

# *Topics in Economic Analysis & Policy*

---

*Volume 2, Issue 1*

2002

*Article 2*

---

## Threats Without Binding Commitment

Steven Shavell  
Harvard Law School

Kathryn E. Spier  
Kellogg School of Management

*Topics in Economic Analysis & Policy* is one of *The B.E. Journals in Economic Analysis & Policy*, produced by The Berkeley Electronic Press (bepress).

Copyright ©2002 by the authors.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, bepress.

# Threats Without Binding Commitment

## Abstract

This paper explores the power of threats in the absence of binding commitment. The threatener cannot commit to carry out the threat if the victim refuses payment, and cannot commit not to carry out the threat if payment is made. An important assumption of the model is that once the threat is carried out it cannot be repeated. If exercising the threat is costly to the threatener, then the threat cannot succeed in extracting money from the victim. If exercising the threat would benefit the threatener, however, then the threat's success depends upon whether the threat may be repeated after a payment is made. In the equilibrium of a finite-period game, the threat is carried out and the victim makes no payments. In an infinite-horizon game, however, it is an equilibrium for the victim to make a stream of payments over time. The expectation of future payments keeps the threatener from exercising the threat.

## 1. INTRODUCTION

Threats are often used successfully to extract money or other things of value. A famous historical example concerns the Vikings, who in 991 stated that they would invade and plunder the English unless tribute was paid. The English submitted to the Viking demands and made a succession of payments of rather extraordinary magnitude. To raise the funds, the English King Ethelred imposed what may have been the first English tax paid in money, called the *Danegeld*.<sup>1</sup>

More generally, when would we expect threats to be successful? (By a *threat* we mean a threatener's expression of intent to behave in such a way as to lower a victim's utility unless the victim acts so as to raise the utility of the threatener.) For a threat to be successful, the victim must believe that the threat will be carried out if and only if he refuses to pay. As is well recognized, the threatener will therefore want to make a binding commitment to carry out his threat if and only if the demand is refused; this will make the threat credible and lead the victim to comply.<sup>2</sup>

What, however, if the threatener is not able to make a binding commitment concerning his threat? One presumes that the Vikings were not able to make a binding commitment to attack the English if and only if they did not comply; quite often the making of binding commitments may be difficult to arrange (as may be the equivalent of acquiring a reputation for carrying out threats).

This question about threats when binding commitments cannot be made is what we address here, using a simple model of threats. The answer that we will give depends on the consequence for the threatener of carrying out his threat: on whether carrying out his threat would impose a positive cost on the threatener, involve no cost to him, or provide a gain to him. (As the reader will see, the chief interest in our analysis concerns the third case.) An important assumption of the model, and one that distinguishes it from the repeated-game literature, is that the game ends once the threat is executed.<sup>3</sup>

The possibility that carrying out a threat involves a cost to the threatener is illustrated by a robber who threatens to assault a victim (and thereby incurs at least the risk of legal sanction) unless he hands over his wallet, by a buyer who threatens to switch to another seller (whose goods are not as well suited to the buyer's needs, imposing a cost on the buyer) unless the seller

<sup>1</sup> See the discussion of *Danegeld* in Hodgkin (1906), pp. 381-383. See also the first several lines of Kipling's (1940) "Dane-Geld," p. 658:

*It is always a temptation to an armed and agile nation,  
To call upon a neighbour and to say:--  
"We invaded you last night--we are quite prepared to fight,  
Unless you pay us cash to go away."*

*And that is called asking for Dane-geld  
And the people who ask it explain  
That you've only to pay 'em the Dane-geld  
And then you'll get rid of the Dane!*

<sup>2</sup> This is emphasized by Schelling (1960), pp. 35-43, in his influential writing on threats. For a discussion of the definition of threat, see also Klein and O'Flaherty (1993), and for a general, wide-ranging treatment of threats, see Milburn and Watman (1981).

<sup>3</sup> Many types of threats do have this feature, such as a robber threatening to kill or maim his victim or a blackmailer threatening to reveal embarrassing information. If the Vikings had carried out their threat to plunder and raze an English town, there would have been substantial losses of built-up wealth making future threats much less effective.

lowers price, by workers who threaten a strike unless employment conditions are improved, and by manifold other situations. Where a threat is costly to execute, a threat without a binding commitment would not be credible in the model, so the threat would be unsuccessful.<sup>4</sup>

The case in which executing a threat does not involve a cost for the threatener is approximately valid in a variety of circumstances. For example, suppose that a blackmailer threatens to reveal embarrassing information (the cost of revelation might be mere postage if the blackmailer is not worried about being caught) or that an extortionist threatens to burn down a building (the cost may be only that of a match if he is not worried about apprehension). Here, a threat without commitment may be credible because it would (just) be rational for the threatener to carry out his threat if and only if he is not paid. Therefore, the threat may be successful.

The situation in which proceeding with a threatened act would benefit the threatener is also often apposite. Suppose that a blackmailer would profit from revealing information (the victim might be a public figure and the information might be sold to the media) or suppose that a country (like Denmark) threatening to invade another would gain by appropriating the victim's land and wealth. In situations like these, the problem of the threatener is not that his threat is empty (the problem that the threatener faces when executing the threat is costly). Rather, the threatener's problem is that he will have an incentive to carry out his threat even if he is paid (the blackmailer's problem is that he will have an incentive to sell his information even if he is paid by the victim). Because this means that the victim will not prevent the threatened act by paying, he will not pay.<sup>5</sup> The threatener cannot overcome this problem in a single (or finite) period setting, and his threat will therefore fail in this version of the model.<sup>6</sup>

Suppose instead that the game is infinitely repeated (or that the threatener can make a threat with a probability in each period), and that the game ends once the threat is executed. When threats can be repeated -- as is often true, consider the typical blackmail threat and many invasion threats -- then the threatener can succeed. It is an equilibrium in the model for the victim to make a payment in each period, as long as the threat has not been carried out. In particular, the expectation of continued receipt of payments makes it rational for the threatener not to carry out his threat, and this very fact also makes it rational for the victim to continue his payments. (The present value of the payment stream is bounded between the value of immediately carrying out the threat to the threatener and the harm to the victim were it carried out.) It is interesting to note that the rationality in the model of the victim's making a stream of payments to prevent the execution of a threat has some counterpart in reality (including in the case of the English and the Vikings).<sup>7</sup>

<sup>4</sup> In contrast, in a standard repeated-game setting where the threat can be carried out more than once the threatener could succeed in extracting payments by establishing a reputation for carrying out costly threats.

<sup>5</sup> Accordingly, the threatener would like to commit *not* to act if he is paid, rather than, in the case where the threat is costly to execute, to commit *to* act if he is not paid.

<sup>6</sup> This point, and the possible solution to the threatener's problem in the repeated threat context to be described, is discussed informally in Shavell (1993).

<sup>7</sup> According to Hodgkin, p. 382, the English made seven payments to the Danes over the years 991 to 1014. When one reads contemporary accounts of bribery, extortion, and blackmail, a pattern of repeated payments is often evident; see, for instance, the discussion of the extortion of corporations overseas in Jacoby, Nehemkis, and Eells (1977), pp. 107-112, the discussion of blackmail in Hepworth (1975), p. 42, and the description of the apparent blackmailing of U.S. Housing Secretary Henry Cisneros in *The New York Times*, September 22, 1994, page 18 column 1.

## 2. THE MODEL

A risk-neutral threatener poses a threat to a risk-neutral victim. If the threat is carried out, the victim suffers a once-and-for-all loss,  $v > 0$ , while the threatener derives a gain,  $b$ . Three cases are considered:  $b > 0$ ,  $b = 0$ , and  $b < 0$ ; the motivation for each of these cases was discussed in the introduction. We assume that  $v > b$  or, equivalently, that carrying out the threat is inefficient: the victim's loss exceeds the threatener's gain. There is a common discount factor,  $\delta \in [0, 1)$ , for the two players.

The timing of the game is simple: in each period,  $t$ , the victim chooses a payment,  $p_t \geq 0$ , to make to the threatener, after which the threatener must decide whether to carry out the threat. If the threat is carried out, the game ends; the victim suffers a loss,  $v$ , and the threatener gains  $b$  (in addition to the stream of past payments,  $p_s$ , where  $s \leq t$ .) If the threatener does not exercise the threat, the game proceeds to period  $t+1$ , and the stage game is repeated.

The analysis is broken into two parts. First, we consider a game with a known finite horizon. Next, we consider the infinitely repeated game. This latter case is formally equivalent to the situation where the game will end with a known probability in each period; the discount factor may be interpreted as the probability that the game will continue into the following period.

### 2.1 Finite Horizon

Suppose that the game has a known finite horizon,  $T$ . To start, if the threatener's gain from carrying out the threat is strictly positive,  $b > 0$ , then there is a unique subgame-perfect equilibrium where no payments are made and the threat is carried out immediately. This outcome can be derived by backwards induction. In the last period, the threatener would strictly prefer to carry out the threat regardless of the previous payments made by the victim. Knowing that his payments cannot influence the threatener's actions, the victim would pay  $p_T = 0$  in period  $T$ . Working backwards, in period  $T-1$  the threatener anticipates receiving  $p_T = 0$  in the following period and therefore carries out the threat immediately. (He prefers to receive a payoff of  $b$  in period  $T-1$  than wait a period to receive the same payoff.) Anticipating this, the victim pays  $p_{T-1} = 0$ . Repeating this argument verifies that the victim never pays a cent and the threatener carries out the threat at his first opportunity.

When  $b > 0$ , this unique subgame-perfect equilibrium is clearly inefficient. Since the victim's loss exceeds the threatener's gain, it would be better for the players to negotiate around the inefficiency and avoid the dead-weight loss,  $v-b$ , associated with carrying out the threat. In the absence of binding commitment, however, this cannot be accomplished in a finitely repeated game.

The inefficient outcome described above is also an equilibrium when the threatener derives no benefit at all from carrying out the threat,  $b = 0$ . If the threatener does not expect to receive any payments, then he is indifferent about carrying out the threat in each period. If the victim expects the threatener to carry out the threat, then he won't make any payments. However, the threatener's strategy of carrying out the threat in the first period is weakly dominated when  $b = 0$ : The threatener would be better off not carrying out the threat in the (off-the-equilibrium-path) event that the victim makes some payment in the future.

When  $b = 0$ , there are many subgame-perfect equilibria that survive the refinement of eliminating weakly dominated strategies. All of the strategies in the refined set of equilibria must have the feature that the threatener never carries out a threat prior to the last period. Since the threatener has the last move, it is reasonable for him to carry out the threat at the very end. The proposition characterizes a set of efficient equilibria for  $b = 0$  where the threatener is

successful in extracting a constant stream of payments from the victim, and never carries out the threat on the equilibrium path.<sup>8</sup>

Finally, if it is costly for the threatener to carry out the threat,  $b < 0$ , then there is a unique subgame-perfect equilibrium where no payments are made and the threat is never exercised. The threatener cannot credibly commit to carry out his threat in period  $T$ , and so the victim will shrug off the threat and refuse to pay. Similarly, the threatener's threats are not credible in earlier periods.

These results are summarized in the following proposition.

*Proposition 1: Suppose that the horizon,  $T$ , is finite.*

- i. *If  $b > 0$ , there is a unique subgame-perfect equilibrium. The victim pays nothing to the threatener, and the threatener carries out the threat in period 1. The payoffs to the victim and the threatener, respectively, are  $(-v, b)$ .*
- ii. *If  $b = 0$ , then for each  $\theta \in [0, 1]$  there is a subgame-perfect equilibrium (which satisfies elimination of weakly dominated strategies) where the victim pays  $p^*(\theta) = [(\delta^{T-1} - \delta^T)/(1 - \delta^T)]\theta v$  to the threatener in each period, and the threatener never carries out the threat. The payoffs to the victim and the threatener, respectively, are  $(-\delta^{T-1}\theta v, \delta^{T-1}\theta v)$ .*
- iii. *If  $b < 0$ , there is a unique subgame-perfect equilibrium. The threatener never carries out the threat and the victim makes no payments. The payoffs to the victim and the threatener, respectively, are  $(0, 0)$ .*

*Proof:*

The cases of  $b < 0$  and  $b > 0$  are easily proven by backwards induction, as described in the text.

For the case of  $b = 0$ , the following strategies support the subgame-perfect equilibrium described in the proposition for a given  $\theta$ : If  $p_t < p^*(\theta)$  in some period  $t$ , then the threatener waits and exercises the threat in period  $T$  with probability  $\theta \in [0, 1]$ . If  $p_t \geq p^*(\theta)$  for all  $t$ , then the threatener does not exercise the threat. The victim pays  $p_t = p^*(\theta)$  if  $p_s \geq p^*(\theta)$  for all  $s < t$ , and pays  $p_t = 0$  if  $p_s < p^*(\theta)$  for some  $s < t$ .

First, we check that the victim's strategy is a best response to the threatener's strategy. Assuming that the victim has not paid less than  $p^*(\theta)$  in the past, the victim can do no better than pay  $p^*(\theta)$  in each subsequent period. Paying  $p_t > p^*(\theta)$  does not reduce the chance that the threat is carried out (if the victim keeps on schedule by paying  $p^*(\theta)$ , the probability is zero). By paying  $p_t < p^*(\theta)$ , he guarantees that the threat will be exercised with probability  $\theta$  in period  $T$ . (Note that since the victim expects that the threat will be carried out with probability  $\theta$  regardless of any (possibly conciliatory) actions that the victim may take, once the victim has deviated from the equilibrium by paying less than  $p^*(\theta)$ , he will pay nothing in all subsequent periods.) The victim will not deviate from this strategy in period  $T$  since  $p^*(\theta) =$

<sup>8</sup> There are many other subgame-perfect equilibria that satisfy the refinement of eliminating weakly dominated strategies. For example, there is an *inefficient* equilibrium where the victim makes no payments and the threatener carries out the threat in the very last period. There is also an *efficient* equilibrium where the victim pays  $v$  in the last period and nothing before, since it is credible for the threatener to exercise the threat unless the victim complies.

$[(\delta^{T-1} - \delta^T)/(1 - \delta^T)]\theta v < \theta v$ .<sup>9</sup> Working backwards through each period, we finally see that he will not deviate in the first period since  $p^*(\theta) + \delta p^*(\theta) + \dots + \delta^{T-1} p^*(\theta) \leq \delta^{T-1} \theta v$ , or  $p^*(\theta) \leq [(\delta^{T-1} - \delta^T)/(1 - \delta^T)]\theta v$ .<sup>10</sup>

Given the victim's strategy, we can check that the threatener's strategy is a best response. Suppose that  $t < T$ , and that the victim has made payments of  $p^*(\theta)$  (or above) in the past. The threatener expects that the victim will continue to make payments in the future, and therefore strictly prefers not to exercise his threat because this would end the game and cut off payments. (If  $t = T$ , then the threatener is indifferent between exercising the threat and not.) If the victim made a smaller payment in the past, then the threatener does not expect any more payments in the future, and is indifferent between exercising the threat and not. If  $t < T$ , then exercising the threat in period  $t$  violates weak dominance; it is acceptable, however, for the threatener to wait until period  $T$  to exercise the threat.

Q.E.D.

## 2.2 Infinite Horizon

We saw in Proposition 1 that when the game is of a known finite duration and  $b > 0$ , the unique subgame-perfect equilibrium is inefficient: The threat is carried out immediately even though the victim's loss exceeds the threatener's gain. We begin by observing that this inefficient outcome is also an equilibrium of the infinite horizon game when  $b > 0$ . Since the victim believes that the threatener will carry out the threat at his first opportunity, he has no incentive to pay the threatener. Since the threatener does not expect to receive any money from the victim, it is rational for him to carry out the threat.<sup>11</sup> We will refer to this equilibrium as the *stationary equilibrium*.<sup>12</sup>

However, when  $b > 0$  there are also many *efficient* subgame-perfect equilibria with successful threats; payments are made by the victim, and threats are never carried out in equilibrium. In particular, we can use the inefficient stationary equilibrium to construct a set of efficient equilibria with a constant payment stream over time,  $p_t = p^*$ , where  $p^* \in [(1 - \delta)/\delta]b, (1 - \delta)v$ . The motivation is as follows: If the victim deviates by paying less than  $p^*$ , the threatener believes that the victim will pay nothing in the future, and the victim believes that the threatener will carry out the threat at his first opportunity. In other words, after a deviation, the continuation equilibrium automatically "switches" to the inefficient stationary equilibrium. In this way, the victim is punished for paying less than  $p^*$ .

With these strategies, we see that any division of the surplus,  $v - b$ , is possible in equilibrium. When the per-period payment is at its lower bound,  $p^* = [(1 - \delta)/\delta]b$ , the threatener does no better than in the stationary equilibrium. His present discounted payoff is  $b$ , and the victim captures the surplus,  $v - b$ . When payment is at the upper bound,  $p^* = (1 - \delta)v$ ,

<sup>9</sup> This is true because  $\delta^{T-1} - \delta^T < 1 - \delta^T$ .

<sup>10</sup> Note that the equilibrium was constructed to have the property that the victim is indifferent about deviating in the first period.

<sup>11</sup> The following strategies support this equilibrium: The victim pays  $p_t = 0$  regardless of the history of the game, and the threatener exercises his threat with probability one at each decision node.

<sup>12</sup> An equilibrium is stationary when the players' strategies are the same at each node of the game, both on and off of the equilibrium path. The stationary equilibrium described in the text is the only stationary equilibrium of the game when  $b > 0$ .

the threatener is extracting all of the surplus (his payoff is  $v$ ); the present discounted value of the payment stream equals what the victim would lose if the threat were actually exercised.<sup>13</sup>

The situation is quite different when  $b = 0$ . Eliminating weakly dominated strategies establishes that the threatener will not exercise his threat at any decision node. (If the victim were to "tremble" and make a future payment, the threatener would be better off not having exercised the threat.) Since the threatener's threats are never credible, the victim will refuse to pay the threatener.

Finally when  $b < 0$ , the foregoing argument is even stronger since the threatener is always strictly better off not exercising the threat. Since the threatener does not have reputation concerns or other means of committing himself to carrying out the threat, when  $b < 0$  the threatener cannot succeed in extracting payments. In the unique subgame-perfect equilibrium, no payments are made and the threat is never carried out.

*Proposition 2: Suppose that the horizon is infinite.*

- i. *If  $b > 0$ , then for each  $p^* \in [(1-\delta)/\delta]b, (1-\delta)v$  there exists a subgame-perfect equilibrium in which the victim pays  $p^*$  to the threatener in each period, and the threatener never carries out the threat. The payoffs to the victim and the threatener respectively, are  $(-p^*/(1-\delta), p^*/(1-\delta))$ .*
- ii. *If  $b = 0$ , then there is a unique subgame-perfect equilibrium that survives the elimination of weakly dominated strategies. The threatener never carries out the threat and the victim makes no payments. The payoffs to the victim and the threatener, respectively, are  $(0, 0)$ .*
- iii. *If  $b < 0$ , there is a unique subgame-perfect equilibrium. The threatener never carries out the threat and the victim makes no payments. The payoffs to the victim and the threatener, respectively, are  $(0, 0)$ .*

*Proof:*

The cases of  $b = 0$  and  $b < 0$  are as described in the text. The case of  $b > 0$  will be proven here.

First we will construct upper and lower bounds on the per-period payment,  $p^*$ . Clearly, the threatener would strictly prefer to carry out the threat if  $b > \delta p^* + \delta^2 p^* + \dots$ . The victim would not be willing to pay  $p^*$  in each period if  $v < p^* + \delta p^* + \delta^2 p^* + \dots$ , or if the loss from having the threat carried out is smaller than the present discounted value of the payments.<sup>14</sup>

<sup>13</sup> This equilibrium resembles that in Klein and Leffler's (1981) analysis of product quality in an infinitely repeated game. There, quality is not observable to the consumer at the time of purchase, so the firm has a natural tendency to reduce quality. In equilibrium, however, the firm has an incentive to maintain high quality in order to protect its reputation. If the firm were to "cheat" on quality, consumer confidence would be lost and future profits would be forfeited. While Klein and Leffler's analysis is a true repeated game, the game presented here is not because it ends when the threat is carried out. A second important difference is that we are concerned with negotiations. So long as the threat is not carried out, all outcomes are equally efficient.

<sup>14</sup> The timing of the game specifies that in each period the victim decides upon the payment, after which the threatener chooses whether or not to carry out the threat. If the threatener does not carry out the threat, then the next payment he will receive is in the *following* period. This is why the first term in the threatener's inequality is discounted by  $\delta$ . If the victim refuses to pay, the threatener will carry out the threat in the same period; that is why, in the victim's inequality, neither  $v$  nor the first payment are discounted.



Rearranging these expressions tell us that any per-period payment must fall within a certain range:  $[(1-\delta)/\delta]b \leq p^* \leq (1-\delta)v$ .<sup>15</sup>

The following strategies support the subgame-perfect equilibrium for each  $p^*$  in this range. The threatener does not carry out the threat in period  $t$  if  $p_s \geq p^*$  for  $s \leq t$ . If  $p_s < p^*$  for some  $s \leq t$ , then the threatener carries out the threat immediately. The victim pays  $p_t = p^*$  if  $p_s \geq p^*$  for all  $s < t$ , and pays  $p_t = 0$  if  $p_s < p^*$  for some  $s < t$ .

The victim cannot do better than pay  $p^*$ : By paying  $p > p^*$  he does not increase the chance of acceptance (it is already unity), and by paying  $p < p^*$  he is (weakly) worse off since the threat will be carried out,  $p + v \geq p^*(1-\delta)$ .<sup>16</sup> (The left hand side denotes the payoff from deviating, and the right hand side is the present discounted value of the equilibrium payments.) So long as  $p_s \geq p^*$  for all  $s \leq t$  the threatener will not carry out the threat because he expects payments to continue in the future and  $b \leq p^*[\delta/(1-\delta)]$ . If  $p_s < p^*$  for some  $s \leq t$ , then the threatener expects no payments in the future and strictly prefers to carry out the threat (since  $b > 0$ ).

Q.E.D.

It is interesting to compare the results from Proposition 2 to those from Proposition 1. In the previous section, we saw that when the horizon was finite, successful threats were possible when  $b = 0$  but not when  $b > 0$ . When  $b = 0$ , the threatener could credibly threaten to carry out the threat in the last period unless payments were made. When  $b > 0$ , this threat was not effective, since the victim knew that the threat would be carried out regardless of his previous payments. When the horizon is infinite, however, these results are reversed: successful threats are possible when  $b > 0$  but not when  $b = 0$ . When  $b > 0$ , the threat is effective in extracting payments since there is another equilibrium where the victim makes no payments and the threat is carried out. A deviation by the victim will trigger switching to this other equilibrium. When  $b = 0$ , the threat is not effective since the strategy of exercising the threat is weakly dominated at all of the threatener's decision nodes.

### 3. CONCLUDING COMMENTS

In conclusion, we offer several remarks about issues that we did not consider.

(a) *Asymmetric information.* We assumed in our analysis that information was symmetric, and this led, among other things, to the conclusion that in many situations threats were made only when they would be successful, so that threats were never actually carried out. If we instead had allowed for asymmetry of information about the threatener's cost or benefit from carrying out a threat and about the victim's harm from the execution of a threat, we would instead have found, realistically, that threats would sometimes be carried out when we indicated they would not be. Notably, a threatener would sometimes make a demand that would be rejected by the victim either because the demand exceeded the true harm the victim would suffer from the threat or else because the victim would incorrectly surmise that the threatener would bear a cost from carrying out his threat.

<sup>15</sup> Note that a range exists only when  $\delta > b/v$ . If the discount factor were very small, then the payers would be myopic. The threatener would prefer to exercise the threat and gain  $b$  than to wait for future payments from the victim.

<sup>16</sup> This is true because  $p^* < (1-\delta)v$ . Note that if  $p^* = (1-\delta)v$  then the victim is indifferent between continuing to make payments and paying  $p = 0$ . For all other cases, however, the victim is strictly worse off when he deviates and pays  $p < p^*$ .

(b) *Emotional motivation.* A plausible emotional motivation provides a reason to believe that a threat might be credible without a binding commitment. Suppose that a threatener will become angry with the victim if his demand is rejected, and thus enjoy carrying out his threat, whereas the threatener will not be angry (and possibly will be favorably disposed toward the victim) if the victim pays him. If this is so and the victim understands that, the threatener's psychological makeup will itself lead him to carry out his threat if and only if he is not paid, so that his threat will be credible.<sup>17</sup>

(c) *Social undesirability and possible illegality of threats.* We considered the question how threats may succeed but did not emphasize the issue of their influence on social welfare. The influence of threats on social welfare may come about in part through ex ante effects: activity on the part of potential victims to avoid becoming prey to threats, and activity on the part of potential threateners to position themselves so as to be able to carry out threats. Such activities may represent a social waste (the activities of extortionists who would threaten to burn down a store if not paid off) or may constitute a social benefit (suppose victims curtail their socially undesirable, possibly criminal acts so that they cannot be blackmailed). Other avenues through which threats affect social welfare are the making of payments to threateners and the actual execution of threats (which may come about through asymmetric information, as just discussed, and also because a threatener would gain from his threat, as in Proposition 1i). These effects too may be socially desirable (for instance, the blackmail payment made by a criminal might be desirable as it would tend to deter crime) or undesirable. In consequence, it can be argued that some types of threat are fairly clearly socially undesirable and ought to be, and are, illegal, while other types of threat are different.<sup>18</sup>

### Acknowledgments

Shavell acknowledges the support of the John M. Olin Center for Law, Economics, and Business at Harvard Law School, and Spier acknowledges summer support from the Kellogg Graduate School of Business, Northwestern University. The authors thank James Dana, Peter Diamond, Louis Kaplow, and Edward Schwartz for comments.

<sup>17</sup> The main point of this paragraph is stressed by Hirshleifer (1987). See also Rabin (1993) and Geanakoplos, Pearce and Stacchetti (1989) for a more general treatment of the role of psychology and emotions in game theory.

<sup>18</sup> For discussion along the lines just sketched of the possible social undesirability and illegality of threats, and especially of blackmail, see Ginsburg and Shechtman (1993), Posner (1993), and Shavell (1993).

## References

- Geanakoplos, John, David Pearce, and Ennio Stacchetti, 1989, "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1: 60 - 79.
- Ginsburg, Douglas H. and Paul Shechtman, "Blackmail: An Economic Analysis of the Law," 1993, *University of Pennsylvania Law Review*, 141: 1849-1876.
- Hepworth, Mike, 1975, *Blackmail: Publicity and Secrecy in Everyday Life*, Routledge, London.
- Hirshleifer, Jack, 1987, "On the Emotions as Guarantors of Threats and Promises," 307-326, in John Dupre, editor. *The Latest on the Best: Essays on Evolution and Optimality*. MIT Press, Cambridge, MA.
- Hodgkin, Thomas, 1906, *The History of England From the Earliest Times to the Norman Conquest*, Longmans, Green, and Co., London.
- Jacoby, Neil H., Peter Nehemkis and Richard Eels, 1977, *Bribery and Extortion in World Business: A Study of Corporate Political Payments Abroad*, Macmillan, New York.
- Kipling, Rudyard, 1940, *Rudyard Kipling's Verse, 1885-1926*, Doubleday, New York.
- Klein, Benjamin, and Keith B. Leffler, "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 81: 615 - 641.
- Klein, Daniel B. and Brendan O'Flaherty, 1993, "A Game-Theoretic Rendering of Promises and Threats" *Journal of Economic Behavior and Organization*, 21: 295-314.
- Milburn, Thomas W. and Kenneth H. Watman, 1981, *On the Nature of Threat*, Praeger, New York.
- Posner, Richard A., 1993, "Blackmail, Privacy, and Freedom of Contract," *University of Pennsylvania Law Review*, 141: 1817-1843.
- Rabin, Matthew, 1993, "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83: 1281 - 1302.
- Schelling, Thomas C. 1960, *The Strategy of Conflict*, Harvard University Press, Cambridge.
- Shavell, Steven, 1993, "An Economic Analysis of Threats and Their Illegality: Blackmail, Extortion, and Robbery," *University of Pennsylvania Law Review*, 141: 1877-1903.

