Contests over Political Authority*

Catherine Hafer†

this version: 04/22/2007

Abstract

This paper analyzes a model of endogenous political authority in which authority may be established by force through a standoff. Two players have a mix of common and contrary interests; the resolution of the dispute is required to be self-sustaining, i.e. there is no external enforcement of agreements; and the players are uncertain about each other's resolve, i.e., about the relative strength of their interests in one outcome over another. An equilibrium solution of the model provides insights into the duration of the contest over authority, its ultimate outcome, and the conditions under which a peaceful resolution is possible. I show that neither cheap talk nor impartial mediation promote peaceful resolutions or enhance efficiency. Costly signaling reduces the incidence of conflict but may consume more resources than the conflict it obviates.

*Comments Welcome. The preliminary draft of this paper was written when I was a visitor at CMS-EMS at the Kellogg School of Business, Northwestern University. I benefitted from comments from participants in the seminars at Kellogg, Stanford GSB, and NYU, and at the Conference on Political Institutions and Economic Policy (Princeton), and Conference on Conflict and Cooperation (Kellogg) and especially from David Austen-Smith, Sandeep Baliga, David Baron, Bruce Bueno de Mesquita, Ethan Bueno de Mesquita, Ernesto Dal Bó, Eric Dickson, Avinash Dixit, Tim Feddersen, Sven Feldmann, Sandy Gordon, Bard Harsted, Dimitri Landa, Adam Meirowitz, and Alastair Smith.

†Assist. Prof. of Politics, NYU, e-mail: catherine.hafer@nyu.edu

1 Introduction

This paper analyzes a model of contested political authority: a situation in which two disputants who wish to agree ultimately on a single (joint) course of action each insist on different courses of action, in spite of the fact that doing so is immediately costly, in the hopes that their opponent will give in first. The model captures a number of features intrinsic to such situations: the players have some degree of common interest, so that they may benefit from coordinating their actions (peaceful settlement is, all else equal, better for each player); they have some degree of contrary interests (each player prefers to be the source of authority given the benefits entailed), so that the dispute between them is not trivial; and the players are uncertain about each other's resolve, i.e., about the relative strength of their interests in one outcome over another. Finally, because the contest is for the prize of political authority, its resolution must be self-sustaining, i.e. there is no external enforcement of agreements.

The first two features - a mix of common and contrary interests - are ubiquitous, and characterize both political disputes, in which people with different policy preferences benefit from coordinating on one public policy, and common economic ones, in which two firms that may profit from trade have opposing preferences over the division of such profits between them. In the case of two firms that wish to do business, many divisions of the profits are possible, in part because the firms can typically rely on some form of third-party contract enforcement to insure that they abide ex post by whatever agreement they make ex ante. However, when the contest in question is for political authority, we cannot assume that the agreement - if it were reached - would be enforced by an already existing common authority, and so the feasible final resolutions must include only those agreements to which both parties could credibly commit to carrying out. The requirement that final resolutions be credible in this way implies a considerable limitation on the set of viable agreements, ruling out a variety of arrangements that otherwise might have been possible.

The requirement that the settlement of authority be self-enforcing also implies a constraint on the analytical model that could be used to capture the underlying distributive game. In particular, it rules out the class of standard bargaining models, which, in effect, assume perfect enforcement of the negotiated outcomes. Indeed, one of the key findings is that in the contest over political authority there may not be enough flexibility – enough room for given and take – to achieve a resolution of the dispute through bargaining. Rather, the parties may find themselves in a costly standoff, each insisting (at least initially) on a different outcome and hoping that the other will be the first to give in.

The conjunction of these factors alone is not sufficient to produce a standoff, however, because if it is common knowledge between the disputants which one of them will ultimately win, the other will concede immediately. Thus a standoff necessarily requires that the participants have some ongoing uncertainty about which of them is willing to hold out longer for her more-preferred outcome, because each of them must have some hope of winning in order to continue the dispute. A model of standoff should, then, incorporate both the possibility of conflict under uncertainty and the possibility of immediate resolution when that uncertainty is removed.

The model presented in this paper is based on a dynamic extension of an incomplete-information asymmetric coordination game. This game captures both the common and the contrary aspects of the interests of parties contesting political authority. After introducing this game, I examine the equilibrium that corresponds to a standoff, i.e., each player insisting upon its own most preferred outcome by taking the action that corresponds to it until one of the players becomes convinced that he is less resolute than the other. This equilibrium may be thought of as corresponding to a war of attrition. (Note that the continuous-time asymmetric coordination game and the war of attrition are not strategically equivalent: in the former, the players may repeatedly change their actions, which is ruled out by assumption in the war of attrition.) Because in this equilibrium play ultimately reveals the type of the less resolute player, and produces the long-term outcome preferred by the more resolute player, neither player has an incentive to rekindle the dispute– the resolution of the dispute is self-sustaining. I then go on to examine the properties of the equilibria in which play reverts to a standoff off the equilibrium path - equilibria that incorporate the absence of (pre-existing) authority and its endogenous determination by contest.

An additional advantage of the model is that, as in a war-of-attrition game, the complete-information version has an asymmetric equilibrium in which the weaker player concedes immediately to the stronger one. I exploit this fact to examine the extent to which mediated and unmediated cheap-talk communication, as well as costly signaling, can be used successfully to end or to avert a standoff as a function of features of the strategic environment.

The key results of the model are, on the whole, negative. I show that peaceful stationary equilibria do not exist generically and that peaceful non-stationary equilibria exist only under limited conditions. Neither cheap-talk communication nor mediation lead to improvements in social welfare. Costly signaling by burning money at the begining of the game can reduce the incidence of conflict, but often at a price greater than the benefit of avoiding conflict.

The paper proceedes as follows. In Section 2, I provide a brief discussion of the relation between the model and the existing literature. Section 3 presents a symmetric equilibrium of the basic model that corresponds to a standoff. Section 4 analyzes the possibility of peaceful equilibria and their efficiency in the presence of cheap-talk, impartial mediation, and costly signaling. Section 5 concludes with a brief discussion. The formal proofs of the results are in the Appendix.

2 Relation to Literature

The model presented below can be seen as a natural complement to the model of political compromise by Dixit, Grossman, and Gul (2000).[1] In the latter, the party in power in a given period is assumed to determine unilaterally the division of benefits between the two parties, and the focus is on the dynamic evolution of the allocation of spoils between the in- and out-of-power parties. In particular, Dixit, Grossman, and Gul ask what arrangements for sharing the political and economic benefits can be supported in equilibrium, and how those arrangements depend on exogenously given probabilities of each party obtaining and retaining office. In contrast, in the model presented below the likelihood of each party's victory is determined by the parties' choices, which depend in equilibrium on their privately known valuations of the possible outcomes, and in each period the outcome is jointly determined by the parties' actions. By assumption, neither party can unilaterally obtain its most preferred outcome for even one period, and preventing an opponent from obtaining their most preferred outcome imposes lower payoffs on both parties, in essence destroying part of the pie. In this model, the focus is on establishing political authority in an environment in which it is initially absent, by force if necessary, rather than on managing its exploitation where it is already well established and can be peacefully transferred from party to party.

The model in this paper is related to two branches of the recent work on appropriative conflict. In the economics literature, Dixit (2004) presents a recent state of the art discussion of the approaches to analyzing economic activity in a decentralized legal and political environment. Hafer (2006) is a model of the state of nature in which parties engage in dynamic continuous-time wars of attrition over a productive resource. In political science, the closely related work is on crisis bargaining games (e.g., Banks 1990, Bueno de Mesquita, Morrow, and Zorick 1997, Powell 2004), in

[1] For a related model, see also Alesina (1988).

which parties attempt to negotiate a resolution to a conflict under the threat of either party's unilateral initiation of a war, and one or both parties are uncertain about the payoffs associated with war.

In the model presented here, the parties are not only uncertain about the value of going to war (because of their uncertainty about their opponents' types), they are also uncertain of their opponents' valuation of the settlement. Both the payoffs to war and the payoffs associated with a given settlement depend on aspects of the player's type, though in different ways. This fact has important implications for the viability of a negotiated long-term resolution in which one player is asked to accept the other's authority, because a player's continued willingness to accept the authority of the other player (in spite of its own changing beliefs about its opponent's type) reveals information about its own type in turn. Furthermore, the fact of that information revelation and its potential consequences for subsequent renegotiation or conflict may limit the players' ability to reach a negotiated settlement (Nalebuff 1987). In this sense, peace under a negotiated settlement is an informative process, just as conflict is.

Prominent political economy applications of incomplete information war of attrition include models of delayed fiscal stabilization (Alesina and Drazen 1991, Drazen and Grilli 1993, Casella and Eichengreen 1996, Spolaore 2004). Relative to these models, the basic model in Section 3.1 generalizes to allow private information over all payoffs, while retaining the potential for selecting inefficient long-run outcomes. The results in the subsequent sections may be interpreted as identifying the extent to which these inefficiencies survive in the presence of cheap-talk communication - a common feature of the empirical settings relevant to these models.

The analysis of the cheap-talk extension of the basic model constitutes a contribution to the literature on cheap-talk games with uncertainty over payoffs, which is mostly of recent vintage (see e.g., Ben-Porath 2002 and Baliga and Morris 1998, 2002). Baliga and Morris (1998, 2002) consider the conditions that guarantee the impossibility of information transmission that would be equilibrium payoff-relevant in the subsequent game. They show that if the game satisfies two general conditions, then, in any equilibrium, all types of each player are indifferent between any of the equilibrium messages sent by any type of that player (though such equilibria may still be informative). These conditions are *type independence* and *common induced preferences*. A game has common induced preferences if, for any pair of a player's possible types, the two types have the same preferences over any pair of distributions of the player's opponent's actions. The game considered above satisfies type independence and violates common induced preferences. While all types prefer the opponent to drop out sooner rather than later, not all types will necessarily have the same preferences over a pair of distributions of stopping times such that in one, both early and late stopping times are more likely than they are in the other.

The results can be instructively compared to those of Banks and Calvert (1992), who ask whether efficient outcomes can be obtained through mediated or unmediated cheap talk in a static incomplete-information battle-of-the-sexes game. They find that, while mediated cheap talk can improve efficiency, unmediated cheap talk cannot. I find that unmediated cheap talk is also fruitless in the dynamic game and, relative to their results for the static game, efficiency gains through mediation are more difficult to sustain. Banks and Calvert also show that, for a one-shot (discrete-time) incomplete-information Battle of the Sexes, there is always an incentive-compatible mechanism that is ex ante efficiency improving compared to the symmetric equilibrium of the game without communication. Baliga and Sjostrom (2004) obtain a non-monotonic cheap-talk equilibrium strategy in a game with some common interest. Neither of these results survives in the model analyzed here.

3 Asymmetric Coordination and Standoff

3.1 The Model

Suppose that two players face an asymmetric coordination problem in continuous time with incomplete information. In particular, suppose that each player has two actions, $\{X, Y\}$, and at every point in time each player must take one of these actions. The matrix below shows the payoff per unit of time for each player from each of the four possible action profiles. Suppose that it is common knowledge that player 1 (weakly) prefers outcome (X, X) to (Y, Y) and strictly prefers (Y, Y) to either (X, Y) or (Y, X). Player 2 prefers (Y, Y) to (X, X), but she also prefers either of these outcomes to either of the others. Suppose, however, that each player is uncertain of the precise payoffs of her opponent (a_i, b_i), and that they have a common prior over the distribution of types. The payoff matrix below corresponds to that of an incomplete information battle-of-the-sexes game, but with the caveat that, because the game takes place in continuous time, the payoffs it shows are flows rather than lump sums. Each player's payoff for the entire infinite-horizon game can be evaluated at any moment in time as the discounted present value of all future payoffs, given the common discount rate r. Throughout the paper, utility is assumed to be non-transferable, though I consider the robustness of the results to allowing transfers in the Discussion.

Player 1 \ Player 2	X	Y
X	a_1, b_2	$0, 0$
Y	$0, 0$	b_1, a_2

$a_i \geq b_i > 0$, $i \in \{1, 2\}$

This dynamic extension of the battle-of-the-sexes game can be understood to be a generalization of the war of attrition, in which the parties are allowed to choose to play X or Y at any point in the game regardless of their past choices. In contrast, in a war of attrition, once a player has "quit," she is assumed to be committed to that choice for the duration of play. While the war of attrition captures all the features of a contest over authority identified above, it arbitrarily limits the opportunities to resolve disputes over authority peacefully in that players cannot take the actions preferred by their opponents without giving up their ability to threaten their opponents with the use of force, and thus giving up the possibility of ever obtaining their own more preferred action profile. Insofar as a player is more willing to subject herself to the limited authority of another (though that limit may be imposed by her own ability and willingness to use force) than to an absolute authority, the dynamic battle-of-the-sexes game better captures the potential for the creation and maintenance of political authority.

3.2 Symmetric Standoff Equilibrium

Suppose that at the beginning of the game, each player must choose an action, X or Y. Actions are observable, and at any moment, either player can switch to the other action. If each player begins by taking the action that corresponds to her own preferred outcome, i.e. 1 chooses action X and 2 chooses action Y, then each player's strategy can be described by the time at which she switches, conceding victory to her opponent. (If a player chooses the action that corresponds to the other player's preferred outcome at the start, then she can be said to "switch" at time $t = 0$.) At each moment in time, a player obtains additional information about her opponent's type from the fact that her opponent has not yet conceded.

The players' utilities can be expressed as the discounted present values of their future streams of payoffs. If player i quits at t_1, while player j stays in, the discounted present value of player i's payoff at $t = 0$ is

$$u_i(t_1, t_2 > t_1) = \int_{t_1}^{\infty} b_i e^{-rt} dt = \frac{b_i}{r} e^{-rt_1}.$$

Similarly, if player j quits at t_2, while player i stays in, the discounted present value of player i's payoff at $t = 0$ is

$$u_i(t_1, t_2 < t_1) = \frac{a_i}{r} e^{-rt_2}.$$

The value of the prize (won at time 0) is the difference between the present value of the infinite repetition of the preferred action profile and of the opponent's preferred action profile, $(a_i - b_i) \int_0^\infty e^{-rt} dt$. The cost of spending another unit of time resisting is the difference between the payoff for that unit of time in discordant play and the payoff for that unit in the opponent's preferred equilibrium, b_i.

The solution concept is Perfect Bayesian Equilibrium, which requires that at each moment in time, a player choose optimally to switch to the other action or to continue, given her beliefs; and that at each moment in time, she update her beliefs about her opponent's type via Bayes' Rule based on the fact that her opponent has not yet conceded.

The first result provides a characterization of this symmetric equilibrium outcome, which I will hereafter call the standoff equlibrium:[2]

Theorem 1 *The game has a symmetric equilibrium which*

(1) in the long run selects the preferred outcome of the player i with the higher value of the ratio $\theta_i := \frac{a_i}{b_i}$

(2) has a standoff of duration $s^*(\theta_i) = \frac{1}{r} \int_{\theta_l}^{\theta_i} (\theta - 1) \frac{p(\theta)}{1 - P(\theta)} d\theta$, where $\theta_i = \min\{\theta_1, \theta_2\}$.

Proof. See Appendix. ■

The argument of the proof proceeds by exploiting the possibility of collapsing the two dimensions of players onto a single dimension induced by the ratio of the benefit

[2]Note that, as in all incomplete-information wars of attrition, there are also two asymmetric equilibria: player 1 "fights forever" and player 2 "quits immediately;" and 1 "quits immediately" and 2 "fights forever." Recall that, because the players' types are *not* common knowledge, their types cannot be invoked to select one equilibrium over the other.

of winning over the opportunity cost of continuing the standoff, θ. This symmetric equilibrium of the game can, then, be characterized as the symmetric equilibrium of the incomplete-information continuous-time war of attrition between opponents defined by their type θ_i.

The standoff equilibrium predicts that the player who has a higher value on that dimension will ultimately prevail, and the coordination outcome that she prefers will be the long-run outcome. Once a player has conceded, neither player has an incentive to renew the dispute by changing her choice of action again; the player who caved has revealed her type and, more to the point, she has revealed that she is of a less resolute type than the winner. Her optimal action, then, is to continue playing the action corresponding to the victor's preferred outcome, and the resolution of the dispute is final.

4 Equilibria without Conflict

4.1 The Selection Argument

Having established the (equilibrium) possibility of creating political authority through force, it is natural to ask if it is possible to obtain permanent peaceful outcomes under the threat of such conflict. More precisely, we will say that an equilibrium is *peaceful* if for at least some pair of types (a_i, b_i) and (a_j, b_j) such that $a_i > b_i$ and $a_j > b_j$, neither action profile (X, Y) nor (Y, X) occurs on the path of play[3]. Note that this notion of peaceful equilibrium is intentionally very weak to strengthen the (mostly negative) key results of the model.

In order to examine the possibility of peacefully establishing (at least limited) political authority where it does not already exist, I restrict attention in what follows to the subset of equilibria that induce the relevant standoff equilibrium (for symmetric

[3]I require $a_i > b_i$ and $a_j > b_j$ to insure that the peace is not that which obtains trivially whenever one of the players has the same payoff from (X, X) and (Y, Y).

or asymmetric beliefs) in subgames off the path of play. In other words, I exclude from the analysis instances of peace that are supported by a pre-existing political authority off the path of play.

Specifically, I restrict attention to equilibria with the following set of features:

1. (*Reversion to Standoff*) In every off-path-of-play subgame in which it is not common knowledge which player is more resolute (i.e. in which it is not common knowledge which player has higher θ), the players play standoff equilibrium strategies.

2. (*Deviations are Trembles*) At the first information set following a deviation, beliefs are the same as those at the point of deviation; and at all subsequent information sets, beliefs are derived from those beliefs and strategies.

3. (*Common Knowledge Peace*) At any information set at which it is common knowledge that $\theta_i > \theta_j$, the players choose i's preferred action pair.

Because the standoff strategies characterized in Theorem 1 constitute an equilibrium of the game, they also constitute an equilibrium of any strategically equivalent subgame, i.e. any subgame in which the players are uncertain of each other's type and share the same prior. A symmetric standoff equilibrium also exists in subgames in which the two players have different beliefs about the distribution of their opponent's type, and again the players' strategies are increasing in θ.[4] Thus equilibria that meet these criteria do exist.

4.2 Equilibria without Communication

We will first ask whether it is possible, on the equilibrium path of play, for one player to be the sole effective political authority and always obtain her most preferred

[4]For the proof of the existence and properties of a symmetric equilibrium in an incomplete-information continuous-time war of attrition with asymmetric beliefs, see Hafer (2006).

outcome. Consistent with the standard usage, we say that equilibria are *stationary* if each player's actions are constant across time on the path of play.

Theorem 2 *1. There is no peaceful stationary equilibrium.*

2. Peaceful non-stationary equilibria exist if and only if

$$\frac{1+\bar{\theta}}{2\bar{\theta}} \geq P(\theta_{\min}) + \int_{\theta_{\min}}^{\bar{\theta}} p(\theta)e^{-rs^*(\theta)}d\theta. \tag{1}$$

3. Condition (1) is satisfied if $p(\theta)$ is increasing and sufficiently convex.

Proof. See Appendix. ■

Because the implementation of power-sharing (non-stationary) mechanisms is often complicated by factors that are outside the present model - e.g., commitment problems associated with the exclusive access to government resources once in power - Theorem 2 is a rather grim result.

The condition for the existence of a peaceful equilibrium is obtained from the conditions under which each type of player will prefer (at every point in time) to switch back and forth at equal time intervals between her most favored action profile and that of her opponent, i.e. between (X, X) and (Y, Y), rather than to initiate a standoff. If any type of player preferred to deviate and thereby initiate a standoff, the incentive to alternate between (X, X) and (Y, Y) would unravel and all types would prefer to initiate a standoff. Thus either there exists an equilibrium in which all possible pairs of types abstain from conflict on the path of play (under the condition specified) or there is no peaceful equilibrium of any kind.

Note that modifying this strategy profile to allow the playing of one action profile for a longer period than the other would necessarily make the incentive compatability constraint more stringent for the less favored player, and would consequently reduce the range of parameter values for which such an equilibrium could be sustained (relative to the equilibrium with equal time intervals). Likewise, any strategy profile that leads to the playing of (X,Y) or (Y,X) with positive probability could not be sustained for as wide a range of parameter values. The existence of a peaceful equilibrium relies on the ability of the players to switch between action profiles sufficiently quickly to ensure the willingness of each player to wait through her less favored action profile in order to obtain the higher payoff of her more favored one.

Condition (1) is satisfied if the probability density function $p(\theta)$ is increasing and sufficiently convex, or, more generally, if the probability density of the lowest types θ is sufficiently small. Intuitively, a peaceful equilibrium exists only if a standoff is sufficiently costly, in expectation, which is the case when low types (who would quit quickly) are sufficiently rare.

4.3 Unmediated Cheap-talk

Recall that the players' uncertainty about each other's types was a necessary ingredient of standoffs. We may wonder, then, if allowing the players to communicate can reduce their uncertainty and affect the duration or occurrence of a standoff. In this subsection I analyze the effects of introducing the possibility of unmediated cheap talk on the standoff equilibrium analyzed above. In *the cheap-talk extension of the game*, players i and j can simultaneously send and observe any of a rich set of messages at any point in the game. Let m_i^t and m_j^t be the messages that i and j, respectively, send at time t. The messages themselves have no direct effect on the players' utilities, although they may, in principle, have an indirect effect via changes in their opponent's beliefs and behavior.

The next theorem shows that unmediated cheap-talk in the cheap-talk extension of the game yields no efficiency gains. I proceed by showing that this holds with respect to the possibility, first, of *monotonic* and, then, of *non-monotonic* informative equilibria. An equilibrium is monotonic if the agents' cheap-talk strategies are monotonic in their type, i.e., for an n-message equilibrium, there exist $n-1$ cutpoints in the type space such that no two types that are separated by any cutpoint prefer to send the same message. An equilibrium is non-monotonic if agents' cheap-talk strategies are not monotonic in their type. I employ the equilibrium selection criteria introduced in 4.1.

Theorem 3 *There is no informative cheap talk in any standoff equilibrium.*

Proof. See Appendix. ■

The key intuition for this theorem is that all types always wish to appear to be as high a type (as resolute in the standoff) as possible, in order to convince their opponent to quit sooner. Because each player can quit at any time, the costs of initially failing to coordinate can be made arbitrarily small by either player, acting unilaterally. Hence the benefits associated with the possibility of getting one's more-preferred outcome by claiming to be a high type always outweigh the benefits of immediate coordination.

4.4 Mediated Cheap-talk

Having established that cheap talk alone cannot affect the occurrence or duration of standoffs, the natural next step is to consider mediation as a potential means of shortening or eliminating them. Consider augmenting the game with a pre-play communication stage in which each player communicates privately with a mediator, who recommends a course of action. Assume that all messages, e.g. claims about one's type, are unverifiable, and that the adoption of the mediator's recommended course of action is strictly voluntary. The main result–that mediation is an ineffective means of eliminating conflict and thereby achieving more efficient outcomes–is obtained by application of the Revelation Principle (Myerson 1985), which states that any outcome associated with equilibrium behavior in any such augmentation of a game is also associated with some *incentive-compatible direct mechanism* (ICDM) based on the same game. A *direct mechanism* consists of a communication protocol and a decision rule for the mediator. The pre-play communication stage has the following form: first, each player sends a private message to the mediator, who is an impartial non-strategic actor, and the set of messages available to each player is identical to the set of possible types. The mediator then privately recommends a strategy to each player, based on their messages and the mediator's commonly known decision rule. A mechanism is said to be *incentive-compatible* if truthfully revealing their types (to the mediator) and adopting the recommended strategies constitutes equilibrium behavior.

As above, we say that an ICDM is stationary if its recommended strategies are stationary, i.e., each player takes the same action, X or Y, at every point in time, independent of the history of play and of calendar time.

Theorem 4 *There is no stationary peaceful incentive-compatible mechanism. If there is no peaceful equilibrium in the baseline game, then there is no peaceful incentive-compatible mechanism.*

Proof. See Appendix. ■

A stationary peaceful incentive compatible mechanism cannot exist in this context because, unless the player who is asked to concede receives information indicating that her opponent is, in fact, a more resolute type than she had previously thought, she will prefer to initiate a standoff unilaterally, since her expected utility from standoff must always be greater than that which she could obtain by simply conceding to her opponent. If the mediator does provide information indicating that her opponent is resolute, than every type of player will have an incentive to misrepresent herself to the mediator as a very resolute type. There is no peaceful incentive-compatible non-stationary mechanism if there is no peaceful nonstationary equilibrium of the baseline game because, in order to elicit compliance under such conditions, the mechanism would have to favor higher types systematically over lower ones. But in that case, the players again have an incentive to misrepresent themselves as being higher types than they are.

4.5 Burning Money

Given that players have over-riding incentives to misrepresent their interests when signaling is costless, it is natural to ask if making communication costly produces the information revelation necessary to avoid conflict. In *the burning-money extension of the game*, players i and j can simultaneously and publicly burn money before playing the baseline game. Let v_i and v_j be the amounts that i and j, respectively, choose to burn. The amount that each player burns is subtracted from her utility. Once again, I use the equilibrium selection criteria introduced in Section 4.1.

Theorem 5 *There exists a unique equilibrium that is semi-separating in the burning of money. In that equilibrium:*

1. *separation reduces the incidence of conflict; and*

2. *conflict may occur on the path of play even when players burn unequal amounts of money.*

Proof. See Appendix. ■

Although there is an informative equilibrium of the burning money game, the information revealed is not enough to obviate the need for conflict in all instances because, although it reveals the value of a function of the players' types, that function is not the same one that governs their success in the standoff equilibrium. Thus burning money alters beliefs and thus behavior in the standoff equilibrium, but it can leave enough uncertainty about which player has the higher value of θ to necessitate a standoff of positive duration in order to resolve the dispute. However, because θ and the amount of money burned in equilibrium are positively correlated, if the difference between the amounts of money burned by the two players is sufficiently large, it does reveal which player has the higher value of θ.

The possibility of burning money as a signal of resolve has another interesting caveat worth noting. Even in those instances in which it obviates the need for conflict entirely, burning money does not necessarily enhance efficiency. To see this, consider the pair of a type with high θ and high v and a type with low θ, e.g., $\theta_{\min}+\varepsilon$, and low v. The war of attrition (without the possibility of burning money) last only as long as the player with the lower θ is willing to fight, which in the case of $\theta_{\min}+\varepsilon$ can be made arbitrarily close to zero as $\varepsilon \to 0$. However, in the game with burning money, the player with the higher type burns a correspondigly large amount of resources even though he is, in fact, facing a much less resolute type.

5 Discussion

The general tenor of the results in this paper suggests that establishing authority endogenously - even in the circumstances with some common interest - is a relatively costly project, and its costs are difficult to mitigate without recorse to a pre-existing or external authority. One of the problems with such a recourse, though, is that it is often non-neutral with respect to the outcome of standoff.[5]

It bears noting that throughout, we have assumed that contestants' utility is non-transferable. Given the focus of the paper on the endogenous establishment of authority in the absence of the pre-existing authority that could dictate the outcome of contestants in the first place, there are susbtantive reasons for assuming non-transferable utility. In particular, the possibility of transfers appears to presuppose the ability to write contracts that rely on third parties for their enforcement - an asusmption that would be contrary to the spirit of the present study.

Still, as an issue of robustness, one might wonder how the possibility of transfers would affect the main results - in particular, the impossibility of stationary peaceful equilibria. A companion paper to the present paper (Hafer 2007) shows that an

[5] The following example provides an illustration. Suppose there are two (a,b) types, (11,6) and (7,4), and the common disagreement payoff increases from 0 to 2. This would flip the ordering of these types in the θ space: with the common disagreement payoff of 0, $\theta(11,6) = 1\frac{5}{6} > \theta(7,4) = 1\frac{3}{4}$ but with the common disagreement payoff of 2, $\theta(11,6) = 2\frac{1}{4} < \theta(7,4) = 2\frac{1}{2}$. From Theorem 1, changing the order of the players in θ-space changes the outcome.

analogous result obtains for transferable utility. The key intuition is that accepting transfers in exchange for playing a less preferred action profile signals irresoluteness, and, furthermore, the continued willingness to accept that deal indicates greater and greater irresoluteness over time. Eventually, the payer realizes that his opponent is sufficiently easy to defeat and prefers to revert to the war of attrition.

Finally, the present results may be usefully compared to the much more positive results obtained by Jackson and Sonnenschein (2007). They consider an environment in which multiple agents face a series of identical collective decision problems in which they have conflicting interests and private information about the strength of their own preferences over outcomes and identify an incentive compatible mechanism that implements efficient (in the limit) outcomes by linking the decision problems through endogenously determined restrictions on the agents' sets of possible messages. The crucial difference between their environment and the one examined above is that, in their model, each agent's preferences are a new draw for each problem, whereas here we have assumed that preferences are fixed across periods. Fixing preferences makes it physically impossible to reveal information consistently across periods in the Jackson-Sonnenschein mechanism and it undermines the incentive to do so, since what a player's behavior reveals about her type in one period can be used against her in future periods. The truth is, no doubt, usually in the middle: in most applications, one would expect type to be neither completely persistent nor completely independent across periods. The stark difference between their results and those obtained here suggests the value of exploring the case of imperfectly persistent types.

6 Appendix

Theorem 1

Proof. Letting $p_j(t)$ be the probability density of opponent j's quitting time, the expected payoff for player i from action t_i is

$$\begin{aligned} E[u_i(t_i)] &= (1-\Pr(t_j \leq t_i))\frac{b_i}{r}e^{-rt_i} + \int\limits_0^{t_i} p_j(t)\frac{a_i}{r}e^{-rt}dt \\ &= \frac{b_i}{r}e^{-rt_i}(1-\Pr(t_j \leq t_i)) + \frac{a_i}{r}\int\limits_0^{t_i} p_j(t)e^{-rt}dt. \end{aligned} \quad (2)$$

Then the first-order condition is

$$-p_j(t_i)\frac{b_i}{r}e^{-rt_i} + (1-\Pr(t_j \leq t_i))(-b_i e^{-rt_i}) + p_j(t_i)\frac{a_i}{r}e^{-rt_i} = 0$$

In order to solve the first-order condition, we must express $\Pr(t_j \leq t_i)$ in terms of primitives, e.g. as a probability of a type (a_j, b_j) that chooses to quite before time t_i. This requires being able to identify a type that corresponds to t, but because type is two-dimensional, multiple types may (and, in equilibrium, do) choose the same stopping time t.[6] This problem can be surmounted by re-expressing payoffs in terms of the ratio of the difference between the value of the prize and the opportunity cost of continuing the conflict, $\frac{a}{b}$. This is algebraically equivalent to dividing the first order condition by b. Let θ represent $\frac{a}{b}$. Figure 1 shows the level curves of θ in the (b, a) space. We solve for players' optimal stopping times as a function of θ.[7]

To solve for i's best response, we must be able to express the opponent's type as a function of her action, i.e. $\Theta(t_j)$. Such a function exists only if her optimal choice of action in equilibrium is strictly monotonic. To see that it is weakly monotonic, recall that, from the definition of Bayesian Nash equilibrium, the equilibrium action of type θ of player i always yields at least as high a payoff for that type of that player than does the equilibrium action of some other type. Thus, if t_i' is the equilibrium action of an agent of type (a_i', b_i'), and hence of type θ_i', and if t_i'' is the equilibrium action of an agent of type (a_i'', b_i''), and hence of type θ_i'', we have the following two

[6] In the game with non-transferable utility, one can normalize b_1 and b_2 without loss of generality, reducing players' types to one dimension. However, later results consider environments in which transfers are possible, requiring that the players' utilities be expressed in terms of a common scale and thus requiring two-dimensional types to maintain generality.

[7] Note that once the problem is expressed in terms of θ, it constitutes a variation of the incomplete information war of attrition in Fudenberg and Tirole (1991, pp. 216-19) with discounting, and it can be solved in the same way.

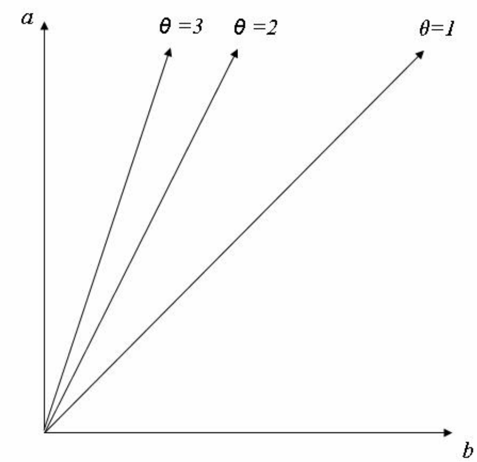

Figure 1: Level curves of θ.

inequalities:

$$\frac{b_i'}{r}e^{rt_i'}(1-\Pr(t_j<t_i'))+\frac{a_i'}{r}\int_0^{t_i'}p_j(t)e^{-rt}dt$$

$$\geq \frac{b_i'}{r}e^{rt_i''}(1-\Pr(t_j<t_i''))+\frac{a_i'}{r}\int_0^{t_i''}p_j(t)e^{-rt}dt;$$

$$\frac{b_i''}{r}e^{rt_i''}(1-\Pr(t_j<t_i''))+\frac{a_i''}{r}\int_0^{t_i''}p_j(t)e^{-rt}dt$$

$$\geq \frac{b_i''}{r}e^{rt_i'}(1-\Pr(t_j<t_i'))+\frac{a_i''}{r}\int_0^{t_i'}p_j(t)e^{-rt}dt.$$

Multiplying each inequality by r and dividing by b_i' and b_i'' respectively, these inequalities can be expressed in terms of θ_i' and θ_i'', respectively. Because these inequalities are of the same sign, the sum of their right-hand sides must be greater than the sum of the left-hand sides. Collecting terms and reducing, we have, then, the following true inequality:

$$\left(\frac{a'}{b'} - \frac{a''}{b''}\right) \int_0^{t'_i} p_j(t)e^{-rt}dt \geq \left(\frac{a'}{b'} - \frac{a''}{b''}\right) p_j(t)e^{-rt}dt.$$

It follows, then, that if $\frac{a'}{b'} > \frac{a''}{b''}$, then $t' \geq t''$. To see that the equilibrium strategy must be strictly increasing and continuous in $\frac{a}{b}$, suppose instead that there is an interval of values of $\frac{a}{b}$ that choose to stop at time t, i.e. that $\Pr(t_j = t) > 0$. Then there must be some $\varepsilon > 0$ such that $\Pr(t_i \in (t - \varepsilon, t)) = 0$. But then $t - \varepsilon$ strictly dominates t for player j. Hence the equilibrium strategy must be strictly increasing in type $\theta := \frac{a}{b}$. Exploiting the symmetry of the game (i.e. the common prior), part 1 is established.

Let $s(\theta)$ be time chosen by type θ, i.e., θ's strategy. Because $s(\theta)$ is continuous and strictly increasing, the inverse of $s(\theta)$, $\Theta(s)$, exists and is continuous. The optimization problem may then be written

$$s_i \in \arg\max \left[\frac{1}{r}\left((1 - P(\Theta_j(s_i)))e^{-rs_i} + \frac{a_i}{b_i}\int_0^{s_i} p(\Theta_j(s))\frac{\partial \Theta_j(s)}{\partial s}e^{-rs}ds\right)\right],$$

where $\Theta_j(s_i)$ is the type of the opponent that would play s_i in equilibrium, and $P(\theta)$ and $p(\theta)$ are the cumulative distribution and the density, respectively, of the opponent's type θ. Replacing $\frac{a_i}{b_i}$ with θ_i, the first-order condition is

$$-p_j(\Theta_j(s_i))\frac{\partial \Theta_j(s_i)}{\partial s}\frac{1}{r}e^{-rs_i} - (1 - P_j(\Theta_j(s_i)))e^{-rs_i} + p_j(\Theta_j(s_i))\frac{\partial \Theta_j(s_i)}{\partial s}\frac{\theta_i}{r}e^{-rs_i} = 0.$$

Collecting terms and cancelling, we get

$$p_j(\Theta_j(s_i))\frac{\partial \Theta_j(s_i)}{\partial s}\frac{\theta_i - 1}{r} = (1 - P_j(\Theta_j(s_i))). \tag{3}$$

Exploiting symmetry and isolating $\frac{\partial s(\theta_i)}{\partial \theta}$ yields $\frac{\partial s(\theta_i)}{\partial \theta} = \frac{p(\theta_i)}{1-P(\theta_i)}\frac{\theta_i - 1}{r}$. Integrating, we obtain a family of solutions that vary only with respect to the constant of integration, k:

$$s^*(\theta_i) = \frac{1}{r}\int_{\theta_l}^{\theta_i}(\theta - 1)\frac{p(\theta)}{1 - P(\theta)}d\theta + k.$$

Given strict monotonicity, if $s(\theta_l)$ were strictly positive, then θ_l could increase her utility by unilaterally switching to $s = 0$. Therefore, $s(\theta_l) = 0$ in equilibrium and, thus, $k=0$. Hence the equilibrium strategy is unique. ■

Theorem 2

Proof. (1) For $i=1,2$ player i's expected utility from deviating and subsequently playing the standoff game is at least $\frac{b_i}{r}$ and for all $\theta_i > 1$ is strictly greater than $\frac{b_i}{r}$. Because some player i receives $\frac{b_i}{r}$ in any peaceful stationary equilibrium, some player has an incentive to deviate.

(2) The proof proceeds by analyzing four collectively exhaustive types of strategy profiles. I first derive the necessary and sufficient conditions for the existence of an equilibrium strategy profile of the first type. I then show that the necessary and sufficient conditions for the existence of an equilibrium strategy profile of each of the other types must be more demanding.

From Theorem 1 and $E[u_i(t_i)]$ (see expression (2)), the expected discounted present value of deviating is

$$\frac{1}{r}(b_i e^{-rs(\theta_i)}(1-P(\theta_i)) + a_i \int_{\theta_{\min}}^{\theta_i} p(\theta')e^{-rs(\theta')}d\theta' + a_i P(\theta_{\min})), \tag{4}$$

where $s(\theta_i) = \frac{1}{r}\int_{\theta_{\min}}^{\theta_i} (\theta'-1)\frac{p(\theta')}{1-P(\theta')}d\theta'$.

Consider first a strategy of the following form. Let n be a natural number and let $T \in \mathbb{R}_{++}$ be a length of time. Let $t \in \mathbb{R}_+$ be a point in calendar time. Then $\forall t$, choose

$$\begin{cases} X \text{ if } \exists n, \text{ s.t.} (n-1)T \leq t < nT \\ Y \text{ else.} \end{cases} \tag{5}$$

The discounted present value of following this strategy, given that the opponent does also, is lowest for player 2 at $t=(n-1)T$ for n odd, and for player 1 at $t=(n-1)T$ for n even. At such a point, the discounted present value is

$$\begin{aligned} & \int_0^T b_i e^{-rt}dt + \int_T^{2T} a_i e^{-rt}dt + \int_{2T}^{3T} b_i e^{-rt}dt + ... \\ = \; & b_i \int_0^\infty e^{-rt}dt + (a_i - b_i)(\int_T^{2T} e^{-rt}dt + \int_{3T}^{4T} e^{-rt}dt + ...) \\ = \; & \frac{1}{r}(b_i - (a_i-b_i)\sum_{k=1}^{\infty}(-1)^k e^{-rkT}) \\ = \; & \frac{1}{r}(b_i + (a_i - b_i)\frac{1}{1+e^{rT}}). \end{aligned} \tag{6}$$

Every type of player prefers using this strategy to deviating from it (given that the opponent follows it) if and only if

$$\begin{aligned}\frac{1}{r}(b_i + (a_i - b_i)\frac{1}{1+e^{rT}}) \quad &\geq \quad \frac{1}{r}b_i e^{-rs(\theta_i)}(1 - P(\theta_i)) \\ &+\frac{1}{r}(a_i \int_{\theta_{\min}}^{\theta_i} p(\theta')e^{-rs(\theta')}d\theta' + a_i P(\theta_{\min})\end{aligned}$$

for all (a_i, b_i). After some algebraic manipulations, we get the following equivalent inequality expressed in terms of θ_i:

$$1 + (\theta_i - 1)\frac{1}{1+e^{rT}} \geq e^{-rs(\theta_i)}(1 - P(\theta_i)) + \theta_i \int_{\theta_{\min}}^{\theta_i} p(\theta')e^{-rs(\theta')}d\theta' + \theta_i P(\theta_{\min}). \quad (7)$$

The LHS of (7) is increasing and linear in θ_i. Using the fact that $\frac{\partial s}{\partial \theta} = \frac{1}{r}(\theta - 1)\frac{p(\theta)}{1-P(\theta)}$, the derivative of the RHS reduces to $\int_{\theta_{\min}}^{\theta_i} p(\theta')e^{-rs(\theta')}d\theta'$, and so the RHS is increasing in θ_i and convex $\forall\theta_i \in (\theta_{\min}, \bar{\theta})$.

Since (7) holds with equality at $\theta_i = \theta_{\min} = 1$, it follows that (7) holds $\forall\theta_i < \theta'$ iff it holds for $\theta_i = \theta'$. Let θ' be $\theta > 1$ s.t. (7) holds at equality (or let $\theta' = \bar{\theta}$ if $\nexists\theta > 1$ s.t. (7) holds at equality). Suppose $\theta' < \bar{\theta}$. Then $\forall\theta_j > \theta'$, j starts a war of attrition; th epossibility of j's doing so lowers i's incentive not to, thus the true cutpoint is lower than θ'. Because not starting a war of attrition indicates that one's type is less than the cutpoint, it also makes starting war of attrition more appealing for one's opponent. Thus, if θ^t represents the cutpoint at time t, then $\theta^{t+1} < \theta^t$ for all t. Thus, lasting peaceful outcomes are possible if and only if (7) holds for $\theta_i = \bar{\theta}$.

Substituting $\theta_i = \bar{\theta}$ into (7) and recognizing that $P(\bar{\theta}) = 1$, (7) implies

$$1 + (\bar{\theta} - 1)\frac{1}{1+e^{rT}} \geq \bar{\theta} \int_{\theta_{\min}}^{\bar{\theta}} p(\theta)e^{-rs(\theta)}d\theta.$$

Multiplying by $(1 + e^{rT})$ and re-arranging terms, we obtain

$$e^{rT}(1 - \bar{\theta} \int_{\theta_{\min}}^{\bar{\theta}} p(\theta)e^{-rs(\theta)}d\theta - \bar{\theta}P(\theta_{\min})) + \bar{\theta}(1 - \int_{\theta_{\min}}^{\bar{\theta}} p(\theta)e^{-rs(\theta)}d\theta - P(\theta_{\min})) \geq 0. \quad (8)$$

From $r > 0$ and $s(\theta) \geq 0$, it follows that $e^{-rs(\theta)} \leq 1$. Hence,

$$\int_{\theta_{\min}}^{\bar{\theta}} p(\theta)e^{-rs(\theta)}d\theta < \int_{\theta_{\min}}^{\bar{\theta}} p(\theta)d\theta,$$

and so the second term in (8) is always positive.

Because $e^{rT} > 1$ with $\lim_{T\to 0} e^{rT} = 1$, it follows that $\exists T$ s.t. the strategies described constitute a PBE iff condition (1) in the statement of the theorem holds.

Consider a similar strategy profile in which $\forall n \in \mathbb{N}$

$$\begin{cases} (X, X) \text{ if } (n-1)(T_1 + T_2) \leq t < (n-1)T_2 + nT_1 \\ (Y, Y) \text{ if } (n-1)T_2 + nT_1 \leq t < n(T_1 + T_2). \end{cases} \tag{9}$$

Suppose $T_2 < T_1$. Then, holding type constant, player 2's present discounted value at $t = (n-1)(T_1+T_2)$ is less than player 1's at $t = (n-1)T_2+nT_1$. Thus, the profile is a weak PBE iff player 2 of type $\bar{\theta}$ prefers not to defect at $t = 0$. While her discounted present value of the war of attrition (expression (4)) is the same as above, her present discounted value of following this strategy is

$$\begin{aligned} &\int_0^{T_1} b_2 e^{-rt}dt + \int_{T_1}^{T_1+T_2} a_2 e^{-rt}dt + \int_{T_1+T_2}^{2T_1+T_2} b_1 e^{-rt}dt + \int_{2T_1+T_2}^{2(T_1+T_2)} a_2 e^{-rt}dt + ... \\ = &\ \frac{1}{r}[b_2 + (a_2 - b_2)\frac{1}{e^{r(T_1+T_2)} - 1}(1 - e^{rT_2})]. \end{aligned}$$

Given that $T_2 < T_1$, and hence $T_2 < \frac{1}{2}(T_1 + T_2)$, this is strictly less than (6) for $T = \frac{1}{2}(T_1 + T_2)$. Thus the necessary and sufficient condition for the existence of an equilibrium of form (9) is always strictly more demanding than the condition for the existence of an equilibrium of form (5). Since the game is symmetric and strategically equivalent $\forall t$, the results are the same for $T_1 < T_2$.

Consider a strategy profile that may be characterized by an infinite sequence of points in time, $\{t^n\}_{n=1}^{\infty}$, such that $\forall n \in \mathbb{N}$

$$\begin{cases} (X, X) \text{ if } t^{n-1} \leq t < t^n \text{ for } n \text{ odd} \\ (Y, Y) \text{ if } t^{n-1} \leq t < t^n \text{ for } n \text{ even.} \end{cases}$$

Now the minimum present discounted value of following this strategy, given that the other player does also, occurs at some $t \in \{t^n : n \text{ is even}\}$ for player 1 and at some $t \in \{t^n : n \text{ is odd}\}$ for player 2, but their present discounted values may vary across these values of t. Since the strategy profile is an equilibrium iff every type of each player prefers not to deviate at every point in time, it is an equilibrium iff the highest type of each player prefers not to deviate at her time of lowest present discounted value. But this condition is strictly more difficult to satisfy than the necessary and sufficient condition for existence for some equilibrium of form (9).

Finally, observe that since $b_1 > 0$ for all types, any strategy profile that requires the play of (X, Y) or (Y, X) for a positive measure of time has a lower present discounted value than some strategy profile in which only (X, X) and (Y, Y) are played.

(3) First observe that $\lim_{\bar{\theta} \to \infty} \frac{1+\bar{\theta}}{2\bar{\theta}} = \frac{1}{2}$ and that LHS of condition (1) is decreasing in $\bar{\theta}$. Second, observe that $e^{-rs^*(\theta)}$ is decreasing in θ (since from Theorem 1, $s^*(\theta)$ is increasing in θ), that $e^{-rs^*(\theta_{\min})} = e^0 = 1$, and that $\lim_{s^*(\theta) \to \infty} e^{-rs^*(\theta)} = 0$. Thus, $\int_{\theta_{\min}}^{\bar{\theta}} p(\theta) e^{-rs^*(\theta)} d\bar{\theta} + P(\theta_{\min}) < 1$, and $\int_{\theta_{\min}}^{\bar{\theta}} p(\theta) e^{-rs^*(\theta)} d\bar{\theta} + P(\theta_{\min})$ decreases as probability mass is redistributed from lower θ to higher θ. hence condition (1) is satisfied if $p(\theta)$ is increasing and sufficiently convex. ■

Theorem 3

Proof. To prove the theorem, I first show that its claim holds for the cheap-talk extension of the standoff game in which communication occurs only before playing the standoff game. (Accordingly, I drop the time superscript on players' messages.) I then argue that the result extends if we allow for the possibility of communication at any point in that game.

Consider first the possibility of a monotonic informative equilibrium in the preplay cheap-talk extension of the standoff game without mediation. Because the existence of such equilibrium with more than two messages implies the existence of one with only two messages, to prove the impossibility of informative monotonic equilibria, it is sufficient to show that there exists no monotonic two-message equilibrium. First consider a symmetric (partial) strategy profile in which the type space is partitioned by $\hat{\theta}$, such that $\forall \theta < \hat{\theta}$, $m = m^1$ and $\forall \theta > \hat{\theta}$, $m = m^2$. If $m_i > m_j$, then it is common knowledge that $\theta_i > \theta_j$ and, from the Common Knowledge Peace selection criterion, they play i's preferred action profile thereafter. If $m_i = m_j$, then they each update their beliefs and play the war of attrition as analyzed in the previous subsection.

This symmetric strategy profile cannot be an equilibrium because $\theta < \hat{\theta}$ prefers to defect to m^2. To see that this is so, we will first establish that, if her opponent (who is using this strategy) sends m^2, then she will choose $s = 0$ in the subsequent war of attrition. Suppose $\theta_2 > \hat{\theta}$ and $\theta_1 < \hat{\theta}$, and both play m^2. (Assume the informative equilibrium but player 1 is deviating in the cheap-talk stage). Player 1 wants to choose t_1 such that

$$-(1 - P(s^{-1}(t_1))) + p(s^{-1}(t_1)) \frac{\partial s^{-1}(t_1)}{\partial t} \frac{1}{r} (\theta_1 - 1) = 0.$$

Recall that $s(\theta|(m^2, m^2))$ is such that $s(\hat{\theta}|(m^2, m^2)) = 0$. Hence,

$$p(s^{-1}(0))\frac{\partial s^{-1}(0)}{\partial t}\frac{1}{r}(\hat{\theta} - 1) = 1 - P(s^{-1}(0)).$$

Note that the left-hand side is increasing in θ, while the right-hand side is constant. It follows that if $\hat{\theta}$ is not getting a sufficient expected benefit from staying in, then $\theta < \hat{\theta}$ is not either. Hence, $\forall \theta < \hat{\theta}$, $s(\theta|(m^2, m^2)) = 0$ and $E[u_1(0)] = \frac{b_1}{r}$.

Note that if $\theta_i < \hat{\theta}$ sends m^1 and her opponent sends m^2, then θ_i quits immediately and obtains the same payoff, $\frac{b_1}{r}$. It follows that which message yields higher expected utility depends entirely on the payoffs obtained when the opponent sends m^1, i.e., when she is also of a type $\theta < \hat{\theta}$. According to the proposed equilibrium, if i sends message m^2 and her opponent sends m^1, then her opponent quits immediately and i obtains her highest possible payoff $\frac{a_1}{r}$. On the other hand, if she were to send m^1, she would then have to play asymmetric war of attrition with a lower expected payoff. Hence, she will always want to choose m^2, and so there is no symmetric monotonic informative equilibrium.

Consider then an asymmetric monotonic strategy profile: Let $(\hat{\theta}_1, \hat{\theta}_2)$ be such that for $i = 1, 2$, $\forall \theta_i < \hat{\theta}_i$, $m_i = m^1$ and $\forall \theta_i > \hat{\theta}_i$, $m_i = m^2$. Let $\hat{\theta}_1 < \hat{\theta}_2$ wlog. Suppose $\theta_1 < \hat{\theta}_1$. Suppose $m_2 = m^2$; then $\theta_2 \geq \hat{\theta}_2 > \hat{\theta}_1$. From the Common Knowledge Peace equilibrium selection requirement, if $m_1 = m^1$ then the players choose action profile (Y, Y) henceforth. If $m_1 = m^2$ then they play the standoff subgame, in which player 1 must do at least as well in expectation as she would if the players chose action profile (Y, Y) henceforth. Suppose instead that $m_2 = m^1$; then $\theta_2 \leq \hat{\theta}_2$. Then m_1 determines 2's beliefs about θ_1 in the standoff game: if $m_1 = m^1$, 2 believes $\theta_1 \leq \hat{\theta}_1$, and if $m_1 = m^2$, 2 believes $\theta_1 \geq \hat{\theta}_1$. Because 2 chooses a lower strategy when she believes $\theta_1 \geq \hat{\theta}_1$, $m_1 = m^2$ dominates $m_1 = m^1$ for all θ_1. Thus there is no asymmetric monotonic informative equilibrium.

Consider next the possibility of a two-message non-monotonic equilibrium. If such an equilibrium exists, then, for some θ', θ'' such that $\theta' < \theta''$, all $\theta < \theta'$ and all $\theta > \theta''$ send message m^1, and all $\theta \in (\theta', \theta'')$ send message m^2. Suppose, without loss of generality, that $m_1 = m^1$ and $m_2 = m^2$. Because strategy is increasing in θ for each player, and because each opponent must be expected to quit with positive probability at every time $s < s(\theta^h)$, there must be a time

$$\hat{s} \equiv \lim_{\theta_1 \to_- \theta'} s(\theta_1 | m_1 = m^1) = \lim_{\theta_1 \to_+ \theta''} s(\theta_1 | m_1 = m^1)$$

such that $\theta_1 < \theta'$ if and only if $s(\theta_1|m^1) < \hat{s}$ and $\theta_1 > \theta''$ if and only if $s(\theta_1|m^1) > \hat{s}$. Then, if player 1 continues at time $\hat{s}$, it becomes common knowledge that $\theta_1 > \theta'' > \theta_2$ and player 2 quits. But then the type θ' strictly prefers m^1 to m^2, and θ'' does also, which yields a contradiction.

Now consider the possibility of a non-monotonic informative equilibrium with any number of messages. For any symmetric nonmonotonic messaging strategy $m(\theta)$, there exists a pair of distinct messages m^1, m^2 and distinct types θ', θ'' that satisfy the following conditions: 1) $\theta' < \theta''$; 2) $s(\theta'|m^1) = s(\theta''|m^1)$; 3) there exists a $\theta \in \{\theta|m(\theta) = m^2\}$ such that $\theta > \theta'$; 4) for every $\theta \in \{\theta|m(\theta) = m^2\}$, $\theta < \theta''$. But then the corresponding partial strategy profile is strategically equivalent to the two-message non-monotonic strategy profile, and, by the same argument, cannot be part of an equilibrium strategy profile.

Finally, to see that allowing symmetric unmediated cheap-talk at any other point in the game will not admit informative communication in equilibrium, first note that, correcting for beliefs, the game is strategically equivalent at any point in time provided that neither player has conceded. Since the proofs rely only on the continuity of $P(\theta)$ and the finiteness of its support, both of which are preserved as the game progresses, informative unmediated cheap talk is not possible at any time. ■

Theorem 4

Proof. Applying the Revelation Principle, it is sufficient to show that there is no ICDM. First consider the stationary case: Suppose the mediator recommends (X, X) independent of the players' messages. The mediator's advice communicates no information about the opponent's type, thus the expected value for $\theta_2 > 1$ of initiating a standoff is strictly greater than $\frac{b_2}{r}$, the payoff from following the mediator's advice. The case in which the mediator recommends (Y, Y) is symmetric (player 1 deviates). Suppose instead the mediator's choice of (X, X) or (Y, Y) is contingent on players' messages. If it is a deterministic rule, then each player has an incentive to misrepresent her type as that which is most likely to yield the choice of her more preferred action profile. If it is a probabilistic rule, then incentives for truthful revelation may be maintained by assigning a higher probability to (X, Y) or (Y, X), each of which has a payoff of zero, in response to higher messages. However, neither player has an incentive to comply with the mediator's recommendations in this case, since each can obtain expected utility from initiating standoff of at least $\frac{b_i}{r} > 0$. Thus truthful revelation cannot be maintained for a stationary mechanism.

Consider the non-stationary case: Suppose parameter values such that there is no peaceful equilibrium in the game without mediation (from Theorem 2). From the derivation of that condition, we see that the mediator must recommend playing the action profile preferred by the agent with higher θ more than half the time for at least some pairs of types in order to insure that players prefer compliance to initiating a standoff. But then every player i prefers to send message $\bar{\theta}$ to message θ_i. ■

Theorem 5

Proof. Existence: Let $W(a_i, b_i, v_i, v_j)$ be i's expected utility in equilibrium from the war of attrition following the burning of v_i and v_j. (Note that v_i and v_j affect W only via their effects on the players' beliefs about their opponents, beliefs that in turn affect their strategies in the war of attrition.) $\frac{\partial W}{\partial s_j} \leq 0$, and $\frac{\partial^2 W}{\partial a \partial s_j} \leq 0$ and $\frac{\partial^2 W}{\partial b \partial s_j} \geq 0$. If $\frac{\partial s_j}{\partial v_i} \geq 0$, then $v_i = 0\ \forall (a_i, b_i)$. Therefore consider $\frac{\partial s_j}{\partial v_i} < 0$. Then $\frac{\partial^2 W}{\partial a \partial s_j} \leq 0$ implies

$\frac{\partial v}{\partial a_i} \geq 0$ and $\frac{\partial^2 W}{\partial b \partial s_j} \geq 0$ implies $\frac{\partial v}{\partial b_i} \leq 0$. Given $\theta = \frac{\alpha}{\beta}$, $\frac{\partial^2}{\partial v_i \partial \theta_i}(\frac{p(\theta|v)}{1-P(\theta|v)}) < 0$. From (3), $\frac{\partial s_j(\theta_j|v_i)}{\partial v_i} < 0$.

We prove Part 2 next. To show that $v(\theta)$ s.t. $\frac{\partial v}{\partial \theta} > 0$ cannot be sustained in equilibrium, suppose that $\frac{\partial v}{\partial \theta} > 0$. Then the inverse of $v(\theta)$ exists; let this be $\Theta(v)$. Then $v_i > v_j$ if and only if $\theta_i > \theta_j$. Then

$$\begin{aligned} E[u_i(v_i; \cdot)] &= -v_i + \Pr(v_j < v_i)\frac{a_i}{r} + \Pr(v_j > v_i)\frac{b_i}{r} \\ &= -v_i + P(\Theta(v_i))\frac{a_i}{r} + (1-P)\Pr(\Theta(v_i))\frac{b_i}{r}. \end{aligned}$$

The FOC is

$$-1 + p(\Theta(v_i))\frac{\partial(\Theta(v_i))}{\partial v}(\frac{a_i - b_i}{r}) = 0.$$

Because this FOC cannot be expressed solely in terms of θ (without a_i, b_i independent of θ_i), it cannot be that v fully reveals θ in equilibrium. Thus, conflict must occur on the path of play.

Part 1: Because $\frac{\partial v}{\partial a} \geq 0$ and $\frac{\partial v}{\partial b} \leq 0$, and $\theta = \frac{a}{b}$, v communicates information about θ. In particular, for $\{(a,b) : v^*(a,b) = v\}$, there is a lowest θ' s.t. $\exists (a', b') \in \{(a,b) : v^*(a,b) = v\}$ s.t. $\frac{a'}{b'} = \theta'$ and, likewise, a highest θ'' s.t. $\exists (a'', b'') \in \{(a,b) : v^*(a,b) = v\}$ s.t. $\frac{a''}{b''} = \theta''$, and $\frac{\partial \theta'}{\partial v} > 0$ and $\frac{\partial \theta''}{\partial v} > 0$. Thus, for pairs of opponents such that $|v_1 - v_2|$ is sufficiently large, it is common knowledge which player has higher θ, and thus the players coordinate immediately on that player's more preferred action profile. ■

References

[1] Abreu, Dilip and David Pearce. 2006. "Reputational Wars of Attrition with Complex Bargaining Postures." Princeton University Mimeo.

[2] Abreu, Dilip and David Pearce. 2002. "Bargaining, Reputation, and Equilibrium Selection in Repeated Games." Princeton University Mimeo.

[3] Alesina, Alberto. 1988. "Credibility and Policy Convergence in a Two-Party System with Rational Voters." *American Economic Review* 78, 796-805.

[4] Alesina, Alberto and Alan Drazen. 1991."Why Are Stabilizations Delayed?" *American Economic Review* 81 (December): 1170-1188.

[5] Baliga, Sandeep and Thomas Sjostrom. 2004. "Arms Races and Negotiations." *Review of Economic Studies.*

[6] Baliga, Sandeep and Stephen Morris. 1998. "Cheap Talk and Coordination with Payoff Uncertainty." Cowles Foundation Discussion Paper #1403.

[7] Baliga, Sandeep and Stephen Morris. 2002. "Cheap Talk, Coordination, and Spillovers." *Journal of Economic Theory.*

[8] Banks, Jeffrey S. 1990. "Equilibrium Behavior in Crisis Bargaining Games." *American Journal of Political Science* 34 (3), 599-614.

[9] Banks, Jeffrey S. and Randall L. Calvert. 1992. "A Battle-of-the-Sexes game with Incomplete Information." *Games and Economic Behavior* 4, 347-72.

[10] Ben-Porath, Elchanan. 2002. "Cheap Talk in Games with Incomplete Information." Hebrew University Mimeo.

[11] Bueno de Mesquita, Bruce. 2000. "Popes, Kings, and Endogenous Institutions: the Concordat of Worms and the Origins of Sovereignty." *International Studies Review* 2 (2), 93-118.

[12] Bueno de Mesquita, Bruce, James Morrow, and Ethan Zorick. 1997. "Capabilities, Perception, and Escalation." *American Political Science Review* 91 (1), 15-27.

[13] Casella, Alessandra and Barry Eichengreen. 1996. "Can Foreign Aid Accelerate Stabilization?" *The Economic Journal* 106, 605-619.

[14] Dixit, Avinash K. 2004. *Lawlessness and Economics: Alternative Modes of Governance*. Princeton: Princeton University Press.

[15] Dixit, Avinash, Gene M. Grossman, and Faruk Gul. 2000. "The Dynamics of Political Compromise." *Journal of Political Economy* 108 (3), 531-568.

[16] Drazen, Alan and Vincent Grilli. 1993. "The Benefits of Crises for Economic Reforms." *American Economic Review* 82, 598-607.

[17] Fudenberg, Drew and Jean Tirole. 1991. *Game Theory.* Cambridge: MIT Press.

[18] Fudenberg, Drew and Jean Tirole. 1986. "A Theory of Exit in Duopoly." *Econometrica* 54 (4), 943-60.

[19] Hafer, Catherine. 2006. "On the Origins of Property Rights: Conflict and Production in the State of Nature." *Review of Economic Studies* 73 (1), 119-144.

[20] Hafer, Catherine. 2007. "Transferable Utility in the Game of Standoff." New York University Mimeo.

[21] Hsieh, Chang-Tai. 2000. "Bargaining Over Reform." *European Economic Review* 44 (9), 1659-76.

[22] Jackson, Mattew O. and Hugo F. Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." *Econometrica* 75 (1), 241-258.

[23] Myerson, Roger. 1985. "Bayesian Equilibrium and Incentive Compatibility." In L. Hurwicz, D. Schmeidler, and H. Sonnenschein, eds., *Social Goals and Social Organization*. Cambridge: Cambridge University Press, 229-59.

[24] Nalebuff, Barry. 1987. "Credible Pretrial Negotiation." *The Rand Journal of Economics* 18 (2), 198-210.

[25] Powell, Robert L. 2004. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48 (2), 344-61.

[26] Spolaore, Enrico. 2004. "Adjustments in Different Government Systems." *Economics and Politics* 16 (2), 117-46.