

Deterrence and Compellence

We begin our foray into the substantive areas of IR, quite appropriately, by looking at an important issue that has not only guided U.S. foreign policy since the end of the Second World War, but that is of great relevance today as well.

Deterrence is a state of being characterized by the absence of war between two opponents and involves an effort to persuade at least one of the opponents not to take an action contrary to the interests of the other by convincing that opponent that doing so would not be worth the effort. Deterrence is an attempt to manipulate the opponent's incentives to challenge the status quo, and it is the business the United States got itself into at the end of the Second World War when it was decided that the Soviet Union's presumed expansion had to be contained for good.

Compellence is similar to deterrence and is yet conceptually distinct. Like deterrence, compellence seeks to manipulate the incentives of the opponent in order to affect his behavior. Unlike deterrence, which is concerned with persuading the opponent not to take some action, compellence is an effort to persuade him to change his action.

Deterrence succeeds when the opponent's expected utility of inaction exceeds his expected utility of action. Compellence succeeds when the opponent's expected utility of changing his action exceeds his expected utility of continuing with his present course. Compellence usually occurs in the wake of failed deterrence. Unlike deterrence, compellence requires the opponent to make concessions or suffer the consequences.

All wars are about compellence: each side is trying to persuade the other that surrendering is better than continuing to fight. That is, the wartime status quo is worse for the other side than capitulating.

1 The Model

Although deterrence and compellence are usually studied separately, we shall study them in one model. We shall see that the two concepts are really quite similar when all is said and done and in fact may not be as easy to distinguish as some claim.

Figure 1 shows a simple model. We shall use $u_i(\cdot)$ to denote the utility for player i for some outcome; i can be either C for the Challenger or D for the Defender. We shall use B_i^o to denote the benefits for player i from outcome o ; and, similarly, we shall use K_i^o to denote the costs for player i from outcome o . For example, B_C^Q denotes the benefits for the challenger for the status quo; K_D^F denotes the costs for the defender for fighting; and $u_D(SQ)$ denotes the utility of the defender for the status quo. As usual, we shall use $EU_i(\cdot)$ to denote the expected utility for player i . The payoffs then are as follows:

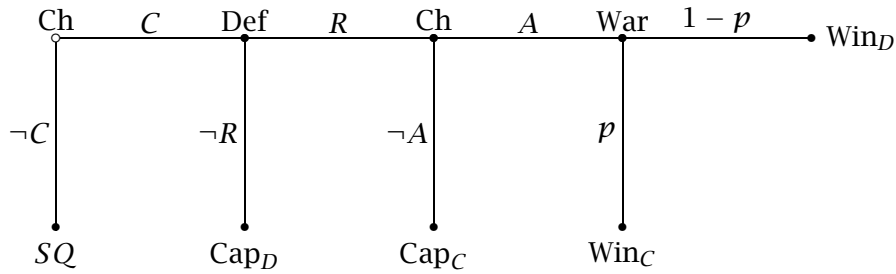


Figure 1: The Deterrence/Compellence Model.

- SQ (status quo): $u_i(SQ) = B_i^Q - K_i^Q$.
- Cap_D (capitulation by the defender): $u_i(Cap_D) = B_i^U - K_i^U$; we shall assume that $K_C^U = 0$ and $K_D^U > 0$; that is, the challenger pays no costs from unopposed gain, and the defender pays (reputational) costs associated with not following through on a promise to resist.
- Cap_C (capitulation by the challenger): $u_i(Cap_C) = B_i^Q - K_i^B$; we shall assume that $K_D^B = 0$ and $K_C^B > 0$; that is, the defender pays no costs if the challenger backs down, and the challenger pays positive costs from retreating under duress
- Win_C (military victory by the challenger): $u_C(Win_C) = B_C^V - K_C^F$, benefits of victory net of the costs of fighting; $u_D(Win_C) = B_D^L - K_D^F$, benefits of loss net of the costs of fighting.
- Win_D (military victory by the defender): $u_C(Win_D) = B_C^L - K_C^F$, benefits of loss net of the costs of fighting; $u_D(Win_D) = B_D^V - K_D^F$, benefits of victory net of the costs of fighting.

The defender wants to deter the challenger from attacking or, having failed that, compel it to back down short of war. To do so, it must influence the challenger's expected utility calculation. That is, the defender must convince the challenger that attacking is not worth it. How? As before, we solve the game by looking forward and reasoning backward (subgame perfection by backward induction).

1.1 The Compellence Subgame

Let's begin from the question of what happens if war should occur. The challenger's expected utility from war is:

$$EU_C(\text{War}) = p u_C(\text{Win}_C) + (1 - p) u_C(\text{Win}_D) = p B_C^V + (1 - p) B_C^L - K_C^F$$

and the defender's expected utility from war is:

$$EU_D(\text{War}) = p u_D(\text{Win}_C) + (1 - p) u_D(\text{Win}_D) = p B_D^L + (1 - p) B_D^V - K_D^F$$

We shall assume that winning is better than losing, and so $B_i^V > B_i^D$ for both players. Since the expected utility of war has four components— p , B_C^V , B_C^L , and K_C^F —the defender can reduce the expected value of war to the challenger by manipulating these components:

1. decrease the benefits of winning B_C^V by scorched earth policy (e.g. the Dutch threaten to blow up the dikes in case of invasion inundating the land and reducing its economic value to invader even if the invasion succeeds);
2. decrease the “benefits” of losing B_C^L by threatening to punish the loser (e.g. laying waste to the country once it has been defeated);
3. increase the costs of fighting K_C^F by punishing his civilian population or destroying its military installations;
4. decrease the probability of winning p by creating a formidable military with high morale that can defeat him.

The defender can manipulate any of these components to reduce the challenger’s expected utility of fighting. However, after the Defender chooses to resist, war may not immediately follow. The Challenger can still back down and avoid fighting. It will do so as long as

$$EU_C(\neg A) > EU_C(\text{War}).$$

It is evident from this equation that reducing the expected value of war is one way of compelling the Challenger to back down. However, another (frequently neglected) option is to *increase* its payoff to avoiding war. That is, increase the Challenger’s expected utility from capitulating.

Although diplomats are quite aware of leaving the opponent the ability to retreat while “saving face,” many analysts focus too much on the military option and neglect to take into account the existence of this component to compellence. If retreat is made bearable, then maybe one does not have to even threaten so much with war. The worse the retreat option for the Challenger, the more the Defender has to threaten to do in war to compel him to back down. Since fighting is costly for the Defender as well, it may do well to manipulate Challenger’s payoff to backing down instead. Putting this in symbols yields

$$B_C^Q - K_C^B > pB_C^V + (1 - p)B_C^L - K_C^F$$

If this condition is met, then the Challenger would back down if threatened and compellence would succeed. Keep this equation in mind when you read about the U.S. compelling its adversaries to back down after mounting a threat to the status quo. Note in particular how the costs of fighting and the costs of backing down have different effects on this decision.

We have seen how the defender can reduce the challenger’s expected utility of war or increase the expected utility of backing down, thus making attacking less preferable. However, given a challenge, the Defender has its own problem to solve: Should it resist it or not? If it does resist, then we’re in the compellence subgame already analyzed. If it does not resist, then the game ends with its capitulation.

1.2 The Decision to Resist: Credibility

The Defender will resist as long as the expected utility from doing so exceeds the expected utility to capitulating to the challenger's demands:

$$EU_D(R) > u_D(\text{Cap}_D).$$

The last inequality illustrates well the Defender's credibility problem. In order to credibly threaten with resistance, it must either *increase* the expected utility of fighting or *decrease* the expected utility of backing down.

Calculating $EU_D(R)$, however, depends on whether the challenger will attack or not. We need quite a bit more of game theory to analyze this properly, but I will give you some intuition with a simplification. Let q denote the defender's belief that the challenger will attack. That is, q is the defender's subjective probability of an attack occurring after resistance. Since q is a probability, $1 - q$ is the corresponding belief that the challenger will not attack. We can write the expected utility as follows:

$$EU_D(R) = qEU_D(\text{War}) + (1 - q)u_D(\text{Cap}_C).$$

We have the expression of the expected utility of war already, which yields

$$q \left[pB_D^L + (1 - p)B_D^V - K_D^F \right] + (1 - q)B_D^Q > B_D^U - K_D^U$$

The left-hand side of this inequality is a weighted average of the expected utility of war and the utility of capitulation by the challenger. We have already seen how the defender can manipulate the costs/benefits of war, and now we add the possibility of manipulating the belief q :

- increase the probability of victory (that is, decrease p); the good news about manipulating this parameter is that it simultaneously decreases the opponent's probability of victory, lowering its expected utility of fighting, and therefore making it easier to compel to back down;
- increase the benefits of winning (e.g. occupation of territory and economic exploitation); the problem with manipulating this parameter is that doing so may conflict with decreasing the challenger's benefits of losing, which, as you should recall, was one of the methods of decreasing its expected utility of fighting; so, threatening to lay waste to the challenger's country may work by reducing its value of fighting but it also reduces the defender's value of victory.
- decrease the costs of fighting; again, the problem with this is that some of the methods of doing so may actually decrease the challenger's costs of fighting as well; on the other hand, modern technology permits the so-called "smart" weapons, which allow the U.S. to simultaneously reduce its costs of fighting while increasing the costs of its opponents; in the First Gulf War, one of the problems of the coalition and one of the reasons it was not successful in compelling Saddam to leave Kuwait short of war was that nobody really knew how effective these weapons would be. Many analysts predicted heavy casualties for the U.S.-led coalition, and there's some evidence that Saddam may have hoped to draw it out in a costly battle and thereby split it apart.

- lower the belief q that the challenger will resist; this is difficult to do but in effect it allows the defender to resist when the expected payoff from fighting is very low; it does so by claiming that the Challenger is not likely to resist, and so war (the bad outcome) is not likely. Giving additional options to the challenger and increasing the payoff from backing down strengthens this tactic. It is important to remember that *one way of enhancing one's own credibility is by destroying the credibility of the opponent's commitments.*

However, these are not the only ways for the Defender to enhance its own credibility. The other one, which should not be surprising to you by now given all that bargaining we've covered, is to reduce its own expected utility from backing down, either by decreasing the benefits associated with the new status quo or increasing the costs of backing down. Looking back at the Defender's inequality, it is clear that even if the expected utility of fighting is low, there will be no credibility problem as long as the utility from backing down is reduced sufficiently. So, the Defender can improve its threat by *imposing costs on itself* in case it backs down.

This is the familiar approach of constraining one's choices. In this scenario, the Defender makes the backing down option completely unattractive and so will have no choice but fight. This, of course, gives it credibility to resist in the first place.

How can it constrain its choices? By staking its reputation on resistance. In the Cuban Missile Crisis, for example, the U.S. publicly proclaimed its policy and staked its reputation on successful resistance to the Soviets. During the Cold War, the U.S. stationed a token force of its best troops in Europe. These had no chances of stopping the Red Army should the Soviets decide to roll, but they did stake the reputation of the U.S. by making virtually sure that the Americans would have to respond to an invasion that kills thousands of their best men. Tying your hands in this way helps.

We are now back to Challenger's initial move: Should it challenge the status quo or should it just sit tight and do nothing?

1.3 The Deterrence Subgame

The Challenger will be deterred when the expected utility of challenging is less than the expected utility of living with the status quo:

$$EU_C(C) < EU_C(-C).$$

There are two parts to the calculation on the left-hand side. First, the challenger must calculate the probability that the defender will resist and the probability that the defender will capitulate (giving the challenger the payoff from unopposed gain). Second, given the probability of war occurring, the challenger must calculate the expected utility of fighting.

Let r be the probability that the defender resists an attack and let q be the probability that the challenger attacks after resistance as before. The expected utility from challenging then is

$$EU_C(C) = r[qEU_C(A) + (1 - q)u_C(\text{Cap}_C)] + (1 - r)u_C(\text{Cap}_D)$$

Thus, r represents the defender's credibility problem, and q represents the challenger's credibility problem. The less credible the defender's threat (the lower the r), the less likely resistance becomes and the higher the payoff from challenging. The less credible the challenger's threat (the lower the q), the less likely war becomes and the lower the payoff from challenging.

As we saw in the previous two cases we discussed, the defender can take steps to manipulate r , which is really the challenger's belief that it will in fact resist. Naturally, the defender may try bluffing here as well. However, obvious commitment of the form we've seen works very well.

Influencing expectations is very important here, and it's all about beliefs. Note in particular that one way the defender can improve its credibility is by denying credibility to the challenger. That is, by claiming that it believes that q is very low. By claiming to believe that the challenger is unlikely to attack when resisted, the defender in fact claims that it believes that its payoff from resistance is higher, and so makes the threat to resist more credible. Now, whether it makes this claim in good faith or not is an important question. Here it can claim that it simply does not understand the challenger's commitment or that it does not believe it anyway. Again, denying the opponent's commitment may enhance your own credibility.

We have already seen how the defender can manipulate the expected utility of war and the probability that it will resist, which leaves:

- reduce value of unopposed gain B_C^U by threatening to escalate the arms buildup and force the victorious opponent into an expensive arms race;

As we saw above, the Challenger faces its own credibility problem in the compellence subgame: it may not be in its interest to attack given resistance by the Defender. Its problem, of course, is that if it lacks credibility in the compellence game, it is unlikely to cause the Defender to back down without resistance in the first place. In other words, this increases the likelihood that when it challenges, the Defender will resist, and the Challenger will have to back down.

The Challenger's relevant calculation, however, is not simply the expected utility of challenging. It is the comparison between that and the utility of the status quo. Since in the absence of challenge the status quo prevails with probability one, the expected utility of the status quo is simply its net value, $B_C^Q - K_C^Q$.

The Defender can manipulate the net value of the status quo and make challenges less attractive, thereby enhancing the prospects of deterrence:

- increase the benefits of the status quo B_C^Q (e.g. agree to some demands, secure challenger's membership in some organization that might be beneficial to him);
- reduce the costs of the status quo K_C^Q (e.g. agreements to eliminate certain expensive types of weapons, reduction of inventories).

Now, even this simple setup shows how complicated deterrence and compellence can be, mostly because actions that enhance one's position in one of the stages may undermine it at another. Here's a common problem with arms races. On one hand, arming improves

one's probability of winning, and therefore enhances the credibility of the threat to go to war. This means that a threat should result in capitulation depending on which side is the one with the more formidable military. Given this logic, an arming by one side immediately makes the other one less secure, and gives it incentives to arm as well. This is the familiar *Security Dilemma* that I've mentioned already. You can probably expect a spiral of arming, an arms race, until "something" happens. That something could be a war, or exhaustion.

The problem from our perspective here is that an arms race increases the costs of the status quo because now the players have to pay the high costs of maintaining the appropriate levels of armaments. But this reduces the net value of the status quo for both and makes it less attractive. Quite apart from the security dilemma, we have a problem with stability here because nobody now likes the status quo sufficiently. War becomes likely because deterrence may fail and it may do so for reasons that are not immediately related to the probability of winning a war.

2 Nuclear Deterrence

One big problem with nuclear weapons is that the costs of using them are enormous, especially if the other side also has them. When nuclear weapons are an option, K_i^F is so high that the expected utility of war is very very low for both players. This means, that the threat to use force is no longer credible, which can become a problem for the players. How can they deal with this issue?

Recall that the expected utility of war for the defender is

$$EU_D(\text{War}) = pB_D^L + (1 - p)B_D^V - K_D^F.$$

With nuclear weapons, especially when the other side has second-strike capability, the costs are paid always regardless of who "wins" the war. Now, if K_D^F is enormous, it swamps out any possible benefits from victory and makes war unattractive, thereby destroying the defender's credibility.

One way the defender can deal with this is by using the second type of commitment strategies we discussed: the manipulation of risk. Suppose the Defender threatens with a limited war that has some chance of escalating into a nuclear one. Let's call the probability of an accident or inadvertent escalation is s and let's assume that the large cost is only paid if nuclear war occurs. The expected utility calculation becomes

$$EU_D(\text{War}) = pB_D^L + (1 - p)B_D^V - sK_D^F.$$

For low values of s , the expression sK_D^F is very small also. Thus, the defender can improve its own credibility by using a threat that leaves something to chance. Some compelling strategies have special names:

1. *punishment* raises the costs of continued resistance, and involves exploiting civilian vulnerabilities of the target; manipulating K_C^F .
2. *risk* raises the probability that the target will suffer the costs by gradually increasing pressure and risk to civilians; manipulating s .

3. *denial* decreases the probability that the target will achieve his objectives by resisting, either through conventional war or guerilla warfare; manipulating p .

Increasing s , the probability of inadvertent escalation, undermines the defender's credibility. However, it also undermines the challenger's credibility as well. The challenger makes a very similar calculation about the nukes. The lower the s , the higher its expected utility from fighting.

The autonomous risk of inadvertent escalation cuts both ways because increasing it lowers the expected utilities for both sides and worsens their credibility problems. This is why it becomes a contest of nerves, a competition in risk taking instead of chess. Either way, however, it is still possible to threaten credibly with war when it would have been impossible to do so without the random component.

Maneuvering yourself into a situation from which you cannot retreat as we discussed can help greatly here because it saddles the opponent with the unenviable choice to take the step that increases the risk. In our simple game today this choice was left to the challenger by design. This means that the Defender can do well by increasing s . Although this also lowers his expected utility from fighting, it does so to the challenger's as well. Given that it is the challenger who has to decide between backing down and starting the fight, the defender wins because the challenger would back down.

When you read about the Cuban Missile Crisis, think about those two components because the strategy that the U.S. picked was one that increased the shared risk of disaster and simultaneously saddled the Russians with the escalatory step, succeeding in compelling them to back down.

3 Summary

- Deterrence occurs when a "defender" tries to manipulate the expectations of a "challenger" such that the challenger is deterred from taking an action contrary to the interests of the defender.
- Compellence occurs when an actor manipulates the expectations of another actor such that the latter reverses its course of action that is detrimental to the former.
- The credibility of the defender and the challenger is important in both the deterrence and compellent stages.
- The challenger compares the status quo to expected utility from action in the deterrence game; improving the status quo can be just as effective as military options.
- The challenger compares backing down to expected utility of continuing its actions in the compellence game; improving its payoff from backing down can be just as effective as military options.
- Beliefs are critical; use commitment tactics (constraining choices, manipulating risk).
- An action that improves one's credibility in the compellent stage may undermine it in the deterrent stage and vice versa.