Conflict Resolution[*]

Helmut Bester[†]   Karl Wärneryd[‡]

August 1999

Abstract

We consider mechanisms for resolving conflicts between two agents who are uncertain about each other's fighting potential. Applications include international conflict, litigation, and elections. Even though only a peaceful agreement avoids a loss of resources, if this loss is small enough, then any mechanism must assign a positive probability of conflict. We show how the likelihood of conflict outbreak depends on the distribution of power between the agents and their information about each other.

Keywords: conflict, asymmetric information

Journal of Economic Literature Classification Numbers: C78, D72, D74, D82, H21, H23.

[†]Department of Economics, Free University of Berlin, Boltzmannstr 20, D—14195 Berlin, Germany.

[‡]Department of Economics, Stockholm School of Economics, Box 6501, S—113 83 Stockholm, Sweden.

Qui desiderat pacem, præparet bellum.[1]

1 Introduction

Economic life is replete with situations where conflicts of interest are resolved not through voluntary agreement, but by means such as theft, warfare, and litigation. In this paper, we develop a theory of when such conflicts may be unavoidable.

Our theory applies to contexts such as the following.

1. Two countries negotiate about the division of the fishing waters located between them. In case negotiations break down, they go to war.

2. A buyer and a seller of a good try to renegotiate the terms of a trading contract. In case they fail to reach an agreement, they go to court.

3. Two political parties try to form a coalition. In case their negotiations are unsuccessful, they enter the elections on uncoordinated policy platforms.

All these examples are bargaining problems where the outside option can be thought of as a probabilistic conflict. The outcome of war, the default option in case 1, is partly a result of the relative military strength of the feuding parties, partly of chance. In case 2, we note (with, e g, Williamson [21] and Grossman and Hart [8]) that contracts may be incomplete and courts fallible. Any already existing contract between the parties must be interpreted by the court in the event of conflict. Since no contract can unambiguously specify what should happen in every possible contingency, and since the court is not omniscient and has no other source of information about the original intentions of the parties, the court must at least partially rely on the cases made by the interested parties. Hence we can think of the option of going

---

[1]Vegetius, *De Re Militari* III, c 375.

to court as stochastic, with win probabilities for the respective parties determined partly by the merits of their cases, partly by their own preparations in presenting their cases. In case 3, if the political parties are uncertain about the preferences of the electorate and each other's policies, then it is natural to view the outcome of the election as a probabilistic contest also. (See, e g, Coughlin [5] for a discussion of such models.)

Put differently, these examples are of bargaining in the absence of well-defined property rights. Economic theory normally deals with situations where property rights are implicitly taken to be well defined and outright conflict never occurs. Yet conflict is pervasive in the real world, and it may well be said that a common procedure for acquiring resources is through coercion rather than market transactions. This is especially true when one recognizes that coercion plays a role in economic affairs that is usually more subtle than actual theft. Since, as noted above, contracts cannot specify what should happen in all possible contingencies, courts often eventually have to rule on the outcome. Thus property rights themselves ultimately rest on coercive power relationships.

In this paper, we explore the limits of peaceful agreement in a setting where agents have the potential for outright conflict, where the latter may take any form from actual armed battle to court proceedings. Although Coase [4] famously listed the lack of well-defined property rights as a potential impediment to reaching efficient agreements, mutually beneficial trade may still be possible in such circumstances. As long as the parties' relative strengths in an outright conflict are common knowledge among them, they would, if risk-neutral, accept an agreement that gave each of them at least their expected value under conflict. In case the agents are risk averse, the scope for ex ante agreement is even greater.

It is well known, however, that asymmetric information may cause the Coase theorem to fail. (See, e g, Farrell [6].) An example of this is the inefficiency result of Myerson and Satterthwaite [16], in which the respective

valuations of a seller and a buyer of a good are private information. Their argument depends on there being a positive probability of trade being inefficient ex post. In this paper, we show that asymmetric information about relative strengths in an outside option conflict may cause agreement to be impossible *even if it is always efficient*.

The intuitive reason for our result is that if the loss of surplus resulting from outright conflict is too low, then an agent of high strength may think himself sufficiently likely to win a conflict to make this his preferred option. This contrast sharply with the world of the Coase theorem, where an efficient agreement will always be reached if it exists and property rights are well defined. In this paper, we present a theoretical framework that allows us to specify *precisely* in what sense and how well defined property rights have to be in order for the Coase theorem to hold.

Since our approach is very general, it can address a broad variety of issues both fundamental and more special. The most fundamental question concerns the stability of society as such. At least since Thomas Hobbes, social theorists have been concerned with the possibility of social contracts that lift society out of endless, multilaterally immiserating conflict. Buchanan [3] noted that any such contract must respect the "natural distribution" of power in the state of nature, i e, it must give each agent at least what he could guarantee himself under anarchy. The present paper can be seen as a contribution toward the formal analysis of this participation constraint of the social contract.

A recent literature (exemplified by Skaperdas [18], Hirshleifer [9], and Grossman and Kim [7]) studies the determination of this equilibrium "natural distribution" of armed conflict in anarchy or the Hobbesian state of nature. These contributions are concerned with determining equilibrium investments in arms. In contrast, our analysis starts at a point where such investments have already been made, but the parties are perfectly informed only about their own arsenals. That is, we study the potential for peace-

ful settlements in situations where fighting potential is no longer a decision variable. This allows us to focus specifically on the probability of the actual outbreak of conflict.

According to one traditional line of thinking in international relations theory, war outbreak is less likely if the parties are of roughly equal strength in expectation (see, e g, Morgenthau [14]). Another school holds that it is less likely if one party is much stronger than the other (see, e g, Blainey [1]). We show in this paper that the likelihood of war may depend on the entire distribution of strength, not just its expectation. As an indication of the importance of asymmetric information in this context, consider the fact that the military capability of the Soviet Union during the Cold War has since been revealed to have been considerably overestimated by NATO.

The paper is organized as follows. Section 2 introduces the framework of probabilistic conflict and mechanisms. In Section 3 we study mechanisms that assign probability zero to conflict occurring. We give precise conditions for when no such incentive compatible and individually rational mechanism exists, and show how the scope for peaceful agreement depends on the nature of the information available to the players. We also consider the total ex ante probability of conflict under any incentive compatible and individually rational mechanism. We provide a lower bound on the limit of the probability of conflict as the surplus loss from conflict goes to zero. We give a complete characterization of mechanisms that are efficient in the sense of minimizing the ex ante probability of conflict for the symmetric two-type case in Section 4. In Section 5 we show by means of an example that efficient mechanisms in general need not be monotonic in the sense that if both agents' types are higher, then the probability of conflict is higher. Section 6 discusses related literature. Finally, Section 7 concludes.

2 The Model

We analyze a situation with two risk neutral agents. For agent $a \in \{1, 2\}$ the set of possible *types* is $T_a = [\underline{t}_a, \overline{t}_a] \subset \mathbb{R}$ with $\underline{t}_a < \overline{t}_a$. We think of an agent's type as his aggressive potential or strength in a conflict. In case the relevant conflict option is warfare, the strength of an agent may be a previously made investment in arms. In case the conflict option is litigation, it may represent the intrinsic strength of the agent's case, or hours of lawyer time, or both.

We assume there is asymmetric information in the sense that an agent always knows his own type, but not the type of his opponent. The agents' types are independently chosen. The distribution of agent a's type is $F_a(t_a)$. We assume $0 < F_a(t_a) < 1$ for all $\underline{t}_a < t_a < \overline{t}_a$. This guarantees that $\underline{t}_a$ and $\overline{t}_a$ are contained in the support of F_a. It should also be noted that we allow for discrete type distributions.

The agents can either fight over some resource or reach a peaceful agreement. If they reach an agreement they share a joint surplus of size 1. In the event of conflict, the joint surplus is $0 < \theta < 1$. That is, we assume that some of the cake that is up for grabs is destroyed if an outright conflict occurs. This means that there is always something for the agents to bargain about, and that only a peaceful settlement is ex post efficient.

For $(t_1, t_2) \in T := T_1 \times T_2$, agent 1's probability of winning the fight is $p(t_1, t_2) \in [0, 1]$. With probability $1 - p(t_1, t_2)$ agent 2 wins the fight. We assume that the conflict technology p is continuous and that for all $(t_1, t_2) \in T$ we have that

$$p(t_1, t_2) < p(t_1', t_2) \text{ if } t_1' > t_1$$

and

$$p(t_1, t_2) > p(t_1, t_2') \text{ if } t_2' > t_2.$$

One can think of p as a measure of the degree to which property rights are well defined, given a pair of types. For example, if we have $p = 1$, then we may say that property rights in the cake are perfectly defined, with agent

1 as the owner. Any situation where p is not equal to 1 or 0 represents a setting where property rights are less than perfectly defined.

If the agents knew each other's types with certainty, a voluntary peaceful settlement would always be possible. Each agent would accept a division of the cake that gave him at least as much as his expected value under outright conflict. What potentially complicates matters in the following is the presence of informational asymmetries.

A *mechanism* determines

1. the division of the surplus under a peaceful settlement, and

2. the probability of conflict.

We allow the mechanism to specify a probability of conflict since we shall show that there may be circumstances such that any incentive compatible and individually rational mechanism must assign a positive probability of conflict occurring. Let $0 \leq \beta \leq 1$ denote agent 1's share of the surplus under a peaceful solution. Agent 2 then gets the share $1 - \beta$. Let $0 \leq \pi \leq 1$ denote the probability of conflict.

By the *revelation principle* (Myerson [15]), we can restrict attention to direct mechanisms that are (Bayesian) incentive compatible. A direct mechanism is a pair of functions $(\beta, \pi): T \to [0, 1] \times [0, 1]$. For each realization (t_1, t_2), with probability $\pi(t_1, t_2)$ conflict occurs. In this event agent 1's expected payoff is $p(t_1, t_2)\theta$ and agent 2's expected payoff is $[1 - p(t_1, t_2)]\theta$. With probability $1 - \pi(t_1, t_2)$ there is no fight; agent 1 then gets $\beta(t_1, t_2)$ and agent 2 gets $1 - \beta(t_1, t_2)$.

Define
$$U_1(t_1'|t_1) := \int_{T_2} [[1 - \pi(t_1', t_2)]\beta(t_1', t_2) + \pi(t_1', t_2)p(t_1, t_2)\theta] \, \mathrm{d}F_2$$
and
$$U_2(t_2'|t_2) := \int_{T_1} [[1 - \pi(t_1, t_2')][1 - \beta(t_1, t_2')] + \pi(t_1, t_2')[1 - p(t_1, t_2)]\theta] \, \mathrm{d}F_1.$$

$U_a(t'_a|t_a)$ is agent a's expected payoff from announcing type t'_a when his actual type is t_a and the opponent always announces his type truthfully.

The mechanism (β, π) is *incentive compatible* if for $a \in \{1, 2\}$ and for all $t_a \in T_a$ we have that

$$U_a(t_a|t_a) \geq U_a(t'_a|t_a) \text{ for all } t'_a \in T_a.$$

The mechanism (β, π) is *individually rational* if for all $(t_1, t_2) \in T$ we have that

$$U_1(t_1|t_1) \geq \bar{U}_1(t_1) := \int_{T_2} p(t_1, t_2) \theta \mathrm{d}F_2$$

and

$$U_2(t_2|t_2) \geq \bar{U}_2(t_2) := \int_{T_1} [1 - p(t_1, t_2)] \theta \mathrm{d}F_1.$$

In what follows we consider exclusively incentive compatible and individually rational mechanisms.

We may interpret a mechanism as a binding contract offered to the agents by a third-party mechanism designer, mediator, or Hobbesian Sovereign. If one or both agents turn down the contract, outright conflict is the default option. Viewed from this perspective, the set of mechanisms describes the outcomes that can be implemented by a mediator. The mechanism design approach, however, also covers the potential solutions of unmediated bargaining between the two players. Regardless of how complicated such a bargaining game might be, possibly involving many stages of actions for each party, any solution concept for the game in the end boils down to a mapping from the sets of possible types of each player to the set of outcomes—in this case divisions of the cake and probabilities of outright conflict. This mapping has to be incentive compatible since each player type can imitate the bargaining strategy of any other type. Moreover, the solution of the game has to be individually rational as each player has the option of choosing outright conflict. Thus any conceivable outcome of bilateral bargaining is contained in the set of mechanisms. If a mediator is unable to assign zero probability to the event of conflict, then this event occurs with positive probability also when the two players bargain about a division of the cake.

3   Peaceful Settlement

We call a mechanism *peaceful* if $\pi(t_1, t_2) = 0$ for all $(t_1, t_2) \in T$. Clearly, a peaceful mechanism is ex post efficient, since it always results in a division of the entire potential cake among the agents. We now show that for certain values of θ, there is no peaceful mechanism. We then investigate how the feasibility of peaceful conflict resolution depends on the distribution of power between the agents and their information about each other.

Lemma 1, which is stated and proved in the Appendix, shows that $\pi(t_1, t_2) = 0$ implies that for each a there is a V_a such that $U_a(t_a|t_a) = V_a$ for all $t_a \in T_a$. Since $\pi(t_1, t_2) = 0$ for all $(t_1, t_2) \in T$, by the definition of $U_a(t'_a|t_a)$ this implies that

$$V_1 = \int_{T_1} [\int_{T_2} \beta(t_1, t_2) \mathrm{d}F_2] \mathrm{d}F_1$$

and

$$V_2 = \int_{T_2} [\int_{T_1} [1 - \beta(t_1, t_2)] \mathrm{d}F_1] \mathrm{d}F_2.$$

This in turn implies $V_1 + V_2 = 1$. By the monotonicity of p, individual rationality is satisfied if and only if

$$V_1 \geq \int_{T_2} p(\bar{t}_1, t_2) \theta \mathrm{d}F_2 \text{ and } V_2 \geq \int_{T_1} [1 - p(t_1, \bar{t}_2)] \theta \mathrm{d}F_1, \tag{1}$$

i e, if it holds for the highest type of each agent, respectively.

Since $V_2 = 1 - V_1$, these inequalities have a solution if and only if

$$\theta \leq \hat{\theta} := \left[\int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}F_2 + \int_{T_1} [1 - p(t_1, \bar{t}_2)] \mathrm{d}F_1 \right]^{-1}.$$

Hence there cannot be a peaceful mechanism if $\theta > \hat{\theta}$. If $\theta \leq \hat{\theta}$, (1) has a solution with $V_1 + V_2 = 1$. It is easily verified that setting $\beta(t_1, t_2) = V_1$ and $\pi(t_1, t_2) = 0$ for all $(t_1, t_2) \in T$ satisfies incentive compatibility and individual rationality. This proves that there is a peaceful mechanism if $\theta \leq \hat{\theta}$.

Furthermore, by the monotonicity of p we have that

$$2 > \hat{\theta}^{-1} = \int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}F_2 + \int_{T_1} [1 - p(t_1, \bar{t}_2)] \mathrm{d}F_1 >$$

$$p(\bar{t}_1, \bar{t}_2) + [1 - p(\bar{t}_1, \bar{t}_2)] = 1.$$

Therefore we must have $1/2 < \hat{\theta} < 1$.

Hence we have proved the following statement.

Proposition 1 *A peaceful mechanism exists if and only if we have that*

$$\theta \leq \hat{\theta} = \left[\int_{T_2} p(\bar{t}_1, t_2) dF_2 + \int_{T_1} [1 - p(t_1, \bar{t}_2)] dF_1 \right]^{-1}.$$

Furthermore, we have $\hat{\theta} \in (1/2, 1)$.

Thus if conflict is sufficiently nondestructive, where what is sufficient is determined by the conflict technology p and the nature of the uncertainty of the situation, a peaceful settlement is impossible. The intuition for this is of course that sufficiently nondestructive conflicts make outright conflict less costly than a compromise solution for at least one agent. In particular, if both highest types are drawn, both agents may think themselves very likely to win a conflict, and would therefore prefer this option.

On the other hand, we note that if we have $\theta < 1/2$, that is, if more than half of the cake would be lost under conflict, then equal division is always incentive compatible and individually rational.

It is natural to ask how the agents' uncertainty about each other's relative strength affects the potential for peaceful agreement. Note that problems arise because higher types may prefer fighting. It is perhaps somewhat counterintuitive, then, that things are actually worse the more likely agents are to be of *low* type. To show this formally, we recall that a distribution $\hat{G}$ dominates another distribution G in the sense of *first order stochastic dominance* (FOSD) if we have that

$$\hat{G}(x) \leq G(x) \text{ for all } x.$$

Let $\hat{F}_2$ dominate F_2 in the sense of FOSD. Intuitively, this means agent 2 is more likely to be strong under $\hat{F}_2$ than under F_2. Since $p(\bar{t}_1, t_2)$ is strictly decreasing in t_2, this implies

$$\int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}\hat{F}_2 < \int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}F_2.$$

Hence $\hat{\theta}$ is increased when F_2 is replaced by $\hat{F}_2$. An analogous argument for F_1 completes the proof of the following proposition.

Proposition 2 *An increase in F_1 and/or F_2 according to the first order stochastic dominance criterion raises $\hat{\theta}$.*

Going back a step to a here unmodelled first stage where an agent chooses his strength—for example, through an investment in arms—this result echoes the contention of Vegetius that "he who desires peace should prepare for war." The stronger is one or both agents in expectation, the greater is the scope for peaceful agreement. Conversely, the prospects for peace are worst when both players are likely to be of low type.

Next, we consider the effect of an increase in uncertainty in the sense of the riskiness of a type distribution. A distribution $\hat{G}$ is dominated by a distribution G in the sense of *second order stochastic dominance* if $\hat{G}$ and G have the same mean and for every nondecreasing concave function h we have that

$$\int h(x) \mathrm{d}\hat{G} \leq \int h(x) \mathrm{d}G \text{ for all } x.$$

Equivalently, $\hat{G}$ is *riskier* than, or a mean-preserving spread of, G.

Let $\hat{F}_2$ be riskier than F_2. Suppose $p(\bar{t}_1, t_2)$ is strictly convex and strictly decreasing in t_2. Strict convexity in t_2 means that there are diminishing returns to strength for agent 2, a natural assumption to make. Then we have that

$$\int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}\hat{F}_2 > \int_{T_2} p(\bar{t}_1, t_2) \mathrm{d}F_2.$$

Hence $\hat{\theta}$ is reduced when F_2 is replaced by $\hat{F}_2$. An analogous argument can be made for F_1. We have therefore proved the following.

11

Proposition 3 *Let $p(t_1, t_2)$ be strictly concave in t_1 and strictly convex in t_2. Then an increase in the riskiness of F_1 and/or F_2 lowers $\hat{\theta}$.*

An increase in the riskiness of the type distribution of one party may be interpreted as a loss of information for the other party. Hence the above result says that receiving a better signal about your opponent's strength enhances the potential for peaceful settlement.

The following result, which is proved in the Appendix, provides a full characterization of the set of peaceful mechanisms.

Proposition 4 *Let $\theta \leq \hat{\theta}$. Then there is an interval $I(\theta)$ such that (β, π) with $\beta(t_1, t_2) = \beta^*$ and $\pi(t_1, t_2) = 0$ for all $(t_1, t_2) \in T$ is a (peaceful) mechanism if and only if $\beta^* \in I(\theta)$. Also, any peaceful mechanism (β, π) is payoff equivalent to some peaceful mechanism (β', π') with $\beta'(t_1, t_2) = \beta^* \in I(\theta)$ for all $(t_1, t_2) \in T$. Moreover, $I(\theta)$ is strictly decreasing in θ, $I(0) = [0, 1]$, and $I(\hat{\theta})$ is a singleton. If $\theta \leq 1/2$, then $1/2 \in I(\theta)$.*

We note, in particular, that if a peaceful mechanism exists, it can always be implemented with a constant sharing rule, i e, one that is independent of the agents' types.

We next consider the probability of conflict in general. Define

$$\tilde{\pi}(\beta, \pi) := \int_{T_1} \int_{T_2} \pi(t_1, t_2) \mathrm{d}F_1 \mathrm{d}F_2,$$

the ex ante probability of conflict under the mechanism (β, π). The following claim is proved in the Appendix.

Proposition 5 *Consider a sequence θ_n with $\theta_n < 1$ and $\theta_n \to 1$. For each θ_n, let (β_n, π_n) be some associated mechanism. Then $\lim_{n \to \infty} \tilde{\pi}(\beta_n, \pi_n) \geq 1 - F_1(\underline{t}_1) F_2(\underline{t}_2)$.*

This says that for any mechanism, as θ approaches 1, i e, as the cost of conflict becomes smaller and smaller, the probability of conflict approaches a

12

number that cannot be less than one minus the ex ante probability of drawing
the two lowest types. In particular, if the distributions are continuous at
the respective lowest types, then the ex ante probability of conflict must
approach 1.

4   The Symmetric Two-Type Case

The ex ante probability of conflict is a natural measure of the ex ante ef-
ficiency of a mechanism. In this section, we characterize completely the
mechanisms that minimize the ex ante probability of conflict in a symmetric
two-type class of problems.

We call a conflict *symmetric* if we have $T_1 = T_2$, $F_1 = F_2$ and $p(t_1, t_2) = 1 - p(t_2, t_1)$ for all $(t_1, t_2) \in T$. Obviously, $p(t, t) = 1/2$ for all t in a symmetric
conflict.

We consider a symmetric example with $T_1 = T_2 = \{\underline{t}, \overline{t}\}$, where $\underline{t} < \overline{t}$. The
probability of type $\underline{t}$ is $\lambda := F_1(\underline{t}) = F_2(\underline{t}) \in (0, 1)$. By symmetry we have

$$p(\underline{t}, \underline{t}) = p(\overline{t}, \overline{t}) = 0.5.$$

By the monotonicity of p we have that

$$\bar{p} := p(\overline{t}, \underline{t}) = 1 - p(\underline{t}, \overline{t}) > 0.5.$$

In this example, we have that

$$\hat{\theta} = \frac{1}{\lambda(2\bar{p} - 1) + 1}.$$

We note that the threshold $\hat{\theta}$ is decreasing in λ, i e, the scope for peaceful
agreement worsens the more likely the agents are to be of low type.

For symmetric conflicts it is natural to look at symmetric mechanisms.
The mechanism (β, π) is *symmetric* if $\beta(t_1, t_2) = 1 - \beta(t_2, t_1)$ and $\pi(t_1, t_2) =$

$\pi(t_2, t_1)$ for all $(t_1, t_2) \in T$. We now derive the symmetric mechanism that minimizes the expected probability of conflict $\tilde{\pi}$. Note that by symmetry we have that

$$\beta(\underline{t}, \underline{t}) = \beta(\bar{t}, \bar{t}) = 0.5,$$

$$\beta(\bar{t}, \underline{t}) = 1 - \beta(\underline{t}, \bar{t}),$$

and

$$\pi(\underline{t}, \bar{t}) = 1 - \pi(\bar{t}, \underline{t}).$$

Therefore, it remains to determine the variables $\pi(\underline{t}, \underline{t})$, $\pi(\underline{t}, \bar{t})$, $\pi(\bar{t}, \bar{t})$ and $\beta(\underline{t}, \bar{t})$.

Minimizing $\tilde{\pi}$ subject to the incentive constraints and the individual rationality constraints implies that the optimal mechanism (β^*, π^*) always satisfies

$$\pi^*(\underline{t}, \underline{t}) = 0.$$

That is, the lowest types never fight each other. Of course, as long as $\theta \leq \hat{\theta}$, also $\pi^*(\underline{t}, \bar{t}) = \pi^*(\bar{t}, \bar{t}) = 0$. For $\theta \geq \hat{\theta}$ we obtain the solution

$$\pi^*(\underline{t}, \bar{t}) = \max\left[0, \frac{\lambda(2\bar{p}\theta - 2\theta + 1) + 2(\theta - 1)}{\lambda(2\bar{p}\theta - 3\theta + 2) + 2(\theta - 1)}\right], \quad \text{and}$$

$$\pi^*(\bar{t}, \bar{t}) = \min\left[\frac{\lambda\theta(2\bar{p} - 1) + \theta - 1}{(1 - \lambda)(\lambda(2\bar{p}\theta - 2\theta + 1) + \theta - 1)}, 1\right].$$

Figure 1 illustrates the situation. There is a critical value θ^* such that the probability of conflict between type $\underline{t}$ and type $\bar{t}$ is zero for $\theta \leq \theta^*$ and positive for $\theta > \theta^*$. Also, for $\theta > \theta^*$ the highest types fight each other with probability one. The critical value θ^* is defined as

$$\theta^* := \frac{2 - \lambda}{2(\lambda(\bar{p} - 1) + 1)}.$$

Hence we have $\hat{\theta} < \theta^* < 1$. We note that θ^*, the threshold beyond which everybody fights a high type opponent with positive probability, is decreasing in λ, the prior probability of the low type. Overall, an increase in λ
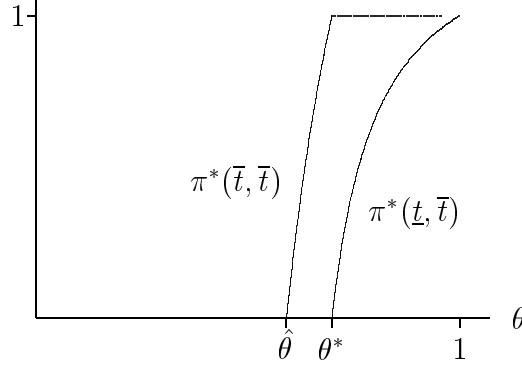
Figure 1: *The conflict probabilities $\pi^*(\bar{t}, \bar{t})$ and $\pi^*(\underline{t}, \bar{t})$*

unambiguously increases the probability of conflict for the high types.

The optimal mechanism is monotone in the sense that $\pi^*(\underline{t}, \underline{t}) \leq \pi^*(\underline{t}, \bar{t}) \leq \pi^*(\bar{t}, \bar{t})$. Moreover, $\pi^*(\underline{t}, \bar{t})$ is strictly increasing in θ when $\theta \in [\theta^*, 1]$. In accordance with Proposition 5, it is the case that $\pi^*(\underline{t}, \bar{t}) \to 1$ as $\theta \to 1$. Similarly, $\pi^*(\bar{t}, \bar{t})$ is strictly increasing in θ for $\theta \in [\hat{\theta}, \theta^*]$ and for $\theta \geq \theta^*$ we have $\pi^*(\bar{t}, \bar{t}) = 1$.

For $\theta \in [\hat{\theta}, \theta^*]$, the solution for $\beta^*(\underline{t}, \bar{t})$ is

$$\beta^*(\underline{t}, \bar{t}) = \frac{\lambda(4\bar{p}^2\theta^2 - 6\bar{p}\theta^2 + 2\theta^2 + \theta - 1) + 2\theta(1-\theta)(1-\bar{p})}{2(1 - \theta - \lambda(2\bar{p}\theta - 2\theta + 1))}.$$

For $\theta \in [\theta^*, 1]$, we have that

$$\beta^*(\underline{t}, \bar{t}) = 1 - \bar{p}\theta.$$

Thus for $\theta = \hat{\theta}$ one has $\beta^*(\underline{t}, \bar{t}) = 0.5$. Moreover, $\beta^*(\underline{t}, \bar{t})$ is continuously decreasing in θ over the range $[\hat{\theta}, 1]$. When confronted with a strong opponent, the weak agent's share of the cake under a peaceful agreement is the smaller the less resources are lost in the event of conflict.

5 An Example

In Section 4, we noted that the probability of conflict of the efficient mechanism in the symmetric two-type case is monotone in the sense that we have $\pi(t_1, t_2) \leq \pi(t'_1, t'_2)$ if $t_1 < t'_1$ or $t_2 < t'_2$ or both. It is perhaps natural to expect that this holds in general for efficient mechanisms. In this section, we show by means of a numerical example that this is not the case. This nonmonotonicity distinguishes our model from most other mechanism design problems, including that of Myerson and Satterthwaite [16]. It also means that we cannot use the standard techniques developed by Myerson and Satterthwaite to generate a more general characterization of efficient mechanisms in this setting.

The following tables present numerical computations for the following example: There are four types of each agent a; $T_a = \{t_{a1}, t_{a2}, t_{a3}, t_{a4}\} = \{1, 2, 3, 4\}$. Each type has the same probability. Agent 1's probability of winning is

$$p(t_{1i}, t_{2j}) := \frac{t_{1i} - 0.9}{t_{1i} + t_{2j} - 1.8}.$$

With this specification, we have $\hat{\theta} = 0.713521$.

The mechanism is chosen to minimize the overall expected probability of conflict.[2] The entries in the tables indicate the probability of conflict $\pi(t_{1i}, t_{2j})$ between type t_{1i} of agent 1 and type t_{2j} of agent 2 for different values of the parameter θ. The parameter $\tilde{\pi}$ represents the overall expected probability of conflict.

---

[2]This mechanism design problem can be formulated as a linear programming problem by defining $z_1(t_1, t_2) := [1 - \pi(t_1, t_2)]\beta(t_1, t_2)$, $z_2(t_1, t_2) := [1 - \pi(t_1, t_2)][1 - \beta(t_1, t_2)]$ and imposing the constraint $z_1(t_1, t_2) + z_2(t_1, t_2) = 1 - \pi(t_1, t_2)$.

	t_{21}	t_{22}	t_{23}	t_{24}
t_{11}	0	0	0	0
t_{12}	0	0.431643	0	0
t_{13}	0	0	0.60293	0
t_{14}	0	0	0	0.735539

$$\theta = 0.73 \Rightarrow \tilde{\pi} = 0.110632$$

	t_{21}	t_{22}	t_{23}	t_{24}
t_{11}	0	0	0	0
t_{12}	0	0.833118	0	0
t_{13}	0	0	0.884755	0.360773
t_{14}	0	0	0.360773	1.0

$$\theta = 0.75 \Rightarrow \tilde{\pi} = 0.21496$$

	t_{21}	t_{22}	t_{23}	t_{24}
t_{11}	0	0	0	0
t_{12}	0	1.0	0.59368	0.34842
t_{13}	0	0.59368	0.75473	1.0
t_{14}	0	0.34842	1.0	1.0

$$\theta = 0.8 \Rightarrow \tilde{\pi} = 0.41493$$

	t_{21}	t_{22}	t_{23}	t_{24}
t_{11}	0	0	0	0
t_{12}	0	1.0	0.92135	0.873874
t_{13}	0	0.92135	0.952524	1.0
t_{14}	0	0.873874	1.0	1.0

$$\theta = 0.85 \Rightarrow \tilde{\pi} = 0.533936$$

17

	t_{21}	t_{22}	t_{23}	t_{24}
t_{11}	0	0.72457	0.72457	0.72457
t_{12}	0.72457	1.0	1.0	1.0
t_{13}	0.72457	1.0	1.0	1.0
t_{14}	0.72457	1.0	1.0	1.0

$$\theta = 0.9 \Rightarrow \tilde{\pi} = 0.834214$$

We note the following about this example.

- The example exhibits certain monotone characteristics. As expected, the ex ante probability of conflict $\tilde{\pi}$ increases as θ increases. In particular, starting at $\theta = .75$, we can identify threshold types such that if both players are above their thresholds, they fight with probability 1. These thresholds get lower as θ increases. Furthermore, higher types are more likely to fight in expectation.

- We observe some striking nonmonotonicities. Consider the case when we have $\theta = .8$, and keep the type of agent 2 fixed at t_{22}. As the type of agent 1 increases, the probability of conflict rises from zero to 1, then drops off again. Nor is it the case that the probability of conflict is monotone when increasing both agents' types at the same type. Along the diagonal, the probability of conflict rises from zero to 1, then drops to .75473, then rises to 1 again.

6   Related Literature

Several papers on mechanism design under asymmetric information state inefficiency propositions closely related to our central impossibility result. The most well-known of these is perhaps that of Myerson and Satterthwaite [16] (henceforth, MS). MS study the scope for trade between a potential seller and a potential buyer of a good, under asymmetric information concerning the valuations of the respective parties. They show that under certain circumstances, specifically when there is a positive probability that the buyer's

valuation is lower than the seller's, there is no ex post efficient incentive compatible and individually rational mechanism, i e, one that prescribes that trade take place whenever there is a surplus. While our framework can also be thought of as a model of trading contracts, the uncertainty in MS concerns the size of the cake, not the probabilities of prevailing in court should the parties fail to reach an agreement. Apart from the difference in context, we wish to stress that in our model agreement may be impossible even though there is *always* a surplus to be shared. MS, in contrast, require that situations exist such that trade would be inefficient ex post.

Other differences between MS and the present model include the following. Since our outside option is a probabilistic conflict, where an agent's probability of winning is a function of both players' types, we have a model with type-dependent reservation utilities. Unlike in most mechanism design problems, including MS, the other agent's type has a direct impact on an agent's outside option payoff. Finally, the MS model requires the assumption of continuous support in the type distributions for the impossibility result, whereas our model does not.

Mailath and Postlewaite [13] discuss binary choice public goods problems with many agents. They exhibit an example where providing the good would be efficient, but the probability of this occurring goes to zero as the number of participants increases. Klibanoff and Morduch [10] study contracting in the presence of externalities, and show that efficient agreements may be impossible when the external effect is too insignificant. Ledyard and Palfrey [11,12] discuss the trade-off of interim for ex ante efficiency in a public goods framework.

The most closely related work that we are aware of is that of Spier [20], who studies the impact of different fee-shifting rules on pretrial negotiations. Spier derives a mechanism that maximizes the probability of settlement and shares some properties with Rule 68 of the Federal Rules of Civil Procedure of the United States. She proves in passing that if the exogenously given

19

costs of going to court are small enough, then there is no incentive compatible and individually rational mechanism for pretrial bargaining that assigns probability zero to going to court. Although this result is clearly an instance of the same phenomenon that we discuss in this paper, it uses a monotone likelihood ratio property of the distribution of the court's awards as a function of the agents' types, something that we do not need here. Hence our framework is more general.

Skaperdas and Syropoulos [19] study contests for a resource that, together with another good of which the agents already control initial endowments, can be used for producing a consumption good. In their model, investments in arms to secure shares of the contested resource are endogenous. They consider the efficiency effects of allowing bargaining and exchange of the contested resource in the shadow of potential conflict. In contrast with the present paper, which may be viewed as a model of a situation where investments in arms have already been made, Skaperdas and Syropoulos focus on a hold-up problem in the form of an inefficient diversion of resources from production into arming. We concentrate attention instead on the effects of asymmetric information on the potential for peaceful settlement.

Alternating offers bargaining with a conflictual outside option and asymmetric information is studied by Powell [17]. Powell assumes that the respective win probabilities in the outside option are known with certainty. The informational asymmetry concerns an agent's individual fixed cost of engaging in conflict. Hence the outside option utilities are not interdependent, as they are in our framework.

A similar setup is found in the comprehensive treatment of arms investment, negotiation, and potential war of Brito and Intriligator [2]. The authors endogenize arms investments, which determine win probabilities in the event of outright conflict (and hence correspond to our "strengths"). Similarly to Powell [17], arms investments are assumed to be observable once they have been made. The authors discuss settlement under one-sided imperfect

information about individual fixed costs of conflict outbreak.

7   Concluding Remarks

We have studied in this paper a class of conflict problems of wide application. We adopted a mechanism design approach and showed that if the loss from actual conflict is too small, there is no incentive compatible and individually rational mechanism that assigns probability zero to conflict occurring. We also showed, by means of an example, that mechanisms that are ex ante efficient in the sense of minimizing the ex ante probability of conflict may be complicated objects that fail to exhibit intuitive monotonicity properties.

At least for some of the applications suggested in this paper, such as international conflict and warfare, the mechanism design perspective may be thought inappropriate. This approach presupposes a third party that can enforce a contract once the parties have agreed to be bound by it. This is crucial, since a mechanism may not be individually rational ex post. In the case of social contract applications, it raises the old question of "who watches the watchmen." Where does one find such a disinterested third party? Yet our analysis gives an upper bound on what voluntary agreement, i e, agreement without the possibility of binding contracts, can achieve. If the conditions of our central impossibility result hold, so that no mechanism exists that assigns probability zero to conflict, then certainly no potential for peaceful agreement can exist in the absence of a mechanism designer either.

This paper may be viewed as a partial investigation of a larger, more interesting problem, namely that of the endogenous allocation of resources between productive activities and activities geared toward establishing property rights. The study of this problem and its associated hold-ups due to contract incompleteness in the form of insecure property rights should be an important task for future research.

Appendix

Lemma 1 *For any mechanism (β, π), it is the case that*

$$\int_{T_2} \pi(t_1', t_2)[p(t_1, t_2) - p(t_1', t_2)]\theta \, dF_2 \leq U_1(t_1|t_1) - U_1(t_1'|t_1') \leq$$

$$\int_{T_2} \pi(t_1, t_2)[p(t_1, t_2) - p(t_1', t_2)]\theta \, dF_2$$

and

$$\int_{T_1} \pi(t_1, t_2')[p(t_1, t_2') - p(t_1, t_2)]\theta \, dF_1 \leq U_2(t_2|t_2) - U_2(t_2'|t_2') \leq$$

$$\int_{T_1} \pi(t_1, t_2)[p(t_1, t_2') - p(t_1, t_2)]\theta \, dF_1$$

for all $t_1, t_1' \in T_1$ and $t_2, t_2' \in T_2$.

Proof. By the definition of $U_a(t_a'|t_a)$ one has

$$
\begin{aligned}
U_1(t_1|t_1) &\geq \int_{T_2} \left[[1 - \pi(t_1', t_2)]\beta(t_1', t_2) + \pi(t_1', t_2)p(t_1, t_2)\theta\right] dF_2 \quad (2) \\
&= U_1(t_1'|t_1') + \int_{T_2} \pi(t_1', t_2)[p(t_1, t_2) - p(t_1', t_2)]\theta dF_2,
\end{aligned}
$$

and

$$
\begin{aligned}
U_1(t_1'|t_1') &\geq \int_{T_2} \left[[1 - \pi(t_1, t_2)]\beta(t_1, t_2) + \pi(t_1, t_2)p(t_1', t_2)\theta\right] dF_2 \quad (3) \\
&= U_1(t_1|t_1) + \int_{T_2} \pi(t_1, t_2)[p(t_1', t_2) - p(t_1, t_2)]\theta dF_2.
\end{aligned}
$$

Rearranging (2) and (3) yields the first result. An analogous argument for $a = 2$ completes the proof. $\square$

Proof of Proposition 4. Let (β, π) satisfy $\beta(t_1, t_2) = \beta^*$ and $\pi(t_1, t_2) = 0$ for all $(t_1, t_2) \in T$. Then $U_1(t_1|t_1) = \beta^*$ for all $t_1 \in T_1$ and $U_2(t_2|t_2) = 1 - \beta^*$ for all $t_2 \in T_2$. By the proof of Proposition 1, (β, π) is a peaceful mechanism if and only if $(V_1, V_2) := (\beta^*, 1 - \beta^*)$ satisfies (1). This is equivalent to requiring that $\beta^* \in I(\theta) := [\int_{T_2} p(\bar{t}_1, t_2)\theta dF_2, 1 - \int_{T_1}[1 - p(t_1, \bar{t}_2)]\theta dF_1]$.

Also, the proof of Proposition 1 shows that for any peaceful mechanism (β, π) there is a (V_1, V_2) with $V_1 = 1 - V_2 \in I(\theta)$ such that $V_a = U_a(t_a|t_a)$ for $a = 1, 2$ and all $t_a \in T_a$. This shows that (β, π) is payoff equivalent to the peaceful mechanism (β', π') with $\beta'(t_1, t_2) = V_1 = \beta^* \in I(\theta)$ for all $(t_1, t_2) \in T$.

Finally, the statements about $I(\theta)$ follow immediately from the definition of $I(\theta)$ and the definition of $\hat{\theta}$. $\qquad\square$

Proof of Proposition 5. Suppose the contrary. Then there is a subsequence (β_m, π_m) such that $\int_{T_2} \pi_m(t_1^*, t_2)\mathrm{d}F_2 \to k < 1$ for some $t_1^* > \underline{t}_1$ in the support of F_1. Let $U_a^m(t_a|t_a)$ denote the payoff of type t_a of agent a under the mechanism (β_m, π_m). Since $\int_{T_2} \pi_m(t_1^*, t_2)\mathrm{d}F_2 \to k < 1$, Lemma 1 implies that

$$\lim_{m \to \infty} [U_1^m(t_1^*|t_1^*) - U_1^m(t_1'|t_1')] < \int_{T_2} [p(t_1^*, t_2) - p(t_1', t_2)]\theta \mathrm{d}F_2 = \bar{U}_1(t_1^*) - \bar{U}_1(t_1')$$

for all $t_1' < t_1^*$. Since by individual rationality we have $U_1^m(t_1^*|t_1^*) \geq \bar{U}_1(t_1^*)$, this yields

$$\lim_{m \to \infty} U_1^m(t_1'|t_1') > \bar{U}_1(t_1') \tag{4}$$

for all $t_1' < t_1^*$. By the individual rationality constraint, $U_a^m(t_a|t_a) \geq \bar{U}_a(t_a)$ for $a = 1, 2$ and for all $t_a \in T_a$. Since $F_1(t_1') > 0$ for all $t_1' \in (\underline{t}_1, t_1^*)$, this together with (4) implies

$$\lim_{m \to \infty} \left[\int_{T_1} U_1^m(t_1|t_1)\mathrm{d}F_1 + \int_{T_2} U_2^m(t_2|t_2)\mathrm{d}F_2 \right] > \int_{T_1} \bar{U}_1(t_1)\mathrm{d}F_1 + \int_{T_2} \bar{U}_2(t_2)\mathrm{d}F_2. \tag{5}$$

By the definition of $U_a(t_a'|t_a)$ and individual rationality we have

$$\int_{T_1} U_1^m(t_1|t_1)\mathrm{d}F_1 + \int_{T_2} U_2^m(t_2|t_2)\mathrm{d}F_2 = 1 - \tilde{\pi}(\beta_m, \pi_m)(1 - \theta_m)$$

and

$$\int_{T_1} \bar{U}_1(t_1)\mathrm{d}F_1 + \int_{T_2} \bar{U}_2(t_2)\mathrm{d}F_2 = \theta_m.$$

Thus (5) implies that $\lim_{m \to \infty}[1 - \theta_m][1 - \tilde{\pi}(\beta_m, \pi_m)] > 0$, a contradiction to $\theta_m \to 1$. $\qquad\square$

23

References

[1] Geoffrey Blainey. *The Causes of War.* The Free Press, New York, 1973.

[2] Dagobert L Brito and Michael D Intriligator. Conflict, war, and redistribution. *American Political Science Review*, 79:943–957, 1985.

[3] James M Buchanan. *The Limits of Liberty: Between Anarchy and Leviathan.* The University of Chicago Press, Chicago, 1975.

[4] Ronald H Coase. The problem of social cost. *Journal of Law and Economics*, 3:1–44, 1960.

[5] Peter Coughlin. Candidate uncertainty and electoral equilibria. In J Enelow and M Hinich, editors, *Advances in the Spatial Theory of Voting.* Cambridge University Press, Cambridge, 1990.

[6] Joseph Farrell. Information and the Coase theorem. *Journal of Economic Perspectives*, 1:113–129, 1987.

[7] Herschel I Grossman and Minseong Kim. Swords or plowshares? A theory of the security of claims to property. *Journal of Political Economy*, 103:1275–1288, 1995.

[8] Sanford Grossman and Oliver Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94:691–719, 1986.

[9] Jack Hirshleifer. Anarchy and its breakdown. *Journal of Political Economy*, 103:26–52, 1995.

[10] Peter Klibanoff and Jonathan Morduch. Decentralization, externalities, and efficiency. *Review of Economic Studies*, 62:223–247, 1995.

[11] John O Ledyard and Thomas R Palfrey. Voting and lottery drafts as efficient public goods mechanisms. *Review of Economic Studies*, 61:327–355, 1994.

[12] John O Ledyard and Thomas R Palfrey. Interim efficiency in a public goods problem. Working paper, California Institute of Technology, 1996.

[13] George Mailath and Andrew Postlewaite. Asymmetric information bargaining problems with many agents. *Review of Economic Studies*, 57:351–367, 1990.

[14] Hans J Morgenthau. *Politics Among Nations*. Alfred Knopf, New York, 1966.

[15] Roger B Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47:61–74, 1979.

[16] Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29:265–81, 1983.

[17] Robert Powell. Bargaining in the shadow of power. *Games and Economic Behavior*, 15:255–289, 1996.

[18] Stergios Skaperdas. Cooperation, conflict, and power in the absence of property rights. *American Economic Review*, 82:720–739, 1992.

[19] Stergios Skaperdas and Constantinos Syropoulos. Insecure property and the stability of exchange. Working paper, University of California at Irvine, 1996.

[20] Kathryn E Spier. Pretrial bargaining and the design of fee-shifting rules. *Rand Journal of Economics*, 25:197–214, 1994.

[21] Oliver E Williamson. *The Economic Institutions of Capitalism*. Free Press, New York, 1985.