Crisis Bargaining and Mutual Alarm

1 Crisis Bargaining

When deterrence fails (that is, when a demand by a challenger is made), an international crisis begins. During this brief and intense period, actors maneuver for better bargaining positions, communicate threats and promises, and try to avoid costly fighting. While deterrence can be seen as bargaining in a peaceful situation and compellence as bargaining in a fighting situation, crisis bargaining occurs in an environment that is not quite war but is less than peace as well. A crisis is the last chance to avoid the transition from peace to war, it is the last chance to settle the contested issue by diplomatic means or limited military deployments. If a crisis is resolved, we're back to deterrence. If it is not, we enter the period of compellence.

The First World War was preceded by an intense crisis that broke out after the assassination of the Austrian Arch-duke Ferdinand in Sarajevo, Serbia. The Austrians demanded that the Serb government produce the co-conspirators, assist in the investigation, and adopt a more pro-Austrian foreign policy. The Russians, who feared further expansion of Austrian influence on the Balkans came to the aid of the Serbs and threatened to intervene, which provoked the Germans. We shall see later how this crisis unfolded.

Another famous crisis was created by Hitler in 1938 when he demanded that Czechoslovakia cede the Sudetenland to Germany because a German minority lived there and was supposedly mistreated by the Czechs. The Great Powers who had guaranteed Czechoslovakia's borders after the First World War (when the state was created out of the defunct Austro-Hungarian Empire) gathered in Munich to negotiate. War was avoided by Britain and France agreeing to Hitler's demands. The helpless Czechs saw their country dismembered in what gave appeasement a bad name.

2 The Game of Chicken

We begin with a very simple formulation of crisis bargaining, one that has been extensively used to describe the tense dynamics of the situation. This is the so-called Game of Chicken.

The original game goes something like this. Two players in souped-up cars race against each other down a narrow road in front of a cheering crowd. Each can choose to keep going or swerve (we assume that whenever they swerve, they swerve to the right). If they both swerve, they avoid collision and neither gains anything. If one swerves but the other keeps going, the first loses face and is declared a wimp, and the other is declared the tough guy, wins in reputation and the admiration of the girls in the crowd. If neither swerves, they collide and die. Here's the game in matrix form:

	Keep Going	*Swerve*
Keep Going	0, 0	3, 1
Swerve	1, 3	2, 2

Figure 1: The Game of Chicken.

The payoff matrix represents the ordinal ranking of the outcomes for each player. Colliding is worst for both, and each prefers to keep going if the other swerves. Swerving at the same time is better than being a wimp (or dying) but is not as good as being the tough guy.

There are two Nash equilibria in pure strategies, (K, S) and (S, K). If a player believes that the other will keep going, then the only rational response is to swerve, to which the rational response is to keep going. These are mutually best responses, and so they are an equilibrium. The outcome is no collision and a winner, with the identity of the winner being different between the two equilibria.

This is not very helpful because it leaves unresolved the crucial question about expectations. How do players form the "appropriate" expectations that determine which equilibrium they end up in?

The first cut is to finish the equilibrium analysis of the static game. The two pure-strategy Nash equilibria may not be very useful, but what about the one in mixed strategies? Let's find it using our payoff-equating method. Let p be the probability that Player 1 keeps going and q be the analogous probability for Player 2.

$$\begin{aligned} EU_1(K) &= EU_1(S) & \qquad EU_2(K) &= EU_2(S) \\ 0q + 3(1-q) &= 1q + 2(1-q) & \qquad 0p + 3(1-p) &= 1p + 2(1-p) \\ q &= \frac{1}{2} & \qquad p &= \frac{1}{2} \end{aligned}$$

So, in the mixed strategy Nash equilibrium, each player keeps going with probability $1/2$. In this equilibrium collision is no longer a zero-probability event. In fact, there is a 25% chance that the game will end in disaster.

A naïve interpretation of playing mixed strategies is that players use some sort of randomizing device, which tells them which pure strategy to use with the correct probability. In this case, a toss of a fair coin would be such a device. Each player would toss the coin and if heads comes up, keep going, or swerve if tails comes up. This interpretation, of course, is stupid.

However, let's think of the situation as a game in which players are not quite sure about each other's payoffs. That is, there is some uncertainty about the true payoffs which are privately known to each player but about which the other player can only form beliefs. Suppose, for example, that your opponent can vary in toughness from being quite wimpish (and therefore likely to swerve) to being quite foolhardy (and therefore likely to keep going).

Intuitively, you would then adjust your strategy to be optimal in a way that is consistent with your beliefs about the opponent. If you think you are more likely to be facing a wimp, you will be more likely to keep going, and if you think you are more likely to be facing a tough guy, you are more likely to swerve. In any case, what you really are playing is

a pure strategy. That is, you either swerve or keep going. However, which one you pick depends on your beliefs, that is, on the probabilities you assign to the different types of opponents you might be facing. From your perspective the strategy is pure, but from the perspective of your opponent, it will appear randomized because he cannot observe your beliefs. Sometimes you will swerve and sometimes you will keep going, making it look from the outside that you are playing a mixed strategy. But in fact you are playing a rational pure strategy consistent with your beliefs.

It can be shown (and the guy who did that first, John Harsanyi, shared the Nobel prize with John Nash and Reinhard Selten) that the mixed strategy Nash equilibrium of the static game of complete information can be obtained from similar games with incomplete information as the players become more and more confident about the type of opponent they might be facing. In other words, the complete information mixed strategy equilibrium is a useful approximate insight into the dynamics of the complicated games of incomplete information.

Given that, what does this equilibrium tell us? The first important thing it tells us is that there's always a danger of unwanted disaster in a crisis. The probability of disaster is always greater than zero, and can actually be quite high depending on the payoffs. This, of course, is why the situation is called a *crisis* and not, for example, "diplomatic picnic" or something. More to the point, it demonstrates that even *in equilibrium*, that is, when both players follow their optimal strategies, this danger continues to exist.

In our example, the equilibrium probability of collision, as I mentioned before, is $1/4$, as is the probability of each other of the three possible outcomes. The expected value of this game for Player 1 is then

$$\begin{aligned} EU_1(G) &= (0)\Pr(K,K) + (3)\Pr(K,S) + (1)\Pr(S,K) + (2)\Pr(S,S) \\ &= (0)\left(\frac{1}{4}\right) + (3)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{4}\right) + (2)\left(\frac{1}{4}\right) = \frac{3}{2} \end{aligned}$$

The calculation for Player 2 is analogous and yields the same result. Thus, each player expects a payoff of 1.5 from playing the equilibrium mixed strategy in the Game of Chicken. This brings us to the second point, namely, that the expected mixed-strategy equilibrium payoff is strictly worse than getting the payoff of 2 from the (S,S) outcome where both players swerve. If players could only commit credibly to swerving, then collision would not occur and they would both do better than engaging in a dangerous crisis.

But this shows the now familiar problem: neither player can credibly commit to swerving because if he does, then the opponent has an incentive to keep going, which destroys its commitment to swerving that made swerving worthwhile in the first place. Bottom line is: You cannot commit to swerving because your opponent will exploit you. So what can you do?

3 Commitment and Risk Strategies in Crises

You can commit not to swerve. Expectations are crucial and so we're now in the familiar realm of forming and influencing expectations. As usual, there're static and dynamic com-

ponents of establishing commitments, with the attendant coordination and communication problems.

The static component involves constraining your options. You can commit to not changing your course if you relinquish control and leave the final clear chance to prevent disaster to the opponent. In the Game of Chicken you can yank your steering wheel and wave it to the other driver. You have now given up the ability to swerve and so could not do it even if you wanted to. Again, making yourself weak can give you a bargaining advantage.

Clearly, it is important that the other driver actually sees you waving the wheel otherwise your commitment is useless. In other words, your commitment must be communicated and understood. If the other driver makes himself unavailable for receiving communication, you may not be able to improve your bargaining position by constraining your choice. For example, the other driver can purposefully put a blindfold on so he cannot see you should you choose to yank your wheel out. As long as you can see that he cannot see you, he's effectively preempted you by committing credibly to not seeing your commitment, and therefore leaving you with the last clear chance to swerve.

This desire to preempt the other's commitment can be strong making the situation rather tricky. Imagine what would happen if you yank the steering wheel and wave it triumphantly out the window just to see the other driver grinning and waving his to you. Now you have both committed with perfect credibility and, since neither can now swerve, have ensured that the disaster outcome will occur. Again, whenever you commit by constraining *your* choice, you must make sure that the other actor has an option to exit. That is, you must make sure he has not committed irrevocably already. We shall return to this dynamic of preemption in a minute.

The dynamic commitment device is the strategy of risk taking. With each second the risk of collision increases. If you start with some belief about how likely the other player is to be a wimp, with every meter that he keeps going you must revise your belief and increase the probability that he might actually be a tough guy. This is continuous dynamic updating: The longer one stays, the "tougher" he appears to the opponent because he appears more willing to run the risk of disaster.

If you think that going one more meter is likely to convince him that you are tough and cause him to swerve, you would probably go that additional meter. If at the end of this meter you observe that he is still in the race, you revise your beliefs and increase the likelihood that he might be tough. Simultaneously, you know that he has also revised his belief about you, so it is tempting to go another meter, and so on.

You are trying to convince him that you think he will swerve and will therefore keep going. Without a static commitment device like the steering wheel or the blindfold, you use the increasing risk of collision to demonstrate that you believe that he will swerve. If you succeed in convincing him that you expect him to swerve, his rational response would be to swerve.

The danger of this process should be evident. It requires genuine risk of disaster to work because neither player will begin revising expectations in the absence of real danger. The costliness of the possible disastrous outcome and the willingness to run the risk of incurring it sends an informative signal to the other player about your expectations. Generally, a signal has to be costly in order to be informative. Otherwise, it's just cheap

talk.

The real problem with this strategy of costly signaling is that the signals become costlier the closer one gets to disaster, and so they are most effective when the cars begin getting close to the collision point. One may have to delay swerving until it's too late. That is, until a point is reached where the cars are so close that even swerving by both can no longer avoid disaster. By using the costly signaling strategy one may find oneself irrevocably committed to a collision course.

In complicated situations it may be difficult or next to impossible to calculate exactly where this point of no return is. In the end, it is entirely possible that both players know that they are both really tough and although at this point they would both prefer to swerve simultaneously, their risk strategies that were necessary to convince them of their new beliefs have made it impossible for them to avoid disaster.

Suppose that a leader pays domestic costs for backing down in a crisis. The longer he stays in the crisis, the higher these costs become because the more national reputation is at stake. So, even though it may be relatively easy to back down at the beginning of a crisis, it may become increasingly difficult to do so once it gets under way.

Two states with leaders who may suffer these "audience costs" may become locked in an unavoidable war that neither one wanted through a dynamic similar to the one we just discussed. Even though the costliness of staying in the crisis is useful for signaling one's toughness to the opponent in the hope of getting a better bargain, at some point it may become impossible to back down. In a highly uncertain environment, one does not know where this point of no return is, and so there exists a real risk that both players would continue the crisis for too long and fight a war that they would have preferred to avoid if they had known at the outset that it would take so much to signal their resolve. This argument about **audience costs** and their use as a signaling (and therefore, commitment) device is quite prominent in international relations theory.

4 The Dynamics of Mutual Alarm

Mobilization is the process through which a country gears up for war. It involves calling the reservists, arming them, and transporting them to the front lines along with piles of equipment, food, fuel, and support personnel. Mobilization is enormously complicated and every country has carefully prepared plans on how to execute its own. It is also terribly expensive because it involves not only removing men from their jobs but also disrupting commercial schedules of railways and, in more modern times, aviation.

Once mobilization is under way, it is hard to stop, and nearly impossible to restart if stopped. Once completed, it cannot be maintained indefinitely. Once its resources and armies are mobilized, a country must use them or lose them. That is, nobody can afford to field armies without action for a long time. The forces either get used or the soldiers must be sent home.

This momentum implies two things. First, a country is vulnerable if it stops its mobilization midway before it is completed because the resulting chaos makes it next to impossible to restart the process quickly. If it stops then, an adversary could use this opportunity to

strike. Second, once mobilized a country becomes a great menace to its potential adversary because it must either strike or demobilize. This brief window of opportunity makes it hard to negotiate at leisure a way out of the crisis.

Now think about the combination of these two effects. A country that begins mobilization will be extremely dangerous to its adversary once mobilization is completed. However, it is also extremely vulnerable during mobilization and in the event it stops the process. Knowing that it will eventually have to face the fully mobilized resources of this country, an adversary might be tempted to strike sooner, making the crisis even more unstable. (Crisis stability refers to the likelihood that the crisis would end up in war.)

Let's look again at that fateful summer of 1914. Austria-Hungary had issued its ultimatum to Serbia and it looked like it would go to war with the little Balkan state. The Russians faced a dilemma. They had to mobilize to threaten the Austrians sufficiently to prevent them from finishing off the Serbs. A full mobilization, however, would also threaten Germany and perhaps provoke it into mobilizing itself.

The Russians did have plans for partial mobilization in the south, which is exactly what they needed to threaten the Austrians only. However, once started, this partial mobilization could not be converted into full mobilization because of the way the railroads were scheduled. This was a problem because initiating partial mobilization, while not threatening to Germany, would expose the Russians to a German attack. The Russians had to trust the Germans not to exploit this opportunity.

Or they could hedge against it and order full mobilization just in case. But full mobilization is preparation for total war and Germany's reaction was, of course, to mobilize itself. Germany also faced a dilemma. The Russians were allied with the French and if Germany attacked Russia, it would find itself fighting on two fronts when the French, in accordance with their agreements with the Russians, attacked from the West while Germany was engaged in the East. Or, even without the alliance, Germany had reasons to fear that France might use the opportunity and try to regain Alsace and Lorraine which she had lost after the Franco-Prussian War of 1871.

At any rate, there was a real danger that if Germany mobilized and threw all its forces in the east, the French would attack across its exposed western borders. The German high command believed that finishing off the French would be quicker and easier than defeating the Russians, and so in an event of a war with Russia, the German war plans called for a surprise attack on France first. The mobilization plans, just like the ones of the Russians, were also impossible to reverse once put into motion, and so once the Russians ordered full mobilization out of fear that Germany might exploit a partial mobilization, the Germans mobilized for war against France out of fear that the French might exploit their potential vulnerability, and the war was destined to become at least European in scope.

The military doctrine at the time emphasized speed of mobilization and surprise attack. It was believed that the country that could finish its mobilization first and attack its opponent before the latter was ready could gain a significant advantage and perhaps even win the war. This creates an awfully dangerous situation. A statesman who has the military instrument at the ready and knows that he must use it or lose and who further knows that his opponent is in the same position, faces a fateful decision where hesitation to strike first may mean national defeat.

Notice how this provides a motivation for war quite apart from its other causes. This one is mechanical, it is produced by the military technology of coercion and planning. A vulnerable military force provides a temptation to the enemy to strike until this window of vulnerability exists. Therefore, a vulnerable military force cannot afford to wait and must attack first.

If striking first carries such an advantage, the other side may think that you want to do it even if you really do not. But if it thinks you might do it, then it is tempted to do it first even though it may not want to do it. But if you know that it might be tempted in this way, you now think that it might strike, and so you might prefer to strike first because you think that it would do so anyway. Both of you provide each other with justification to strike first. These interacting expectations produce a chain of the now familiar logic: he thinks that I think that he thinks that I think... he thinks that I think he will attack, so he will, so I must.

The end result is war that neither may have wanted, an accidental war that is not due to some mechanical failure but to the expectations that shift in such a way due to the constraints of technology that both sides become convinced that war is inevitable, making it truly inevitable in the process. In a way, because technology commits the players to following certain strategies, they may become victims of circumstance and make the fateful decision to start fighting even though they would rather not.

It is the fear of surprise attack that influences expectations in this way, and this fear is generated by one's own vulnerability and that of its opponent. Especially that of his opponent because what generates the escalating reciprocity of fear is the expectation that because the opponent is vulnerable, he might strike first.

We reach the somewhat paradoxical conclusion that to increase crisis stability one must work to *decrease* the vulnerability of its opponent's military forces. But compelling one's opponent requires destroying a significant portion of these forces, which makes it desirable to *increase* their vulnerability. Herein lies the problem: An action that is designed to reduce the likelihood of war makes it more difficult to win the war should the war occur. Conversely, an action that increases the likelihood of war also makes it easier to win the war. You can see how a prudent state would probably hedge against losing a war and will choose a strategy of the second type, making crises less stable and far more dangerous.

Still, during the Cold War, the two superpowers pursued strategies that decreased the vulnerability of the military forces and increased the vulnerability of the civilian population, thereby providing powerful incentives not to jump the gun in a crisis. Once each side acquired second-strike capability, the era of mutually assured destruction (MAD) began. Each country could absorb a first strike by the enemy and then return a devastating counterblow.

Acquiring this capability involved (a) building a lot more missiles—what some people mistakenly called "overkill" in the belief that once the U.S. had enough nuclears to blow up the Russians it was unnecessary to build more, completely missing the point that the relevant quantity was not the total number of nuclears but the number that could survive a surprise attack by the Russians; and (b) rendering the existing forces invulnerable to enemy bombs. The second strategy involved dispersing of missile sites and bombers, hardening missile silos, and, once it became technologically possible, placing nuclear weapons on hard to detect submarines.

In addition to making their military forces less vulnerable, the two superpowers made their civilian populations more vulnerable when they agreed not to build anti-ballistic missile systems (ABMs). This venerable treaty persisted until George W. Bush unilaterally withdrew the U.S. from it. The purpose, however gruesome, was to supplement the stability-inducing invulnerability of the military. If you have second strike capability and your enemy's cities are vulnerable, then your enemy is unlikely to attack you first by jumping the gun in a crisis. But if your enemy is unlikely to launch a surprise attack, then you have no reason to launch one either, and so crises become much more stable.

5 Summary

- A **crisis** is the last chance to avoid the transition from peace to war. It is a period of time where intense negotiations take place along with preparations for fighting.
- **Crisis stability** refers to the probability that a crisis will end in war. A crisis is stable if the likelihood of war is small, and is unstable if the likelihood is high.
- **Crisis bargaining** refers to the bargaining process through which opponents try to avoid war while securing their demands.
- The **Game of Chicken** is the simplest model of crisis bargaining and although its two pure strategy equilibria are not useful in analyzing strategic behavior, the mixed-strategy equilibrium yields important insights:
  - the mixed strategies are useful approximations of how rational players would behave in the presence of uncertainty;
  - the probability of war in such situations is strictly positive;
  - the expected utility from the crisis is less than the utility from the outcome where both sides back down but the latter outcome cannot occur because neither side can credibly commit not to exploit the other.
- One can use the familiar *constraining choice* and *risk* strategies to increase the credibility of one's commitments in a crisis. A **static** commitment device would remove the option of backing down or would make one unavailable to receive communication from the other. A **dynamic** commitment device would gradually increase the shared risk of disaster to persuade the opponent that one expects it to back down first. Both strategies can inadvertently result in a *lock-in* where players are unable to back down and war is inevitable.
- When there are advantages to striking first, the *reciprocal fear of surprise attack* may generate an additional risk of war quite apart from its other causes.
- Reducing the vulnerability of the opponent's military forces increases crisis stability but decreases the effectiveness of compellence strategies should war occur. Increasing the vulnerability decreases crisis stability but increases likelihood of compellence success.