economic and political situations, such as those mentioned in the Introduction, in which the uniqueness example is reasonable: oligopolies, public goods, lobbying, customs unions ....

12. The notation $\mathbf{u}_S$ denotes the projection of the vector $\mathbf{u}$ on S.

An implicit feature of this assumption is that a formed coalition cannot help but play its "best response" given the ambient coalition structure (see Ray and Vohra, 1997, for more discussion). In a sense, this is what it means to form a coalition: that "across coalitions", no cooperation or commitment is possible. One might formulate a model in which a new coalition precommits to an action vector. Newly formed coalitions would then be acting as Stackelberg leaders, and such a model would blend both the ability to form a coalition and the ability to commit to actions *vis-à-vis* other coalitions. In some situations this may be a useful formulation, but it is not one that we follow here.

2.3. *Binding agreements*

In this paper, we study "binding" agreements. The appropriate way to define the concept is to place restrictions on the approval committees for each proposed move. Loosely speaking, we would like to say that agreements are *binding* if an individual is on the approval committee for every proposed move that will "affect" an ongoing agreement enjoyed by that individual.

The operative word is "affect", and we break this into two parts. First, if a player's coalitional *membership* is affected as a consequence of a proposed move, the move *must* be disrupting some previous agreement to which that player was a signatory. Existing coalitional membership is, after all, the product of some past agreement. In this case we assume that the individual in question must be on the approval committee for the move.

Second, a proposed move might affect the (ongoing) *pay-off* to a particular agent, *without* altering her coalitional membership. Must consent be sought from that agent? The situation here is more subtle. It may be that the pay-off is affected simply because a fellow member of a coalition wishes to reallocate the worth of that coalition. In that case—given that the existing allocation is in force—it is only reasonable that our agent be on the approval committee for the move. On the other hand, our agent's pay-off may be affected because of a coalitional change elsewhere in the system, which then affects our agent's coalition via an externality. Our agent is "affected", but need not be on the approval committee because she was not part of the agreement "elsewhere" in the first place.[13]

We may summarize more formally as follows. For any move from x to y, let $C(x, y)$ denote the set of individuals whose coalitional membership is altered by the move, and $P(x, y)$ the set of individuals whose one-period pay-offs are altered by the move. Say that agreements are *binding* if the following restrictions on approval committees are satisfied:

- (B.1) For every state x and proposed move y, $C(x, y)$ is a subset of any approval committee for the move.
- (B.2) Consider a coalition S with membership entirely untouched by a move from state x to state y. Then, provided either that there has been no change at all in the coalitional structure or that pay-offs are described by a characteristic function, every member of $S \cap P(x, y)$ must belong to any approval committee for the move.

The discussion above indicates that (B.2) is the more subtle of the two restrictions. The idea behind (B.2) is simply this: fix a coalition and a move that does not alter this coalition. If, moreover, there is no change in the *entire* coalition structure, or if the situation is describable by a characteristic function to begin with, how could the pay-off of a particular agent in the unchanged coalition possibly change? The answer implicit in (B.2) is that it could *only* have

13. Notice that we would not insist that our player should *not* be on that approval committee; it is just that our definition of binding agreements is silent on the matter.

changed because there is a deliberate reallocation within that coalition, and then (B.2) demands that all individuals affected by that reallocation must approve the move.[14]

Thus (B.1) and (B.2) formalize binding agreements, and we maintain these restrictions throughout the paper. Sometimes—mainly in the examples—we invoke the *sufficiency* of these restrictions. Say that approval committees are *minimal* if any coalition respecting (B.1) and (B.2) can serve as approval committee for a proposed move.

We study binding agreements in this paper, but an entirely different theory can, in principle, be written down when the restrictions (B.1) and (B.2) are not met. Consider two variants.

First, a theory of "temporary agreements" can be constructed by assuming that agreements only bind for, say, one period. For any move from x to y, any approval committee must contain all members of at least $m-1$ of the m new coalitions that form, and in particular, must include all new coalitions in y that are not subsets of former coalitions in x.[15]

As a second variant, allow a coalition to break up or change if some given fraction (say a majority) of the members in that coalition permit that change. Some political voting games or legislative bargaining would come under this category. Now any approval committee must consist of at least a majority from *every* coalition affected by the move from one state to another.

It is also worth noting that our definition of binding agreements is not the strongest such definition one can write down. For instance, one might require that a coalition once formed can *never* break up again (as in Seidmann and Winter, 1998, or even more strongly, that an agreement once written can never be renegotiated as in Ray and Vohra, 1999).

2.4. *A technical restriction on protocols*

A mild technical restriction on proposer protocols will also be maintained throughout the paper.

[P]. For each i, let H_i be the set of histories after which player i is asked to make a proposal with positive probability. Then this probability is *uniformly* positive on H_i.

All that P rules out is the rather arcane possibility that some player may be asked to propose along a sequence of histories with a corresponding sequence of positive probabilities that converges to 0. P is satisfied for every reasonable protocol that we can think of, including all deterministic and history-independent random protocols. Indeed, we shall sometimes restrict the protocol more strongly, asking that it be entirely independent of history, simply assigning a fixed proposer probability vector to each going state.

2.5. *Strategies and equilibrium*

At each stage of the proceedings, we keep track of past proposers, proposals, and rejectors (if any). A *history* at some stage of the game is a list of such objects up to, but not including, the events that will occur at that stage. Such stages may be of various kinds: A proposer is about to be chosen, or a proposal about to be made, or a responder about to respond, or—such matters concluded—a state about to be implemented. We use obvious nomenclature to distinguish between the different types: "proposer histories", "responder histories", "implementation histories", and so on.

14. This restriction would make less sense if there were multiple equilibria across coalitions. For then a changed pay-off in coalition S could be compatible with no change in the coalition structure if somehow, the selection of across-coalition equilibria were affected by the move. No "deliberate reallocation" within S is involved.

15. The interpretation is that if a new coalition is formed by taking members from more than one erstwhile group, then all the members of the new coalition must approve the move. At best one coalition may be left out of the approval process, and this coalition must be a subset of an ertswhile coalition. It is to be interpreted as a "residual" left by the other "perpetrating coalitions" (compare with the definition of perpetrators and residuals in Ray and Vohra, 1997).

At proposer or responder histories players have to take deliberate actions. A full listing of a particular player's actions for all such histories is a *strategy* for that player. Notice that we are being deliberately quite general here by allowing for all history-dependent strategies. We will see why in the next section.

To describe strategies more formally, consider an individual k. For a proposer history h at which k is meant to propose, she must choose a (possibly new) state y and an approval committee S for the proposed move. She could employ a behaviour strategy, which would be a probability distribution over (y, S). Denote by $\mu_k(h)$ the probability distribution that she uses at proposer history h.[16]

Likewise, at a responder history h at which k is meant to respond, denote by $\lambda_k(h)$ the probability that k will accept the going proposal under that history. The full collection $\sigma = \{\mu_k, \lambda_k\}$ over all players k is a *strategy profile*.

A strategy profile σ induces *value functions* for each player. These are defined at all histories of the game, but the only ones that we will need to track are those just prior to the implementation of a fresh state (or the unaltered continuation of a previous state). Call these *implementation histories*. On the space of such histories, every strategy profile σ (in conjunction with the given proposer protocol) defines a stochastic process P^σ as follows. Begin with an implementation history. Then a state is indeed "implemented". Subsequently, a new proposer is determined. The proposer proposes a state. The state is then accepted or rejected. (The outcome in each of these last three events may be stochastic.) At this point a new implementation history h' is determined. The entire process is summarized by the transition P^σ on implementation histories.

For each person i and given an implementation history h, the *value* for i at that date is given by

$$V_i^\sigma(h) = (1-\delta)u_i(x) + \delta \int V_i^\sigma(h')P^\sigma(h, dh'), \tag{1}$$

where x is the state implemented at h. Given any transition P^σ, a standard contraction mapping argument ensures that V_i^σ is uniquely defined.

Say that a strategy profile σ is an *equilibrium* if two conditions are met for each player i:

(a) At every proposer history h for i, $\mu_i(h)$ has support within the set of proposals that maximize the expected value $V_i(h')$ of i, where h' is the subsequent implementation history induced by i's actions and the given responder strategies.
(b) At every responder history for i, $\lambda_i(h)$ equals 1 if $V_i^\sigma(h') > V_i^\sigma(h'')$, equals 0 if the opposite inequality holds, and lies in $[0, 1]$ if equality holds, where h' is the implementation history induced by acceptance, and h'' the implementation history induced by rejection.

In the case in which the proposer protocol is history independent, say that strategies are *Markovian* if h can be replaced by the going state x everywhere in the definitions above. A *Markov equilibrium* is an equilibrium involving Markov strategies.

This is a well-defined game of perfect information. Given that X is compact and u_i is continuous for every i, the existence of equilibrium is guaranteed (see, for example, Harris, 1985). The existence of Markov equilibrium is easy enough to establish this if X is finite or countable (see the Supplementary Notes).

As discussed in the Introduction and in more detail below, we will sometimes employ a mild refinement. Say that an individual is *benign* if she prefers an outcome in which other individuals are better off, provided that she is not worse off, where the terms "better off" and "no worse

16. Notice that we are allowing any proposer to make a proposal to any committee.

off" are defined with respect to equilibrium value functions, in just the same way as equilibrium pay-offs are. Viewed alternatively as an equilibrium refinement, say that an equilibrium strategy profile is *benign* if for no individual and no history is there a deviation which increases the pay-offs of some players while leaving all other pay-offs (including that of the deviating player) unchanged.

2.6. *Absorption and efficiency*

An equilibrium induces a stochastic process on the space of implementation histories. Consider the stochastic process of one-period pay-off vectors $\mathbf{u}(x_t)$ thus generated. Say that an equilibrium is *absorbing* if $\mathbf{u}(x_t)$ converges almost surely (a.e.) from every initial history.

A vector of pay-offs $\mathbf{u}$ *Pareto dominates* another vector $\mathbf{u}'$ if $\mathbf{u} \gg \mathbf{u}'$. A pay-off vector exhibits (static) *efficiency* if it is not Pareto dominated by any pay-off vector associated with some other state.

We can easily apply this concept to absorbing equilibria with well-defined pay-off limits. Specifically, say that absorbing equilibria are *asymptotically efficient* if their pay-off limits are (static) efficient.

To be sure, we can be more demanding in our efficiency requirement. Say that an equilibrium is *dynamically efficient* from some initial history h if the vector $\mathbf{V}^{\sigma}(h)$ is not Pareto dominated by the infinite-horizon pay-off arising from some, conceivably stochastic, sequence of states.

Whether dynamic or static, our notion of efficiency must respect the same constraints that the players themselves face. In particular, if pay-offs cannot be freely transferred across players it would be inappropriate to label an equilibrium as inefficient if it fails to maximize, say, the sum of total surplus. So lack of transferability, for instance, should not be judged as a *prima facie* correlate of inefficiency. The efficiency definition itself must be suitably modified.

3. CHARACTERISTIC FUNCTION GAMES

3.1. *Preliminaries*

Recall that for characteristic functions, the dependence of coalitional worth on the ambient partition can be dropped, so that the set of feasible pay-off vectors for each coalition S is given simply by $U(S)$. We will prove that every equilibrium of every game of coalition formation derived from a characteristic function must be absorbing (Proposition 1). Moreover, the absorbing limit is efficient (Propositions 2 and 3). These findings substantially extend Seidmann and Winter (1998) and Okada (2000), who prove similar results for *Markovian* equilibria in transferable-utility superadditive games in which coalition structures grow ever coarser.

We require neither transferable utility nor superadditivity (though we do not rule such possibilities out either), nor do we assume that equilibria are Markovian. The possible lack of transferability means that the limit outcome need not maximize the sum of pay-offs. The possible lack of superadditivity means that there is no necessity for the grand coalition to form in the limit. Most important, however, is the lack of restriction to Markovian equilibrium. As already discussed in the Introduction, it is well known that in bargaining games with three or more players, there are history-dependent equilibria with inefficient outcomes (part of the pie may be wasted). Our model includes these scenarios as special cases with the one difference that negotiations—while binding—can always be reopened. With the help of some other restrictions (see below) it is this difference that forces static inefficiency to ultimately disappear, whether or not strategies are conditioned on history.

3.2. *Absorption*

Proposition 1. *Assume* [B.1] *and* [B.2]. *In a game of coalition formation derived from a characteristic function, all equilibria are absorbing.*

Proof. Fix any equilibrium strategy profile σ and initial condition x_{-1}, and consider the stochastic process on histories thus generated. Conditions B.1 and B.2 tell us that for every player i, and for every history h_t with going state x_t,

$$V_i^\sigma(h_{t+1}) \geq u_i(x_t), \tag{2}$$

for every equilibrium realization of the state h_{t+1} conditional on h_t. We claim that the induced stochastic process on V_i^σ is a submartingale. To prove this, recall the functional equation

$$V_i^\sigma(h_t) = (1-\delta)u_i(x_t) + \delta \int V_i^\sigma(h_{t+1}) P^\sigma(h_t, dh_{t+1}), \tag{3}$$

and use (2); it is easy to see that

$$V_i^\sigma(h_t) \geq u_i(x_t), \tag{4}$$

as well. Now suppose, contrary to our assertion, that

$$\mathbf{E}[V_i^\sigma(h_{t+1}) \mid h_t] < V_i^\sigma(h_t),$$

for some history h_t. Then the functional equation (3) implies that

$$V_i^\sigma(h_t) < (1-\delta)u_i(x_t) + \delta V_i^\sigma(h_t),$$

which directly contradicts (4). This proves the claim.

Because V_i^σ is a bounded function on histories, the Martingale Convergence Theorem (see, for example, Ash, 1972, theorem 7.4.3) implies that the induced sequence of random variables $V_i^\sigma(h_t)$ converges a.s. to some limit random variable; call it V^*.

Next, observe that the random variable $Z(h_t) \equiv \mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t)$ is also a submartingale.[17] To see this, recall that $Z(h_{t+1}) \geq V_i^\sigma(h_{t+1})$, so that $\mathbf{E}(Z(h_{t+1}) \mid h_t) \geq \mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t) = Z(h_t)$. It follows that $\mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t)$ converges a.s. to a limit.

Finally, recalling (3) and writing it along any sample path for which both $V_i^\sigma(h_t)$ and $\mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t)$ converge, we must conclude immediately that $u_i(x_t)$ converges along the very same sample path. Hence $u_i(x_t)$ converges a.s., and the equilibrium is absorbing. ‖

3.3. *Efficiency*

In this subsection the following restrictions apply: (i) we study *pure-strategy* equilibria, though we allow for arbitrary forms of history dependence in the strategies, (ii) we assume that the state space X is finite, and (iii) we suppose that players are *benign* in the sense defined in Section 2.5—assuming that they are no worse off, they prefer an outcome path in which no one else is worse off either, and at least one individual is strictly better off. The main result of this section states that under the restrictions described above, *every* limit pay-off of every equilibrium—well defined by Proposition 1—must be efficient.

Proposition 2. *Suppose that the set of states is finite. Then in characteristic function games with permanently binding agreements, every pure-strategy benign equilibrium is asymptotically efficient: every limit pay-off is static efficient.*

17. To be sure, we employ the regular version of conditional expectations in defining Z here.

Proof. Consider any equilibrium σ. The proof of Proposition 1 tells us that both $V_i^\sigma(h_t)$ and $\mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t)$ converge to random variables V^* and $\hat{V}^*$ a.s.

By the submartingale property, $\hat{V}^* \geq V^*$ a.s., but indeed equality must hold. To see this, recall the notation $Z(h_t) \equiv \mathbf{E}(V_i^\sigma(h_{t+1}) \mid h_t)$. Observe that $\mathbf{E}(V_i^\sigma(h_t))$ and $\mathbf{E}(Z(h_{t-1}))$ converge to $\mathbf{E}(V^*)$ and $\mathbf{E}(\hat{V}^*)$, respectively (by the dominated convergence theorem), and that $\mathbf{E}(Z(h_{t-1})) = \mathbf{E}[\mathbf{E}(V_i^\sigma(h_t) \mid h_{t-1})] = \mathbf{E}(V_i^\sigma(h_t))$ for every $t \geq 1$. So $\mathbf{E}(\hat{V}^*) = \mathbf{E}(V^*)$. Because $\hat{V}^* \geq V^*$ a.s., equality must hold a.s.

Consider, then, any path $\{h_t\}$ for which the above equality holds. Then the associated sequence of pay-off vectors $\mathbf{u}(x_t)$, values $\mathbf{V}^\sigma(h_t)$, and conditional expectations $\mathbf{E}(\mathbf{V}^\sigma(h_{t+1}) \mid h_t)$ *all* converge to the same limit $\mathbf{u}^*$. Because there are finitely many states, the limit of one-period pay-offs is a.s. attained after finitely many dates. We claim that the same is a.s. true for $\mathbf{E}(\mathbf{V}^\sigma(h_{t+1}) \mid h_t)$ (and trivially for $\mathbf{V}^\sigma(h_t)$ as a consequence).

Suppose that the assertion is false. Then there is a positive measure of sample histories[18] such that one-period pay-offs converge in finite time but the same is not true for $\mathbf{E}(\mathbf{V}^\sigma(h_{t+1}) \mid h_t)$. Indeed, because there are countably many dates, there is an integer S such that a positive measure of histories exists satisfying all the requirements in the preceding sentence and the additional requirement that one-shot pay-offs converge by date S. Let $\mathbf{\Omega}$ denote this distinguished set of sample paths, and let $\mathbf{\Omega}^c$ be its complement. For each path $\{h_t\} \in \mathbf{\Omega}$ there is an individual i and a subsequence t_k such that for every k,

$$\mathbf{E}(V_i^\sigma(h_{t_k+1}) \mid h_{t_k}) > u_i, \tag{5}$$

where u_i is the *particular* limit of one-shot pay-offs for i along this path. The strict inequality in (5) implies that there is a another subsequence s_k of dates, with $s_k > S$ for all k, such that at each of those dates, some proposer makes a proposal which yields a higher pay-off to player i than the normalized value of u_i. Because we only study pure strategies, such a proposal must be made and accepted with probability at least $\zeta > 0$, where ζ is uniform across histories and individuals and is given by the restriction (P). Now observe that such a proposal, if accepted, *must* subsequently lead to paths that are not in $\mathbf{\Omega}$. This is because (i) $\mathbf{\Omega}$ contains only those paths for which one-period pay-offs have already converged by date S, (ii) every s_k exceeds S, and (iii) an accepted proposal must lead to a change in the (by-then) stationary path of one-period pay-offs. More formally,

$$\text{Prob}(\mathbf{\Omega}^c \mid h_{s_k}) \geq \zeta > 0 \quad \text{for all } k,$$

whenever the path $\{h_t\}$ lies in $\mathbf{\Omega}$. It is easy to see that this must imply $\text{Prob}(\mathbf{\Omega}^c) = 1$, a contradiction. This proves the claim that $\mathbf{u}(x_t)$, $\mathbf{V}^\sigma(h_t)$, and $\mathbf{E}(\mathbf{V}^\sigma(h_{t+1}) \mid h_t)$ all converge in finite time to the same limit $\mathbf{u}^*$, a.s.

We complete the proof by showing that $\mathbf{u}^*$ must be efficient. Suppose not; then there is a state x such that $\mathbf{u}(x) > \mathbf{u}^*$. Suppose a player were to propose x. The offer must be rejected, otherwise we are not in equilibrium. Consider all the rejectors: all the players who will reject conditional on all previous responders accepting. Number these players $1, \ldots, R$ in order of their appearance. For each rejector i, let h_i' denote the history following her acceptance and h_i'' the history following her rejection. Because the last rejector R rejects, it is easy to see that $V_i(h_R'') \geq u_i(x) > u_i(x^*)$. Moreover, no other player can be worse off compared to $\mathbf{u}^*$: $V_i(h_R'') \geq u_i(x^*)$ for all i. In summary,

$$V_i(h_R'') \geq u_i(x^*) \text{ for all } i, \text{ with strict inequality for some } i. \tag{6}$$

18. To be sure, this positive measure is generated by the protocol as well as equilibrium strategies.

Now consider player $R-1$. She, too, rejects the offer. Therefore the first part of (6) holds for the history h''_{R-1}. In general, no more can be said, but because $R-1$ is benign and (6) holds for the history h''_R, it must hold too for the history h''_{R-1}. Continuing recursively in this way, we see that (6) must holds for the history h''_1. But now we have a contradiction. By benignness, then, it is profitable for our proposer to propose x irrespective of whether it is accepted or rejected. ‖

This result is a substantial generalization and extension of other literature, in the main because it establishes efficiency *with no restriction on the degree of history dependence in strategies*. In contrast, the existing literature that establishes bargaining efficiency in characteristic function games restricts attention to Markov-perfect equilibria. Such a restriction is made for good reason: the results of Herrero (1985) and others show that multiple equilibria—many with inefficiency—are endemic when the Rubinstein bargaining model is extended to three or more players. Proposition 2 establishes that under some conditions—which we feel are acceptably mild—the multiplicity may still continue but all limit inefficiency must disappear.

3.4. *Discussion of the efficiency theorem*

We briefly discuss the assumptions behind the efficiency proposition and the extent to which they can be dropped.

3.4.1. Transferable utility and finite state spaces. First, observe that even though our proposition is stated for finite state spaces, we can approximate arbitrarily high degrees of transferable utility. It should therefore not be concluded that our efficiency result is somehow linked to the presence or absence of transferability in pay-offs.

The reader may nevertheless wonder if the proposition goes through if the state space is allowed to be infinite. We are not sure of the answer to this question in general, though we would conjecture that it is in the affirmative. For instance, here is a version of Proposition 2 when the proposer protocol is restricted to be deterministic.

Proposition 3. *Suppose that every individual is benign, the proposer protocol is deterministic and the set of states is compact. Then in characteristic function games with permanently binding agreements, every limit pay-off of every pure-strategy equilibrium is efficient.*

The proof is similar to that of Proposition 2 and is made available in the Supplementary Notes.

3.4.2. Ongoing negotiations and benignness. If negotiations are not permitted to continue indefinitely, then inefficiency is endemic. For n-person bargaining games, where $n \geq 3$, Herrero and Shaked provide the required analysis. The inefficient equilibria in their construction cannot be Markovian. For more general n-person characteristic functions, in which intermediate coalitions have worths, even Markov equilibria can yield inefficiency when ongoing renegotiation is not possible (Chatterjee *et al.*, 1993).

To illustrate these points (as well as the discussion of benignness below), consider a four-player characteristic function. Assume that the grand coalition of all players can achieve a total worth of 1, allocable in any way among the players. Next, suppose that every two-person coalition has a (transferable) worth of x, where x is a positive number. All other coalitions have zero worth.

Suppose that an initial proposer is chosen at random, and after this the first rejector of a going proposal (if any) gets to counterpropose. Suppose further *once* an agreement is made, it cannot be further renegotiated. Then, using arguments similar to Herrero (see Osborne and Rubinstein, 1994, p. 130), and provided that discount factors are close enough to 1, one can easily construct an equilibrium in which players 1 and 2 divide x, while players 3 and 4 divide their x, even if x is "small" (less than 1/2, say). This outcome is supported by rewarding a player for rejecting a deviant offer by receiving the entire unit pie in the next (and, therefore, every future) period.

This equilibrium is inefficient, and what is more, the imposition of benignness will not get rid of it. The reason is that a deviant proposer is *strictly* punished by the above strategies (given that x is strictly positive).

This illustrates the possible inefficiency of history-dependent equilibria. But Markovian equilibria may be inefficient as well. Consider a variant of this example in which only those two-player coalitions that contain player 1 obtain x, and the remaining subcoalitions get 0 (the grand coalition gets 1 as before). Following the arguments in Chatterjee *et al.* (1993), it is now possible to show—for discount factors close to 1—that *every* Markovian equilibrium of this bargaining game is inefficient provided that $1/2 < x < 1$. Such an equilibrium *must* involve a two-person coalition (1 included) approximately splitting x, while the other two players get 0. Moreover, such equilibria are robust to the imposition of benignness.

In both these examples, an agreement once arrived at shuts down all further negotiations among the players who agree. It should therefore be clear that the ability to conduct ongoing negotiations is, generally speaking, necessary for Proposition 2. But on its own, it is not enough. It is true that the Okada–Seidmann–Winter analysis implies the restoration of asymptotic efficiency for *Markovian* equilibria. But there are history-dependent equilibria that continue to display inefficiency.

In particular, add "ongoing negotiation" to the examples above. Provided that discount factors are high enough, one can construct equilibria in which *every* state is absorbing. Such absorption is obvious for any status-quo state involving the grand coalition. For other states, the status-quo outcome is supported as an absorbing state as follows: Reward a player for rejecting a deviant offer by allowing her to ask for (and receive) the grand coalitional pie in the next (and, therefore, every future) period, net of what the other players are currently receiving in the status quo. It continues to be true—as guaranteed by Proposition 1—that we have pay-off absorption (and in finite time) from every initial condition. However, for some initial conditions, convergence is to an inefficient state.

It is easy to verify directly that such equilibria are not benign. Indeed, our proposition implies that benignness must restore efficiency.

Is the benignness restriction on player preferences reasonable? Obviously, like every assumption it is open to scrutiny. We only mention that benignness has found support in a number of different experimental settings (including bargaining); see, for example, Charness and Grosskopf (2001), Andreoni and Miller (2002), and Charness and Rabin (2002) among others. Indeed, these studies suggest something stronger than our benignness condition: people are sometimes willing to *sacrifice* their own pay-off in order to achieve a socially efficient outcome.

In summary, our overall view is that asymptotic efficiency appears to be broadly guaranteed when *both* negotiations are forever ongoing (in principle) and all agents are benign, in that they do not grudge others a pay-off improvement provided that they do not personally lose in the process. The result holds whether or not pay-offs are transferable. At the same time, these two conditions are critical. Without them, counterexamples to asymptotic efficiency can easily be constructed.

It is true that other assumptions have been made to prove the proposition. But we believe that these are essentially technicalities. For instance, Proposition 3 establishes a variant of the efficiency result when finiteness of the state space is dropped. The Supplementary Notes comment on another assumption: the restriction to pure strategies.

4. GAMES WITH EXTERNALITIES

The ubiquitous absorption and efficiency results reported for characteristic functions break down when there are externalities across players. Equilibrium pay-offs may cycle, and even if they do not do so, inefficiencies occur. Such outcomes are not driven by the self-fulfilling contortions of history dependence. They occur even for Markovian equilibria.

Let us sidestep a common pitfall right away. It is tempting to think of inefficiencies as entirely "natural" equilibrium outcomes when externalities exist. Such an observation is true, of course, for games in which there are no binding agreements. When agreements can be costlessly written, however, no such presumption can or should be entertained. These are models of *binding* agreements, a world in which the so-called "Coase Theorem" is relevant. For instance, in all that we have done so far, two-player games invariably yield efficiency, quite irrespective of whether there are externalities across the two players. This is not to say that the "usual intuition" has no role to play. It must, because the process of negotiation is itself modelled as a non-cooperative game. But that is a very different object from the "stage game" over which agreements are sought to be written. Indeed, while positive results are possible for three-player games (see Proposition 4 below), we are generally in murky waters as soon as the number of players exceeds two.

4.1. *Three-player games*

Three-player games represent an interesting special case. Even when externalities are allowed for, such games share a central feature with their characteristic function counterparts: each player possesses, in effect, a high degree of veto power in all changes which alter her pay-off. This will allow us to prove an efficiency result even for the restricted class of Markovian equilibria. To be sure, the veto power is not absolute, which blocks off the possibility of obtaining stronger efficiency results.

In passing, we note that three-player situations have been the focus of study in several applied models of coalition formation (see, for example, Krishna, 1998; Aghion *et al.*, 2004; Kalandrakis, 2004; and Seidmann, 2005).

4.1.1. The failed partnership. This example attempts to capture the following situation. There are three agents, any two of whom can become "partners". For instance, two of three countries could form a customs union, or a production cartel, or an R&D coalition with a commitment to share ideas. A three-player partnership is assumed not to be feasible (or has very low pay-offs), but see the remarks on superadditivity below.

As partners two players earn a pay-off, while the outsider is reduced to a relatively low pay-off. In the example below, the failed partner is player 1. Partnerships between him and any other agent are dominated—both for the partners themselves and certainly for the outsider—by all three standing alone. In contrast, the partnership between agents 2 and 3 is rewarding (for those agents).

In the example below and those to follow, we simply record those states with non-trivial pay-off vectors and omit any mention of the remaining states, with the presumption that the pay-offs in those states are zero to all concerned. We shall also be somewhat cavalier in our description of equilibrium and ignore these trivial states: equilibrium transitions from those states are implicitly defined in obvious ways.

$$\begin{aligned}
x_0 &: \pi_0 = \{\{1\},\{2\},\{3\}\}, && \mathbf{u}(x_0) = (6,6,6)\\
x_1 &: \pi_1 = \{1,\{23\}\}, && \mathbf{u}(x_1) = (0,10,10)\\
x_2 &: \pi_2 = \{2,\{13\}\}, && \mathbf{u}(x_2) = (5,0,5)\\
x_3 &: \pi_3 = \{\{12\},3\}, && \mathbf{u}(x_3) = (5,5,0).
\end{aligned}$$

Observation 1. *For δ sufficiently close to 1, inefficient outcomes x_2 and x_3 must be absorbing states in every equilibrium.*

We omit a formal proof of this simple observation; the discussion to follow makes the argument clear. Despite the fact that x_2 (or x_3) is Pareto dominated by x_0, player 1 will not accept a transition to x_0. If she did, players 2 and 3 would surely initiate a further transition to x_1. Player 1 *might* accept such a transition if she is very myopic and prefers the short-term pay-off offered by x_0, but if she is patient enough she will surely see ahead to the infinite phase of "outsidership" that will surely follow the short-term gain. In that situation it will be impossible to negotiate one's way out of x_2 or x_3. This inefficiency persists in *all* equilibria, history dependent or otherwise.

Notice that x_2 or x_3 would not be *reached* starting from any other state. This is why the interpretation, the "failed partnership", is useful. The example makes sense in a situation in which players 2 or 3 have been locked in with 1 on a past deal, on expectations which have failed since. To be sure, this interpretation is unnecessary for the formal demonstration of persistent inefficiency from some initial state.

Why can players not negotiate themselves out of, say, x_2? They could, if players 2 and 3 could agree never to write an agreement while at x_0. Are such contracts reasonable? In theory they certainly are possible. However, it may be difficult to imagine that from a legal point of view, player 1, who has voluntarily relinquished all other contractual agreements between 2 and 3, could actually hold 2 and 3 to such a meta-agreement.

This raises a delicate issue to which we return below. Does one interpret the stand-alone option (x_0) as an *agreement* from which further deviations require universal permission? Or does "stand-alone" mean freedom from all formal agreement, in which case further bilateral deals only need the consent of the two parties involved? This example takes the latter view.

4.1.2. Transfers. How does the ability to make transfers affect the example? It is important to distinguish between two kinds of transfers. Coalitional or partnership worth could be freely transferred between the players in a coalition. Additionally, players might be able to make large upfront payments in order to induce certain coalitions to form. In all cases, of course, the definition of efficiency should match the transfer environment.[19]

As we have noted earlier, our model allows for any degree of within-coalition transferability. Indeed, it is very easy to verify that if we allowed for transfers within coalitions, then nothing changes in the failed partnership example.

Upfront transfers *across* coalitions are not allowed in the framework we consider. This is what distinguishes the exercise of this section from Gomes and Jehiel (2005)—GJ hereafter. GJ consider transferable-utility coalitional bargaining games in which any player may make (large) upfront transfers to any set of players. Whether such transfers are reasonable is a question that requires a contextual answer. While in some situations that involve a long time horizon and substantial liquidity constraints, the assumption may be problematic, we do not want to take an explicit position on it. The two exercises are complementary. But a discussion of this issue will clarify the differences between GJ and this section of the paper.

If unlimited upfront transfers are introduced into the failed partnership example of Section 4.1.1, efficient outcomes result from every initial state. Players 2 and 3 could make an upfront payment to 1 to have 1's partner released. But it is easy to write down a variant of the example in which inefficiency persists. For instance, suppose that

19. For instance, if transfers are not permitted, it would be inappropriate to demand efficiency in the sense of aggregate surplus maximization. If a non-transferable utility game displays inefficiency in the sense that "aggregate surplus" is not maximized, this is of little interest: aggregate surplus is simply the wrong criterion.

$$x_0 : \pi(x_0) = \{\{1\}, \{2\}, \{3\}\}, \quad \mathbf{u}(x_0) = (6, 6, 1)$$
$$x_1 : \pi(x_1) = \{1, \{2, 3\}\}, \quad \mathbf{u}(x_1) = (0, 7, 2)$$
$$x_3 : \pi(x_3) = \{\{1, 2\}, 3\}, \quad \mathbf{u}(x_3) = (5, 5, 0),$$

and all other states yield zero pay-offs.

This variant retains the essential features of the previous example, though it is asymmetric. Players 1 and 2 still form a "failed partnership" at x_3. At that state player 1 continues to cling to player 2. They would prefer to move to state x_1, but player 1 rationally fears the subsequent switch to x_2.

But the introduction of upfront transfers in this example has a perverse effect. Instead of taking the inefficiency away (as it does in the previous example), it generates inefficiency from *every* initial condition. Notice that with transfers, we must use aggregate surplus as our defining feature of efficiency, so that only the state x_0 is efficient. An equilibrium displaying asymptotic efficiency can stay away from x_0 for a finite number of dates at best. However, players 2 and 3 invariably have the incentive to move away from x_0 to x_1. Of course, the fact that x_1 is itself inefficient will cause further movement across states as upfront transfers continue to be made along an infinite subsequence of time periods. The precise computation of such transfers is delicate, but the assertion that efficiency cannot be attained should be clear.

Thus the presumption that unlimited transfers act to restore or maintain efficiency is wrong, and GJ show this explictly in their analysis.[20]

Without upfront transfers (but irrespective of whether or not there are transfers within coalitions), we obtain a different result for three-player games, and to this we now turn.

4.1.3. An efficiency result for three-person games. The failed partnership or its later variant is not the only form of inefficiency that can arise. The Supplementary Notes record three other forms of inefficiency, including one which can even arise from the stand-alone starting point of no agreements: the structure of singletons. In the light of these several examples, it is perhaps of interest that a positive efficiency result holds for *every* three-person game satisfying a "minimal transferability" restriction. To state this restriction, let $\bar{u}(i, \pi)$ be the maximum one-period pay-off to player i over all states with the same coalition structure π.

[T] If two players i and j both belong to the same coalition in coalition structure π, then $\bar{u}(i, \pi)$ and $\bar{u}(j, \pi)$ are achieved at different states.

Proposition 4. *Consider a three-person game with a finite number of states and satisfying condition* (T), *with history-independent proposer protocols and minimal approval committees. Then for all δ close enough to 1, there exists an initial state and a stationary Markov equilibrium with efficient absorbing pay-off limit from that state.*

The proof of Proposition 4 exhaustively studies different cases and is therefore relegated to the appendix.[21] But we can provide some broad intuition for the result. Pick any player i and consider her maximum pay-off over all conceivable states. If this maximum is attained at a state x^* in which i belongs to a coalition with two or more players, then observe that i's consent *must* be given for the state to change. (This step is not true when there are four or more players, and invalidates the proposition, as we shall see later.) Because the pay-off in question is i's maximum, it is easy enough to construct an equilibrium in which x^* is an absorbing state.

20. Inefficiency in this example (at least for discount factors close to 1) is also a corollary of GJ, proposition 7.
21. It should be noted that in most cases the result is stronger in that it does not insist upon $\delta \to 1$; in only one case do we rely on $\delta \to 1$.

It therefore remains to consider games in which for every player, the maximum pay-off is attained at states in which that player stands alone. If no such state is absorbing in an equilibrium, one can establish the existence of a cyclical equilibrium path, the equilibrium pay-offs along which are uniquely pinned down by the pay-offs at the state in which all players stand alone. With the transferability condition T, one can now find pay-off vectors for other coalitions (doubletons or more) such that *some* player in those coalitions prefer these pay-offs to the cyclical equilibrium pay-offs. The associated states then become absorbing, and a simple additional step establishes their efficiency.

We conjecture that neither the minimality of approval committees nor the history independence of proposer protocols is needed for this result, but do not have a proof.

4.2. *Four or more players*

Section 4.1.1 shows that with externalities, we cannot expect every absorbing state to be Pareto efficient. At the same time, in three-player games, asymptotic efficiency from *some* initial state holds in at least one equilibrium. However, the same conclusion does not apply when there are four or more players. We now present an example that displays the most severe form of inefficiency: *every* absorbing state in every equilibrium is static inefficient, and every non-convergent equilibrium path in every equilibrium is dynamically inefficient.

Consider the following four player game with minimal approval committees, in which a proposer is chosen with uniform probability at every stage.

$$x_1 : \pi(x_1) = \{\{1,2\},\{3\},\{4\}\}, \quad \mathbf{u}(x_1) = (4,4,4,4)$$
$$x_2 : \pi(x_2) = \{\{1\},\{2\},\{3\},\{4\}\}, \quad \mathbf{u}(x_2) = (5,5,5,5)$$
$$x_3 : \pi(x_3) = \{\{1\},\{2\},\{3,4\}\}, \quad \mathbf{u}(x_3) = (0,0,10,10)$$
$$x_4 : \pi(x_4) = \{\{1,2\},\{3,4\}\}, \quad \mathbf{u}(x_4) = (2,2,2,2).$$

Observation 2. *For δ sufficiently close to* 1, *every stationary Markov equilibrium is inefficient starting from any initial state.*

A formal proof of this result is relegated to the Supplementary Notes, but some discussion may be useful. In part, a logic similar to the failed partnership is at work here. Consider state x_1, in which players 1 and 2 are partners and 3 and 4 are outsiders. If the {12}-partnership disbands, the state moves to x_2, which is better for all concerned. But once at x_2, we see other latent, beneficial aspects of the erstwhile partnership between 1 and 2: if players 3 and 4 now form a coalition, they can exploit 1 and 2 for their own gain; this is the state x_3.

So far, the story is not too different from that of the failed partnership. But the similarity ends as we take up the story from the point at which 3 and 4 "counterdeviate" to x_3. Their gains can be reversed if players 1 and 2 form (or depending on the dynamics, re-form) a coalition. Balance is now restored; this is the state x_4. Finally, in this context, the partnership between 3 and 4 is more a hindrance than a help (just as {12} was in the state x_1), and they have an incentive to disband. We are then "back" to x_1. This line of reasoning helps to prove that all equilibria must be inefficient.

Despite the apparent circularity of the argument above, equilibria do exist that are absorbing from every initial condition. The Supplementary Notes show that there is an equilibrium that converges to the state x_1 regardless of the starting point. There is also another equilibrium with no absorbing states, which displays (dynamic) inefficiency from every initial condition, because it *must* spend non-negligible time at the inefficient states x_1 and x_4.

Three final remarks are in order regarding this example. First, the strong form of inefficiency is robust to (at least) a small amount of transferability in pay-offs. The reason is simple: Regardless of transferability, the state x_3 (respectively x_4) cannot be absorbing since players 1 and 2 (respectively 3 and 4) can guarantee themselves a higher pay-off by initiating a transition to x_4 (respectively x_1).

Second, if players can commit to irreversible exit in the spirit of Bloch and Gomes (2006), an efficient equilibrium exists in this example. At x_1, players 3 and 4 could exit and so commit not to be opportunistic at x_2. Players 1 and 2 can now confidently transit to x_2 without fear of exploitation. (To be sure, once irreversible exit options enter the picture, other instances of inefficiency can easily be constructed.)

Finally, inefficiency in this example can also be overcome by history-dependent strategies. Indeed, x_2 can be supported as an absorbing state provided that deviations from x_2 are punished by a return to the inefficient stationary equilibrium in which x_1 is absorbing.

This last remark creates an interesting contrast between models based on characteristic functions and those based on partition functions. In the former class of models, Seidmann and Winter (1998) and Okada (2000) assure us that ongoing negotiations lead to efficiency under Markovian equilibrium. It is the possibility of history dependence that creates the inefficiency problem, albeit one that we successfully resolved with the help of the benignness condition. In contrast, partition functions are prone to inefficiency under Markovian equilibrium, while history dependence might help to alleviate this problem (it does in this example, though not in the failed partnership).

5. SUPERADDITIVE GAMES

An important feature of the examples in Section 4.2 is that they employ a subadditive pay-off structure. One can make further progress under the assumption of *grand coalition superadditivity*:

[GCS] For every state $x = (\mathbf{u}, \pi)$, there is $x' = (\mathbf{u}', \{N\})$ such that $\mathbf{u}' \geq \mathbf{u}$.

It is worth recording that GCS restores (Markovian) efficiency, at least if the existence of an absorbing limit pay-off is assumed:

Proposition 5. *Under GCS, every absorbing pay-off limit of every Markovian equilibrium must be static efficient.*

The proof follows a much simpler version of the argument for Proposition 2 and we omit it.[22]

Is GCS a reasonable assumption? If we view the merger of two coalitions in a literal way, so that the two erstwhile groups now attempt to cooperate as one, there are many interesting and important problems in which GCS might fail. Factors as varied as antitrust laws, ideology, ancient hatreds, geography, or the competitive spirit may conspire against cooperation in such mergers. For instance, the entire doctrine of healthy competition is based on the failure of physical superadditivity, at least after a point.[23]

22. Note that GCS does not guarantee long-run efficiency in all situations. The example of Section 4.2 can be modified so that GCS holds, *but* there is an inefficient cycle over non-grand coalition states. In order to guarantee that even this form of inefficiency does not persist in the long run, one needs a sufficient amount of transferability in the grand coalition. Indeed, it can be proved that under GCS and the additional assumption that the pay-off frontier for the grand coalition is continuous and concave, *every* Markov equilibrium must be absorbing—and therefore asymptotically efficient.

23. As another example, the "grand coalition" may not include all players; there could always be some exogenous set of agents that cannot form coalitions with our players, but always interact non-cooperatively with them. For this reason

Some game theorists might argue that such a literal merging of coalitions is not entailed by superadditivity at all. They have in mind a different notion of GCS, which is summarized in the notion of the *superadditive cover*. After all, the grand coalition—by cooperatively breaking up—can replicate the pay-offs obtainable in some other coalition structure.[24] But we would have to be careful about applying Proposition 5 in such a situation, for that would presume, in effect, that *future* changes in the strategy of one of the subgroups *must* require the consent of all players. For instance, in the example of the failed partnership, it would debar individuals 2 and 3 from cooperating without the express permission of 1.[25] In many cases, it is hard to see how such a requirement can be legally feasible, at least in the long term.

6. CONCLUSIONS

Our study of coalitional bargaining problems in "real time" yields a number of implications. For characteristic function form games, a very general result for all pure-strategy equilibria (whether history dependent or not) can be established: every equilibrium path of states must eventually converge to some absorbing state, and this absorbing state must be static efficient. Perhaps this the strongest formal vindication of Coase's "theorem", at least for characteristic functions.

In contrast, in games with externalities, matters are more complicated. It is possible to find a three-person example in which there is persistent inefficiency from some initial state, whether or not equilibria are allowed to be history dependent. At the same time, in every three-person game, there is *some* Markovian equilibrium, which yields asymptotic efficiency from some initial condition.

Yet, even this limited efficiency result is not to be had in four-person games. Section 4.2 demonstrates the existence of games in which *every* recurrent class in *every* Markovian equilibrium exhibits static inefficiency. The ability to make unlimited upfront transfers may worsen matters, as Gomes and Jehiel (2005) have shown, and we discuss this issue. We also discuss the implications of imposing superadditivity, at least at the level of the grand coalition, and show that some efficiency properties are restored.

The main open question for games with externalities is whether there always exists some history-dependent equilibrium, which permits the attainment of asymptotic efficiency from *some* initial state (that there is no hope in obtaining efficiency from *every* initial state is made clear in Section 4.1.1). We are convinced that the answer should be in the affirmative: assuming—by way of contradiction—that equilibria are inefficient from every initial state, one should be able to employ such equilibria as continuation punishments in the construction of some efficient strategy profile. Such a result would be intuitive: after all, one role of history-dependent strategies is to restore efficiency when simpler strategy profiles fail to do so.

Finally, the general set-up in Section 2 may be worthy of study, with or without binding agreements. For instance, the general set-up nests games in which agreements are only temporarily binding, or in which unanimity is not required in the implementation of a proposal. We believe that there is merit in exploring these applications in future work.

alone, GCS may fail. To be sure, all failures of GCS must be based on some non-contractible factor, such as the creativity or productivity created by the competitive urge, or the presence of agents with whom contracts cannot be written.

24. For instance, companies might spin off certain divisions, or organizations might set up competing R&D groups, but they might do so "cooperatively". In a word, the grand coalition can agree not to cooperate, if need be.

25. As a possible example, consider no-compete contract clauses that typically debar an executive from working for a competitor firm for a number of years.

APPENDIX

Proof of Proposition 4.

In what follows, we denote by π_0 a singleton coalition structure, by π_i a coalition structure of the form $\{\{i\},\{j,k\}\}$, and by π_G the structure consisting of the grand coalition alone. Use the notation $\pi(x)$ for the coalition structure at state x and $S^i(x)$ for the coalition to which i belongs at x. Subscripts will also be attached to states (*e.g.* x_i) to indicate the coalition structure associated with them (*e.g.* $\pi(x_i) = \pi_i$).

For each i, let $X_i^* = \text{argmax}\{u_i(x) \mid x \in X\}$, with x_i^* a generic element. Finally, we will refer to $\pi(x_i^*) = \pi$ as a *maximizing* (coalition) structure (for i).

Case 1: There exists $i = 1,2,3$ and $x_i^* \in X_i^*$ such that $|S^i(x_i^*)| \geq 2$.

Pick $x_i^* \in X_i^*$ as described and consider the following "pseudo-game". From x_i^*, there does not exist an approval committee capable of initiating a transition to *any* other state. Notice that a Markovian equilibrium exists for this pseudo-game (see the Supplementary Notes for the general existence proof) and that x_i^* is absorbing. Denote by σ^* the equilibrium strategies for the pseudo-game. Return now to the actual game and suppose that players use the strategies σ^*; suppose also from x_i^*, that player i always proposes x_i^* and rejects *any* other transition. For other players $j \neq i$, any proposal and response strategies may be specified. Denote this new strategy profile σ'. Notice that σ^* and σ' specify the same transitions for the pseudo-game and actual game and no player has a profitable deviation from x_i^*. Therefore, σ' constitutes an equilibrium of the actual game. This equilibrium has an efficient absorbing state, x_i^*.

Case 2: For all i and for all $x_i^* \in X_i^*$, $|S^i(x_i^*)| = 1$. A number of sub-cases emerge:

(a) $\pi(x_1^*) = \pi(x_2^*) = \pi(x_3^*) = \pi_0$ for some (x_1^*, x_2^*, x_3^*), but the maximizing structures are not necessarily unique.
(b) $\pi(x_1^*) = \pi(x_2^*) = \pi_0$ and $\pi(x_3^*) = \pi_3$, and while the maximizing structures are not necessarily unique, Case 2(a) does not apply.
(c) For all players $i = 1,2,3$, π_i is the unique maximizing structure.
(d) $\pi(x_1^*) = \pi_0$, $\pi(x_j^*) = \pi_j$, $j = 2,3$ and each maximizing structure is unique.

We now prove the proposition for each of these cases.

Case (a). Here x_0, the unique state corresponding to π_0, is weakly Pareto dominant and we construct an equilibrium as follows. From any state x, every player proposes a transition to x_0 and every player accepts this proposal. A deviant proposal y is accepted if $V_i(y) \geq V_i(x) = (1-\delta)u_i(x) + \delta u_i(x_0)$. This is clearly an equilibrium with x_0 efficient and absorbing.

Case (b). The proof is similar to Case 1. Consider a pseudo-game in which there is no approval committee that can initiate a transition away from x_0. Again, we are assured of a Markovian equilibrium for the pseudo-game; denote the equilibrium strategies by σ^* and notice that x_0 is absorbing. In the actual game, suppose that all players use the strategies given by σ^*, and suppose that, at x_0, players 1 and 2 always propose x_0 and reject any transition from x_0. Call these strategies σ'.

As in Case 1, notice that σ^* and σ' specify the same transitions for the pseudo-game and the actual game, and no player has a profitable deviation from x_0. Therefore, σ' constitutes an equilibrium of the actual game.

The following preliminary result will be useful for cases (c) and (d):

Lemma 1. *Suppose that player i's maximizing structure $\hat{\pi}$ is unique, and that $\hat{\pi} \in \{\pi_0, \pi_i\}$. Let $Y = \{y \mid \pi(y) \in \{\pi_0, \pi_i\} \setminus \hat{\pi}\}$. Then in any equilibrium such that $x \in X_i^*$ is not absorbing, $V_i(x) > V_i(y)$ for all $y \in Y$.*

Proof. We prove the case for which $\hat{\pi} = \pi_i$. The proof of the case for which $\hat{\pi} = \pi_0$ is identical. Let $x \in X_i^*$. Note that $Y = \{x_0\}$. Suppose on the contrary that $V_i(x_0) \geq V_i(x)$. We know that

$$V_i(x) = (1-\delta)\bar{u}_i + \delta \int_X V_i(z) P(x, dz).$$

Now, there could be—with probability μ—a transition to the singletons, which player i need not approve. All other transitions must be approved by i, and she must do weakly better after such transitions. Using $V_i(x_0) \geq V_i(x)$, it follows that

$$V_i(x) \geq (1-\delta)\bar{u}_i + \delta[\mu V_i(x_0) + (1-\mu)V_i(x)] \geq (1-\delta)\bar{u}_i + \delta V_i(x),$$

so that $V_i(x) \geq \bar{u}_i$. Strict inequality is impossible since $\bar{u}_i$ is i's maximal pay-off. So $V_i(x) = \bar{u}_i$, but this means that x is absorbing. ‖

The next two lemmas prepare the ground for the case (c).

Lemma 2. *Assume Case* 2(c). *Let y be not absorbing, and $\pi(y) = \pi_j$. Then y transits one-step to x_0 with positive probability.*

Proof. Suppose not. Then player j is on the approval committee for every equilibrium transition from y. Therefore

$$V_j(y) \geq u_j(y).$$

At the same time, y is not absorbing by assumption. But then the above inequality is impossible, since $u_j(y)$ is the uniquely defined maximal pay-off for j across all coalition structures. ||

Lemma 3. *Assume Case* 2(c). *Suppose that a state x^i, with coalition structure π_i, is part of a non-degenerate recurrence class* (*starting from x^i*). *Then $V_j(x_0) = V_j(x^i)$ for all $j \neq i$.*

Proof. First, since x^i is not absorbing, by Lemma 2, x^i transits one-step to x_0 (with positive probability) and both players $j \neq i$ must approve this transition. Therefore

$$V_j(x_0) \geq V_j(x^i). \tag{7}$$

Next, consider a path that starts at x_0 and passes through x^i (there must be one because x^i is recurrent). Assume without loss of generality that it does not pass through x_0 again. If both individuals $j \neq i$ must approve every transition between x_0 and x^i, we see that $V_j(x_0) \leq V_j(x^i)$, and combining this with (7), the proof is complete.

Otherwise, some $k \neq i$ does not need to approve some transition. This can only be a transition from x_0 to a state x^k with coalition structure π_k, with subsequent movement to x^i without reentering x_0. So $V_k(x^i) \geq V_k(x^k)$. But x^k itself is not absorbing and so by Lemma 2 transits one-step to x_0 (with positive probability). By Lemma 1, $V_k(x^k) > V_k(x_0)$. Combining these two ineqalities, $V_k(x^i) > V_k(x_0)$, but this contradicts (7). ||

Case (c). We divide up the argument into two parts. In the first part, we assume that for some i, some state x^i (with coalition structure π_i) is part of a non-degenerate recurrence class. Suppose that no efficient pay-off limit exists. We first claim that

$$V_j(x_0) = V_j(x^i) = u_j(x^i) \text{ for all } j \neq i. \tag{8}$$

To prove this, consider an equilibrium path from x^i. If this path *never* passes through x_0, then it is easy to see that all three players must have their value functions monotonically improving throughout, so one-period pay-offs converge. Moreover, the limit pay-off for player i must be at the maximum, so this limit is efficient. Given our presumption that there is no efficient limit, the path does pass through x_0, so consider these alternatives:

(i) For some $j \neq i$, the path passes a state y^j (with structure π_j) before it hits x_0. Moreover, y^j is not absorbing, and so by Lemma 2 it must transit one-step to x_0 with positive probability. However, player i must approve all these moves; so $V_i(x_0) \geq V_i(y^j) \geq V_i(x^i)$. But this contradicts Lemma 1. So this alternative is ruled out.

(ii) Otherwise, the path *either* transits one-step to x_0, *or* passes through a sequence of moves, *all* of which must be approved by both players $j \neq i$. So for any one-step transition from x^i to a state y, we have

$$V_j(x_0) \geq V_j(y) \geq V_j(x^i),$$

for $j \neq i$. But by Lemma 3, $V_j(x_0)$ *equals* $V_j(x^i)$ for $j \neq i$. It follows that for every one-step transit y,

$$V_j(y) = V_j(x^i),$$

for $j \neq i$. Consequently, for each such j,

$$V_j(x^i) = (1-\delta)u_j(x^i) + \delta \int V_j(y) P(x^i, dy) = (1-\delta)u_j(x^i) + \delta V_j(x^i).$$

Using this, and $V_j(x_0) = V_j(x^i)$ for $j \neq i$, the claim is proved.

We now show that there is an efficient absorbing state, contrary to our initial presumption. Consider a state x^i (with structure π_i) to which the claim just established applies. By condition (T), there is some other state x^*, also with coalition structure π_i, such that for some $j \neq i$, $u_j(x^*) > u_j(x^i)$. Because j must approve every transition from x^*, $V_j(x^*) \geq u_j(x^*) > u_j(x^i) = V_j(x_0)$, where this last equality uses the claim. So x^* cannot have an equilibrium transition to x_0, but then i must approve *every* equilibrium transition. However, since π_i gives player i his unique maximal pay-off, he will reject every transition to a different coalition structure. Therefore, x^* is both absorbing and efficient.

For the second part of case (c), suppose now that all recurrence classes are singletons. Assume by way of contradiction that all these are inefficient. This immediately rules out all absorbing states x^i with $\pi(x^i) = \pi_i$ for some i, and it also rules out x_0.[26]

Now consider any state x^i with $\pi(x^i) = \pi_i$. Since it is not absorbing, $V_j(x^i) \geq u_j(x^i)$ for $j \neq i$. Also Lemma 2 tells us that x^i transits one-step to x_0 with positive probability, so $V_j(x_0) \geq V_j(x^i) \geq u_j(x^i)$. In particular, $V_j(x_0) \geq \max\{\bar{u}(j, \pi_i), i \neq j\}$ for all $j = 1, 2, 3$. Moreover, because π_j is maximal for j and j must approve all other transitions from x_0 (as well as to all states from the structure π_j except for x_0), we have $V_j(x_0) \geq u_j(x_0)$, so $V_j(x_0) \geq \max\{u_j(x_0), \bar{u}(j, \pi_i), i \neq j\}$ for all $j = 1, 2, 3$. Since x_0 and x^i are transient, there must be a path from x_0 to an absorbing state x_G, but this implies that any such absorbing state must satisfy $u_j(x_G) \geq \max\{u_j(x_0), \bar{u}(j, \pi_i), i \neq j\}$ for all $j = 1, 2, 3$. Therefore, x_G is efficient, contradicting our initial supposition.

Case (d). Proceed again by way of contradiction; assume there is no Markov equilibrium with efficient absorbing pay-off limit. It is immediate, then, that any state x such that $\pi(x) \in \{\pi_0, \pi_2, \pi_3\}$ is not absorbing. It also gives us the following preliminary result:

Lemma 4. *If any state x^1 with $\pi(x^1) = \pi_1$ is absorbing, then x^1 is not dominated by any state y with $\pi(y) \in \{\pi_0, \pi_G\}$.*

Proof. Suppose this is false for some x^1. It is trivial that x^1 cannot be dominated by any grand coalition state; otherwise x^1 would not be absorbing. So x_0 dominates x^1. Consider any player $j \neq 1$. From x_0, there may be a transition to z^j with $\pi(z^j) = \pi_j$, which j need not approve. She must approve *all* other transitions from x_0. Thus, along the lines of Lemma 1, we see that $V_j(x_0) \geq u_j(x_0) > u_j(x^1) = V_j(x^1)$ for $j = 2, 3$, but this contradicts the presumption that x^1 is absorbing (given the minimality of approval committees, 2 and 3 will jointly deviate). ‖

As in case (c), divide the analysis into two parts.

(i) *All recurrence classes are singletons.*
Since all absorbing states are assumed inefficient, it is clear that all absorbing states must either have coalition structure π_1 or π_G (since all states with coalition structures π_0, π_2, and π_3 are efficient). Consider x_0; it is transient. Let $\hat{x}$ be an absorbing state reached from x_0. By Lemma 2, we know that there must be a transition from any state with coalition structure π_2 or π_3 to x_0, and—because π_0 is maximal for player 1—from x_0 to some state with coalition structure π_1 with strictly positive probability. Therefore, we may conclude that

$$\begin{aligned} V_1(x_0) &\geq \max\{\bar{u}(1, \pi_2), \bar{u}(1, \pi_3)\} \\ V_2(\hat{x}) &\geq V_2(x_0) \geq \max\{\bar{u}(2, \pi_3), u_2(x_0)\} \\ V_3(\hat{x}) &\geq V_3(x_0) \geq \max\{\bar{u}(3, \pi_2), u_3(x_0)\}. \end{aligned} \tag{9}$$

This implies that $\hat{x}$ is not Pareto dominated by π_0, π_2, or π_3.
Now, if $\pi(\hat{x}) = \pi_1$, then Lemma 4, the fact that π_0, π_2, and π_3 are not absorbing, and (9) allow us to conclude that $\hat{x}$ must be efficient, a contradiction. So suppose that $\pi(\hat{x}) = \pi_G$. Note that $\hat{x}$ cannot be dominated by a state y such that $\pi(y) = \pi_1$. For $V_i(y) \geq u_i(y)$ for $i = 2, 3$. Moreover, an argument along the same lines as Lemma 1 easily shows that $V_1(y) \geq u_1(y)$. Therefore, if $\hat{x}$ were dominated by y, there would be a profitable move from $\hat{x}$, contradicting the presumption that it is absorbing. Therefore, $\hat{x}$ must be efficient, but this contradicts our assumption that no absorbing state is efficient.

(ii) *There is some non-degenerate recurrence class (and all other states are either transient or inefficient).*
Observe that analogues to Lemmas 2 and 3, and the first part of case 2(c) can be established for case (d). However, whereas in case 2(c), we were able to pin down the equilibrium pay-off of *two* players along some non-degenerate recurrence class, now we can only pin down the equilibrium pay-off of *one* player; that is, for a recurrent class that transits from x_0 to x_i, $i \neq 1$, we have: $V_j(x_0) = V_j(x_i)$ and $V_j(x_0) = u_j(x_i)$ for $j \neq 1, i$.

Observe that if, for the player j whose pay-offs we have pinned down, $u_j(x_i)$ *equals* $\bar{u}(j, \pi_i)$ *and* for the other player k who is part of the doubleton coalition with j, $V_i(x_0) \geq \bar{u}(k, \pi_i)$, the argument based on condition (T) will not go through.[27] That is, we cannot find another efficient state which one player (whose consent would be required for any transition) prefers to x_0. In this case, we must construct an equilibrium with some efficient absorbing state, and this is our remaining task.

26. If x_0 is absorbing and inefficient, then it is dominated either by a state for the grand coalition or by a state with coalition structure π_i for some i. Either way, by minimality of approval committees, x_0 will fail to be absorbing.

27. Of course, if these conditions are not satisfied, the same argument as in case 2(c) implies the existence of an efficient absorbing state.

First suppose that there does not exist a state x such that $u_i(x) > u_i(x_0)$ for $i = 2, 3$ and for all y such that $\pi(y) = \pi_1$, $u_1(x) \geq u_1(y)$.[28] In the construction of the equilibrium, the following sets of states, will be important: x_0 and

$$
\begin{aligned}
&S_2^u = \{x \mid \pi(x) = \pi_2, u_3(x) > u_3(x_0)\}\\
&S_2^d = \{x \mid \pi(x) = \pi_2, u_3(x_0) \geq u_3(x)\}\\
&S_3^u = \{x \mid \pi(x) = \pi_3, u_2(x) > u_2(x_0)\}\\
&S_3^d = \{x \mid \pi(x) = \pi_3, u_2(x_0) \geq u_2(x)\} \qquad (10)\\
&S^u = \{x \mid \pi(x) \in \{\pi_1, \pi_G\}, x \text{ efficient}\}\\
&S_1^D = \{x \mid \pi(x) \in \{\pi_1, \pi_G\}, \exists z \in \{S_2^d, S_3^d\} : \mathbf{u}(z) > \mathbf{u}(x), \exists j \in \{2, 3\} : u_j(x_0) < u_j(x)\} \setminus S_2^D\\
&S_2^D = \{x \mid \pi(x) \in \{\pi_1, \pi_G\}, \exists z \in \{S_2^u, S_3^u, S^u, x_0\} : \forall i, \mathbf{u}(z) > \mathbf{u}(x)\}.
\end{aligned}
$$

Consider the following description of strategies:

(i) For all players $i = 1, 2, 3$, from x_0 all players offer x_0 and accept a transition to another state y only if $V_i(y) > V_i(x_0)$.

(ii) From all states $x \in S_2^d \cup S_3^d$, players i such that $|S^i(x) = 2|$ propose and accept x_0, while player i such that $|S^i(x)| = 1$ proposes the status quo. An arbitrary player k accepts a transition to another state y only if $V_k(y) > V_k(x)$.

(iii) From all states $x \in S_2^u \cup S_3^u \cup S^u$, all players propose the status quo and an arbitrary player k accepts a transition to another state y only if $V_k(y) > V_k(x)$.

(iv) From all states $x \in S_1^D$, all players propose the status quo and an arbitrary player k accepts a transition to another state y only if $V_k(y) > V_k(x)$.

(v) From all states $x \in S_2^D$, all players propose a state $z(x) \in S_2^u \cup S_3^u \cup S^u \cup x_0$ and an arbitrary player k accepts a transition to another state y only if $V_k(y) > V_k(x)$. If x is dominated by x_0, we require $z(x) = x_0$.

It is easy to see that these strategies constitute an equilibrium in which the singletons are absorbing. Moreover, every other state is either absorbing itself or transits (one-step with positive probability) to some absorbing state. Note that the states in S_1^D are absorbing for δ high enough and are inefficient. The reason they are absorbing is clear: if a transition to a dominating state were allowed, there would eventually be a transition to the singletons, which, by assumption, hurts one of the players whose original consent is needed. That the strategies defined in (v) above constitute an equilibrium for δ high enough follows from arguments similar to van Damme, Selten and Winter (1990): with a finite number of states and sufficiently patient players any such absorbing state could be implemented; one player will always prefer to reject any other offer.

Now suppose that there exists a state x such that $u_i(x) > u_i(x_0)$ for $i = 2, 3$ and for some y such that $\pi(y) = \pi_1$, $u_1(x) \geq u_1(y)$. In this case, the singletons clearly cannot be absorbing for δ high enough. However, with a finite number of states one can easily construct an equilibrium with a positive probability path from x_0 to some efficient absorbing state for δ high enough. In particular, from x_0, there is a positive probability transition to a state $y \in \pi_1$; from y there is a probability 1 transition to some efficient state y' which dominates y (if such a state exists; if not, y is absorbing).[29] ‖

Acknowledgements. Discussions with Douglas Gale, Hideo Konishi, Rahul Roy, and Ennio Stacchetti were extremely valuable. We are also grateful to a Managing Editor and two anonymous referees for helpful suggestions that significantly improved the paper. Ray is grateful for the warm hospitality of the Department of Economics and STICERD at the London School of Economics, where a first draft of the paper was written. This research was supported by the National Science Foundation under grant 0241070 to Ray.

28. That is, there is no state that players 2 and 3 prefer to the singletons, which they can achieve, either directly or indirectly (by initiating a preliminary transition to the coalition structure π_1).

29. From x_0, there may also be a positive probability transition to some other state z. However, if $\pi(z) \in \{\pi_2, \pi_3\}$ it is clearly efficient since at these states players 2 and 3 obtain their unique maximum. Moreover, for δ high enough, it cannot be that $\pi(z) = \pi_G$, since then this would imply that z Pareto dominates y'.

REFERENCES

AGHION, P., ANTRAS, P. and HELPMAN, E. (2004), "Negotiating Free Trade" (Mimeo, Department of Economics, Harvard University).

ANDREONI, J. and MILLER, J. (2002), "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica*, **40**, 737–753.

ASH, R. B. (1972) *Real Analysis and Probability* (Orlando: Academic Press).

AUMANN, R. and MYERSON, R. (1988), "Endogenous Formation of Links between Players and of Coalitions: An Application of the Shapley Value", in A. Roth, (ed.) *The Shapley Value: Essays in Honor of Lloyd Shapley* (Cambridge: Cambridge University Press) 175–191.

BLOCH, F. (1995), "Endogenous Structures of Association in Oligopolies", *Rand Journal of Economics*, **26**, 537–556.

BLOCH, F. (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division", *Games and Economic Behavior*, **14**, 90–123.

BLOCH, F. and GOMES, A. (2006), "Contracting with Externalities and Outside Options", *Journal of Economic Theory*, **127**, 172–201.

BLOCH, F., SÁNCHEZ-PAGÉS, S. and RAPHAEL, S. (2006), "When Does Universal Peace Prevail? Secession and Group Formation in Conflict", *Economics of Governance*, **7**, 3–29.

CHARNESS, G. and GROSSKOPF, B. (2001), "Relative Payoffs and Happiness: An Experimental Study", *Journal of Economic Behavior and Organization*, **45**, 301–328.

CHARNESS, G. and RABIN, M. (2002), "Understanding Social Preferences with Simple Tests", *The Quarterly Journal of Economics*, **117** (3), 817–869.

CHATTERJEE, K., DUTTA, B., RAY, D. and SENGUPTA, K. (1993), "A Noncooperative Theory of Coalitional Bargaining", *Review of Economic Studies*, **60**, 463–477.

CHWE, M. (1994), "Farsighted Coalitional Stability", *Journal of Economic Theory*, **63**, 299–325.

ESTEBAN, J. and RAY, D. (1999), "Conflict and Distribution", *Journal of Economic Theory*, **87**, 379–415.

ESTEBAN, J. and SÁKOVICS, J. (2004), "Olson vs. Coase: Coalitional Worth in Conflict", *Theory and Decision*, **55**, 339–357.

GOMES, A. and JEHIEL, P. (2005), "Dynamic Processes of Social and Economic Interactions: On the Persistence of Inefficiencies", *Journal of Political Economy*, **113**, 626–667.

HARRIS, C. (1985), "Existence and Characterization of Perfect Equilibrium in Games of Perfect Information", *Econometrica*, **53**, 613–628.

HERRERO, M. J. (1985), "N-Player Bargaining and Involuntary Unemployment" (Ph.D. Dissertation, London: London University).

KALANDRAKIS, A. (2004), "A Three-Player Dynamic Majoritarian Bargaining Game", *Journal of Economic Theory*, **116**, 294–322.

KONISHI, H. and RAY, D. (2003), "Coalition Formation as a Dynamic Process", *Journal of Economic Theory*, **110**, 1–41.

KRISHNA, P. (1998), "Regionalism vs Multilateralism: A Political Economy Approach", *Quarterly Journal of Economics*, **113**, 227–250.

LEVY, P. (2004), "A Model of Political Parties", *Journal of Economic Theory*, **115** (2), 250–277.

OKADA, A. (2000), "The Efficiency Principle in Non-Cooperative Coalitional Bargaining", *Japanese Economic Review*, **51**, 34–50.

OSBORNE, M. and RUBINSTEIN, A. (1994) *A Course in Game Theory* (Cambridge, MA: MIT Press).

RAY, D. (1989), "Credible Coalitions and the Core", *International Journal of Game Theory*, **18**, 185–187.

RAY, D. and VOHRA, R. (1997), "Equilibrium Binding Agreements", *Journal of Economic Theory*, **73** (1), 30–78.

RAY, D. and VOHRA, R. (1999), "A Theory of Endogenous Coalition Structures", *Games and Economic Behavior*, **26**, 286–336.

RAY, D. and VOHRA, R. (2001), "Coalitional Power and Public Goods", *Journal of Political Economy*, **109**, 1355–1384.

RUBINSTEIN, A. (1982), "Perfect Equilibrium in a Bargaining Model", *Econometrica*, **50**, 97–109.

SEIDMANN, D. J. (2005), "Preferential Trading Arrangements as Strategic Positioning" (Mimeo, School of Economics, University of Nottingham).

SEIDMANN, D. J. and WINTER, E. (1998), "Gradual Coalition Formation", *Review of Economic Studies*, **65**, 793–815.

THRALL, R. M. and LUCAS, W. L. (1963), "N-Person Games in Partition Function Form", *Naval Research Logistics Quarterly*, **10**, 281–298.

VAN DAMME, E., SELTEN, R. and WINTER, E. (1990), "Alternating Bid Bargaining with a Smallest Money Unit", *Games and Economic Behavior*, **2**, 188–201.