14

Brinkmanship

The Cuban Missile Crisis

In Chapter 1, we explained that our basic approach was neither pure theory nor pure case study, but a combination in which theoretical ideas were developed by using features of particular cases or examples. Thus, we ignored those aspects of each case that were incidental to the concept being developed. However, after you have learned the theoretical ideas, a richer mode of analysis becomes available to you in which factual details of a particular case are more closely integrated with game-theoretic analysis to achieve a fuller understanding of what has happened and why. Such *theory-based case studies* have begun to appear in diverse fields—business, political science, and economic history.[1]

Here we offer an example from political and military history—namely, nuclear brinkmanship in the Cuban missile crisis of 1962. Our choice is motivated by the sheer drama of the episode, the wealth of factual information that has become available, and the applicability of an important concept from game theory.

The crisis, when the world came as close to an unaccidental nuclear war as it ever has, is indeed often offered as the classic example of brinkmanship. You may think that the risk of nuclear war died with the dissolution of the

[1] Two excellent examples of theory-based studies are Pankaj Ghemawat, *Games Businesses Play: Cases and Models* (Cambridge, Mass.: MIT Press, 1997), and Robert H. Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast, *Analytic Narratives* (Princeton: Princeton University Press, 1998). A broader analysis of the approach can be found in Alexander L. George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences* (Cambridge, Mass.: MIT Press, 2005).

Soviet Union and that therefore our case is a historical curiosity. But nuclear arms races continue in many parts of the world, and such rivals as India and Pakistan or Iran and Israel may find use for the lessons taken from the Cuban crisis. More important for many of you, brinkmanship must be practiced in many more common situations, from political negotiations to business-labor relations to marital disputes. Although the stakes in such games are lower than those in a nuclear confrontation between superpowers, the same principles of strategy apply.

In Chapter 9, we introduced the concept of brinkmanship as a strategic move; here is a quick reminder of that analysis. A *threat* is a response rule, and the threatened action inflicts a cost on both the player making the threat and the player whose action the threat is intended to influence. However, if the threat succeeds in its purpose, this action is not actually carried out. Therefore, there is no apparent upper limit to the cost of the threatened action. But the risk of *errors*—that is, the risk that the threat may fail to achieve its purpose or that the threatened action may occur by accident—forces the strategist to use the minimal threat that achieves its purpose. If a smaller threat is not naturally available, a large threat can be scaled down by making its fulfillment probabilistic. You do something in advance that creates a probability, but not certainty, that the mutually harmful outcome will happen if the opponent defies you. If the need actually arose, you would not take that bad action if you had the full freedom to choose. Therefore, you must arrange in advance to let things get out of your control to some extent. *Brinkmanship* is the creation and deployment of such a probabilistic threat; it consists of a deliberate loss of control.

In our extended case study of the Cuban missile crisis, we will explain the concept of brinkmanship in detail. In the process, we will find that many popular interpretations and analyses of the crisis are simplistic. A deeper analysis reveals brinkmanship to be a subtle and dangerous strategy. It also shows that many detrimental outcomes in business and personal interactions—such as strikes and breakups of relationships—are examples of brinkmanship gone wrong. Therefore, a clear understanding of the strategy, as well as its limitations and risks, is very important to all game players, which includes just about everyone.

1 A BRIEF NARRATIVE OF EVENTS

We begin with a brief story of the unfolding of the crisis. Our account draws on several books, including some that were written with the benefit of documents and statements released since the collapse of the Soviet Union.[2] We cannot hope to do justice to the detail, let alone the drama, of the events. President Kennedy said at the time of the crisis: “This is the week when I earn my salary.” Much more than a president’s salary stood in the balance. We urge you to read the books that tell the story in vivid detail and to talk to any relatives who lived through it to get their firsthand memories.[3]

In late summer and early fall of 1962, the Soviet Union (USSR) started to place medium- and intermediate-range ballistic missiles (MRBMs and IRBMs) in Cuba. The MRBMs had a range of 1,100 miles and could hit Washington, D.C.; the IRBMs, with a range of 2,200 miles, could hit most of the major U.S. cities and military installations. The missile sites were guarded by the latest Soviet SA-2–type surface-to-air missiles (SAMs), which could shoot down U.S. high-altitude U-2 reconnaissance planes. There were also IL-28 bombers and tactical nuclear weapons called Luna by the Soviets and FROG (free rocket over ground) by the United States, which could be used against invading troops.

This was the first time that the Soviets had ever attempted to place their missiles and nuclear weapons outside Soviet territory. Had they been successful, it would have increased their offensive capability against the United States manyfold. It is now believed that the Soviets had fewer than 20, and perhaps as few as “two or three,” operational intercontinental ballistic missiles (ICBMs) in their own country capable of reaching the United States (*War*, 464, 509–510). Their initial placement in Cuba had about 40 MRBMs and IRBMs, which was a substantial increase. But the United States would still have retained vast

[2] Our sources include Robert Smith Thompson, *The Missiles of October* (New York: Simon & Schuster, 1992); James G. Blight and David A. Welch, *On the Brink: Americans and Soviets Reexamine the Cuban Missile Crisis* (New York: Hill and Wang, 1989); Richard Reeves, *President Kennedy: Profile of Power* (New York: Simon & Schuster, 1993); Donald Kagan, *On the Origins of War and the Preservation of Peace* (New York: Doubleday, 1995); Aleksandr Fursenko and Timothy Naftali, *One Hell of a Gamble: The Secret History of the Cuban Missile Crisis* (New York: W. W. Norton & Company, 1997); and last, latest, and most direct, *The Kennedy Tapes: Inside the White House During the Cuban Missile Crisis*, ed. Ernest R. May and Philip D. Zelikow (Cambridge, Mass.: Harvard University Press, 1997). Graham T. Allison’s *Essence of Decision: Explaining the Cuban Missile Crisis* (Boston: Little Brown, 1971) remains important not only for its narrative, but also for its analysis and interpretation. Our view differs from his in some important respects, but we remain in debt to his insights. We follow and extend the ideas in Avinash Dixit and Barry Nalebuff, *Thinking Strategically* (New York: W. W. Norton & Company, 1991), ch. 8.

When we cite these sources to document particular points, we do so in parentheses in the text, in each case using a key word from the title of the book followed by the appropriate page number or page range. The key words have been underlined in the sources given here.

[3] For those of you with no access to firsthand information or those who seek a beginner’s introduction to both the details and the drama of the missile crisis, we recommend the film *Thirteen Days* (2000, New Line Cinema). A new, relatively short book by Sheldon Stern uses the evidence from the Kennedy administration tapes to present as accurate a view of the crisis and its later analysis as possible. His book is perhaps the best short read for interested parties. See Sheldon Stern, *The Cuban Missile Crisis in American Memory: Myths versus Reality* (Stanford, Calif.: Stanford University Press, 2012).

superiority in the nuclear balance between the superpowers. Also, as the Soviets built up their submarine fleet, the relative importance of land-based missiles near the United States would have decreased. But the missiles had more than mere direct military value to the Soviets. Successful placement of missiles so close to the United States would have been an immense boost to Soviet prestige throughout the world, especially in Asia and Africa, where the superpowers were competing for political and military influence. Finally, the Soviets had come to think of Cuba as a "poster child" for socialism. The opportunity to deter a feared U.S. invasion of Cuba and to counter Chinese influence in Cuba weighed importantly in the calculations of the Soviet leader and Premier, Nikita Khrushchev. (See *Gamble*, 182–183, for an analysis of Soviet motives.)

U.S. surveillance of Cuba and of shipping lanes during the late summer and early fall of 1962 had indicated some suspicious activity. When questioned about it by U.S. diplomats, the Soviets denied any intentions to place missiles in Cuba. Later, faced with irrefutable evidence, they said that their intention was defensive, to deter the United States from invading Cuba. It is hard to believe this, although we know that an offensive weapon *can* serve as a defensive deterrent threat.

An American U-2 "spy plane" took photographs over western Cuba on Sunday and Monday, October 14 and 15. When developed and interpreted, they showed unmistakable signs of construction on MRBM launching sites. (Evidence of IRBMs was found later, on October 17.) These photographs were shown to President Kennedy the following day (October 16). He immediately convened an ad hoc group of advisers, which later came to be called the Executive Committee of the National Security Council (ExComm), to discuss the alternatives. At the first meeting (on the morning of October 16), he decided to keep the matter totally secret until he was ready to act, mainly because if the Soviets knew that the Americans knew, they might speed up the installation and deployment of the missiles before the Americans were ready to act, but also because spreading the news without announcing a clear response would create panic in the United States.

Members of ExComm who figured most prominently in the discussions were the Secretary of Defense, Robert McNamara; the National Security Adviser, McGeorge Bundy; the Chairman of the Joint Chiefs of Staff, General Maxwell Taylor; the Secretary of State, Dean Rusk, and Undersecretary George Ball; the Attorney General, Robert Kennedy (who was also the President's brother); the Secretary of the Treasury, Douglas Dillon (also the only Republican in the Cabinet); and Llewellyn Thompson, who had recently returned from being U.S. Ambassador in Moscow. During the two weeks that followed, they would be joined by or would consult with several others, including the U.S. Ambassador to the United Nations, Adlai Stevenson; the former Secretary of State and a senior statesman of U.S. foreign policy, Dean Acheson; and the Chief of the U.S. Air Force, General Curtis LeMay.

In the rest of that week (October 16 through 21), the ExComm met numerous times. To preserve secrecy, the President continued his normal schedule, including travel to speak for Democratic candidates in the midterm congressional elections that were to be held in November 1962. He kept in constant touch with ExComm. He dodged press questions abut Cuba and persuaded one or two trusted media owners or editors to preserve the facade of business as usual. ExComm's own attempts to preserve secrecy in Washington sometimes verged on the comic, as when almost a dozen of them had to pile into one limo, because the sight of several government cars going from the White House to the State Department in a convoy could cause speculation in the media.

Different members of ExComm had widely differing assessments of the situation and supported different actions. The military Chiefs of Staff thought that the missile placement changed the balance of military power substantially; Defense Secretary McNamara thought it changed "not at all" but regarded the problem as politically important nonetheless (*Tapes*, 89). President Kennedy pointed out that the first placement, if ignored by the United States, could grow into something much bigger and that the Soviets could use the threat of missiles so close to the United States to try to force the withdrawal of the U.S., British, and French presence in West Berlin. Kennedy was also aware that it was a part of the *geopolitical* struggle between the United States and the Soviet Union (*Tapes*, 92).

It now appears that he was very much on the mark in this assessment. The Soviets planned to expand their presence in Cuba into a major military base (*Tapes*, 677). They expected to complete the missile placement by mid-November. Khrushchev had planned to sign a treaty with Castro in late November, then travel to New York to address the United Nations and issue an ultimatum for a settlement of the Berlin issue (*Tapes*, 679; *Gamble*, 182), using the missiles in Cuba as a threat for this purpose. Khrushchev thought Kennedy would accept the missile placement as a *fait accompli*. Khrushchev appears to have made these plans on his own. Some of his top advisers privately thought them too adventurous, but the top governmental decision-making body of the Soviet Union, the Presidium, supported him, although its response was largely a rubber stamp (*Gamble*, 180). Castro was at first reluctant to accept the missiles, fearing that they would trigger a U.S. invasion (*Tapes*, 676–678), but in the end he, too, accepted them. The prospect gave him great confidence and lent some swagger to his statements about the United States (*Gamble*, 186–187, 229–230).

In all ExComm meetings up to and including the one on the morning of Thursday, October 18, everyone appears to have assumed that the U.S. response would be purely military. The only options that they discussed seriously during this time were (1) an air strike directed exclusively at the missile sites and (probably) the SAM sites nearby, (2) a wider air strike including Soviet and Cuban aircraft parked at airfields, and (3) a full-scale invasion of Cuba. If anything, attitudes hardened when the evidence of the presence of the longer-range IRBMs arrived. In fact, at the Thursday meeting, Kennedy discussed a timetable for air strikes to commence that weekend (*Tapes*, 148).

McNamara had first mentioned a blockade toward the end of the meeting on Tuesday, October 16, and developed the idea (in a form uncannily close to the course of action actually taken) in a small group after the formal meeting had ended (*Tapes*, 86, 113). Ball argued that an air strike without warning would be a "Pearl Harbor" and that the United States should not do it (*Tapes*, 115); he was most importantly supported by Robert Kennedy (*Tapes*, 149). The civilian members of ExComm further shifted toward the blockade option when they found that what the military Joint Chiefs of Staff wanted was a massive air strike; the military regarded a limited strike aimed at only the missile sites so dangerous and ineffective that "they would prefer taking no military action than to take that limited strike" (*Tapes*, 97).

Between October 18 and Saturday, October 20, the majority opinion within ExComm gradually coalesced around the idea of starting with a blockade, simultaneously issuing an ultimatum with a short deadline (from 48 to 72 hours was mentioned), and proceeding to military action if necessary after this deadline expired. International law required a declaration of war to set up a blockade, but this problem was ingeniously resolved by proposing to call it a "naval quarantine" of Cuba (*Tapes*, 190–196).

Some people held the same positions throughout these discussions (from October 16 through 21)—for example, the military Chiefs of Staff constantly favored a major air strike—but others shifted their views, at times dramatically. Bundy initially favored doing nothing (*Tapes*, 172) and then switched toward a preemptive surprise air attack (*Tapes*, 189). President Kennedy's own positions also shifted away from an air strike toward a blockade. He wanted the U.S. response to be firm. Although his reasons undoubtedly were mainly military and geopolitical, as a good domestic politician he was also fully aware that a weak response would hurt the Democratic party in the imminent congressional elections. In contrast, the responsibility of starting an action that might lead to nuclear war weighed very heavily on him. He was impressed by the CIA's assessment that some of the missiles were already operational, which increased the risk that any air strike or invasion could lead to the Soviets' firing these missiles and to large U.S. civilian casualties (*Gamble*, 235). In the second week of the crisis (October 22 through 28), his decisions seemed constantly to favor the lowest-key options discussed by ExComm.

By the end of the first week's discussions, the choice lay between a blockade and an air strike, two position papers were prepared, and in a straw vote on October 20 the blockade won 11 to 6 (*War*, 516). Kennedy made the decision to start by imposing a blockade and announced it in a television address to the nation on Monday, October 22. He demanded a halt to the shipment of Soviet missiles to Cuba and a prompt withdrawal of those already there.

Kennedy's speech brought the whole drama and tension into the public arena. The United Nations held several dramatic but unproductive debates. Other world leaders and the usual busybodies of international affairs offered advice and mediation.

Between October 23 and October 25, the Soviets at first tried bluster and denial; Khrushchev called the blockade "banditry, a folly of international imperialism" and said that his ships would ignore it. The Soviets, in the United Nations and elsewhere, claimed that their intentions were purely defensive and issued statements of defiance. In secret, they explored ways to end the crisis. This exploration included some direct messages from Khrushchev to Kennedy. It also included some very indirect and lower-level approaches by the Soviets. In fact, as early as Monday, October 22—before Kennedy's TV address—the Soviet Presidium had decided not to let this crisis lead to war. By Thursday, October 25, they had decided that they were willing to withdraw from Cuba in exchange for a promise by the United States not to invade Cuba, but they had also agreed to "look around" for better deals (*Gamble*, 241, 259). The United States did not know any of the Soviet thinking about this.

In public as well as in private communications, the USSR broached the possibility of a deal concerning the withdrawal of U.S. missiles from Turkey and of Soviet ones from Cuba. This possibility had already been discussed by ExComm. The missiles in Turkey were obsolete; so the United States wanted to remove them anyway and replace them with a Polaris submarine stationed in the Mediterranean Sea. But it was thought that the Turks would regard the presence of U.S. missiles as a matter of prestige and so it might be difficult to persuade them to accept the change. (The Turks might also correctly regard missiles, fixed on Turkish soil, as a firmer signal of the U.S. commitment to Turkey's defense than an offshore submarine, which could move away on short notice; see *Tapes*, 568.)

The blockade went into effect on Wednesday, October 24. Despite their public bluster, the Soviets were cautious in testing it. Apparently, they were surprised that the United States had discovered the missiles in Cuba before the whole installation program was completed; Soviet personnel in Cuba had observed the U-2 overflights but had not reported them to Moscow (*Tapes*, 681). The Soviet Presidium ordered the ships carrying the most sensitive materials (actually the IRBM missiles) to stop or turn around. But it also ordered General Issa Pliyev, the commander of the Soviet troops in Cuba, to get his troops combat-ready and to use all means except nuclear weapons to meet any attack (*Tapes*, 682). In fact, the Presidium twice prepared (then canceled without sending) orders authorizing him to use tactical nuclear weapons in the event of a U.S. invasion (*Gamble*, 242–243, 272, 276). The U.S. side saw only that several

Soviet ships (which were actually carrying oil and other nonmilitary cargo) continued to sail toward the blockade zone. The U.S. Navy showed some moderation in its enforcement of the blockade. A tanker was allowed to pass without being boarded; another tramp steamer carrying industrial cargo was boarded but allowed to proceed after only a cursory inspection. But tension was mounting, and neither side's actions were as cautious as the top-level politicians on both sides would have liked.

On the morning of Friday, October 26, Khrushchev sent Kennedy a conciliatory private letter offering to withdraw the missiles in exchange for a U.S. promise not to invade Cuba. But later that day he toughened his stance. It seems that he was emboldened by two items of evidence. First, the U.S. Navy was not being excessively aggressive in enforcing the blockade. It had let through some obviously civilian freighters; they boarded only one ship, the *Marucla*, and let it pass after a cursory inspection. Second, some dovish statements had appeared in U.S. newspapers. Most notable among them was an article by the influential and well-connected syndicated columnist Walter Lippman, who suggested the swap whereby the United States would withdraw its missiles in Turkey in exchange for the USSR's withdrawing its missiles in Cuba (*Gamble*, 275). Khrushchev sent another letter to Kennedy on Saturday, October 26, offering this swap, and this time he made the letter public. The new letter was presumably a part of the Presidium's strategy of "looking around" for the best deal. Members of ExComm concluded that the first letter was Khrushchev's own thoughts but that the second was written under pressure from hard-liners in the Presidium—or was even evidence that Khrushchev was no longer in control (*Tapes*, 498, 512–513). In fact, both of Khrushchev's letters were discussed and approved by the Presidium (*Gamble*, 263, 275).

ExComm continued to meet, and opinions within it hardened. One reason was the growing feeling that the blockade by itself would not work. Kennedy's television speech had imposed no firm deadline, and as we know, in the absence of a deadline a compellent threat is vulnerable to the opponent's procrastination. Kennedy had seen this quite clearly and as early as Monday, October 22, in the morning ExComm meeting preceding his speech, he commented, "I don't think we're gonna be better off if they're just sitting there" (*Tapes*, 216). But a hard, short deadline was presumably thought to be too rigid. By Thursday, others in ExComm were realizing the problem; for example, Bundy said, "A plateau here is the most dangerous thing" (*Tapes*, 423). The hardening of the Soviet position, as shown by the public "Saturday letter" that followed the conciliatory private "Friday letter," was another concern. More ominously, that Friday, U.S. surveillance had discovered that there were tactical nuclear weapons (FROGs) in Cuba (*Tapes*, 475). This discovery showed the Soviet presence there to be vastly greater than thought before, but it also made invasion more dangerous to U.S. troops. Also on Saturday, a U.S. U-2 plane was shot down over Cuba.

(It now appears that this was done by the local commander, who interpreted his orders more broadly than Moscow had intended [*War*, 537; *Tapes*, 682].) In addition, Cuban antiaircraft defenses fired at low-level U.S. reconnaissance planes. The grim mood in ExComm throughout that Saturday was well encapsulated by Dillon: "We haven't got but one more day" (*Tapes*, 534).

On Saturday, plans leading to escalation were being put in place. An air strike was planned for the following Monday, or Tuesday at the latest, and Air Force reserves were called up (*Tapes*, 612–613). Invasion was seen as the inevitable culmination of events (*Tapes*, 537–538). A tough private letter to Khrushchev from President Kennedy was drafted and was handed over by Robert Kennedy to the Soviet Ambassador in Washington, Anatoly Dobrynin. In it, Kennedy made the following offer: (1) The Soviet Union withdraws its missiles and IL-28 bombers from Cuba with adequate verification (and ships no new ones). (2) The United States promises not to invade Cuba. (3) The U.S. missiles in Turkey will be removed after a few months, but this offer is void if the Soviets mention it in public or link it to the Cuban deal. An answer was required within 12 to 24 hours; otherwise "there would be drastic consequences" (*Tapes*, 605–607).

On the morning of Sunday, October 28, just as prayers and sermons for peace were being offered in many churches in the United States, Soviet radio broadcast the text of a letter that Khrushchev was sending to Kennedy, in which he announced that construction of the missile sites was being halted immediately and that the missiles already installed would be dismantled and shipped back to the Soviet Union. Kennedy immediately sent a reply welcoming this decision, which was broadcast to Moscow by the Voice of America radio. It now appears that Khrushchev's decision to back down was made before he received Kennedy's letter through Dobrynin but that the letter only reinforced it (*Tapes*, 689).

That did not quite end the crisis. The U.S. Joint Chiefs of Staff remained skeptical of the Soviets and wanted to go ahead with their air strike (*Tapes*, 635). In fact, the construction activity at the Cuban missile sites continued for a few days. Verification by the United Nations proved problematic. The Soviets tried to make the Turkey part of the deal semipublic. They also tried to keep the IL-28 bombers in Cuba out of the withdrawal. Not until November 20 was the deal finally clinched and the withdrawal begun (*Tapes*, 663–665; *Gamble*, 298–310).

2 A SIMPLE GAME-THEORETIC EXPLANATION

At first sight, the game-theoretic aspect of the crisis looks very simple. The United States wanted the Soviet Union to withdraw its missiles from Cuba; thus the U.S. objective was to achieve compellence. For this purpose, the United States deployed a threat: Soviet failure to comply would eventually lead to a nuclear war between the superpowers. The blockade was a starting point of this inevitable process and an action that demonstrated the credibility of U.S. resolve. In other words, Kennedy took Khrushchev to the brink of disaster. This was sufficiently frightening to Khrushchev that he complied. The prospect of nuclear annihilation was equally frightening to Kennedy, but that is in the nature of a threat. All that is needed is that the threat be sufficiently costly to the other side to induce it to act in accordance with our wishes; then we don't have to carry out the bad action anyway.

A somewhat more formal statement of this argument proceeds by drawing a game tree like that shown in Figure 14.1. The Soviets have installed the missiles, and now the United States has the first move. It chooses between doing nothing and issuing a threat. If the United States does nothing, this is a major military and political achievement for the Soviets; so we score the payoffs as −2 for the United States and 2 for the Soviets. If the United States issues its threat, the Soviets get to move, and they can either withdraw or defy. Withdrawal is a humiliation (a substantial minus) for the Soviets and a reaffirmation of U.S. military superiority (a small plus); so we score it 1 for the United States and −4 for the Soviets. If the Soviets defy the U.S. threat, there will be a nuclear war. This outcome is terrible for both, but particularly bad for the United States, which as a democracy cares more for its citizens; so we score this −10 for the United States and −8 for the Soviets. This quantification is very rough guesswork, but the conclusions do not depend on the precise numbers that we have chosen. If you disagree with our choice, you can substitute other numbers you think to be a more accurate representation; as long as the *relative* ranking of the outcomes is the same, you will get the same subgame-perfect equilibrium.

Now we can easily find the subgame-perfect equilibrium. If faced with the U.S. threat, the Soviets get −4 from withdrawal and −8 by defiance; so they prefer to withdraw. Looking ahead to this outcome, the United States reckons on

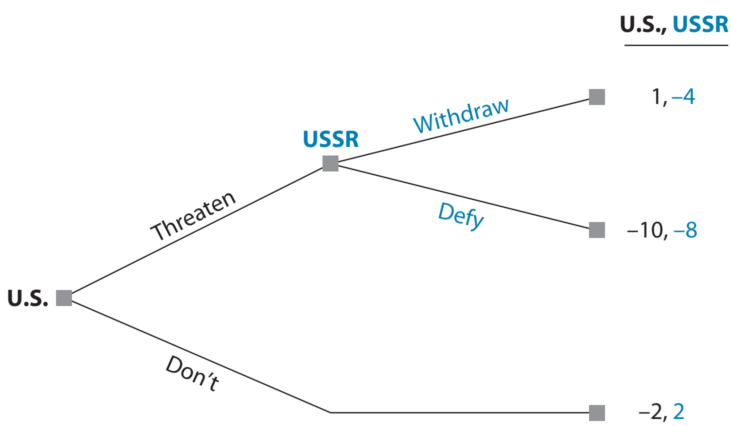

FIGURE 14.1 The Simple-Threat Model of the Crisis

getting 1 if it issues the threat and −2 if it does not; therefore it is optimal for the United States to make the threat. The outcome gives payoffs of 1 to the United States and −4 to the Soviets.

But a moment's further thought shows this interpretation to be unsatisfactory. One might start by asking why the Soviets would deploy the missiles in Cuba at all, when they could look ahead to this unfolding of the subsequent game in which they would come out the losers. But more important, several facts about the situation and several events in the course of its unfolding do not fit into this picture of a simple threat.

Before explaining the shortcomings of this analysis and developing a better explanation, however, we digress to an interesting episode in the crisis that sheds light on the requirements of successful compellence. As pointed out in Chapter 9, a compellent threat must have a deadline; otherwise the opponent can nullify it by procrastination. The discussion of the crisis at the U.N. Security Council on Tuesday, October 23, featured a confrontation between U.S. Ambassador Adlai Stevenson and Soviet Ambassador Valerian Zorin. Stevenson asked Zorin point-blank whether the USSR had placed and was placing nuclear missiles in Cuba. "Yes or no—don't wait for the translation—yes or no?" he insisted. Zorin replied: "I am not in an American courtroom. . . . You will have your answer in due course," to which Stevenson retorted, "I am prepared to wait for my answer until hell freezes over." This was dramatic debating; Kennedy, watching the session on live television, remarked, "Terrific. I never knew Adlai had it in him" (*Profile*, 406). But it was terrible strategy. Nothing would have suited the Soviets better than to keep the Americans "waiting for their answer" while they went on completing the missile sites. "Until hell freezes over" is an unsuitable deadline for compellence.

3 ACCOUNTING FOR ADDITIONAL COMPLEXITIES

Let us return to developing a more satisfactory game-theoretic argument. As we pointed out before, the idea that a threat has only a lower limit on its size—namely, that it be large enough to frighten the opponent—is correct only if the threatener can be absolutely sure that everything will go as planned. But almost all games have some element of uncertainty. You cannot know your opponent's value system for sure, and you cannot be completely sure that the players' intended actions will be accurately implemented. Therefore, a threat carries a twofold risk. Your opponent may defy it, requiring you to carry out the costly threatened action; or your opponent may comply, but the threatened action may occur by mistake anyway. When such risks exist, the cost of threatened action to oneself becomes an important consideration.

The Cuban missile crisis was replete with such uncertainties. Neither side could be sure of the other's payoffs—that is, of how seriously the other regarded the relative costs of war and of losing prestige in the world. Also, the choices of "blockade" and "air strike" were much more complex than the simple phrases suggest, and there were many weak links and random effects between an order in Washington or Moscow and its implementation in the Atlantic Ocean or in Cuba.

Graham Allison's excellent book *Essence of Decision* brings out all of these complexities and uncertainties. They led him to conclude that the Cuban missile crisis cannot be explained in game-theoretic terms. He considers two alternatives: one explanation based on the fact that bureaucracies have their set rules and procedures; another based on the internal politics of U.S. and Soviet governance and military apparatuses. He concludes that the political explanation is best.

We broadly agree but interpret the Cuban missile crisis differently. It is not the case that game theory is inadequate for understanding and explaining the crisis; rather, the crisis was *not a two-person game*—United States versus USSR, or Kennedy versus Khrushchev. Each of these two "sides" was itself a complex coalition of players with differing objectives, information, actions, and means of communication. The players within each side were engaged in other games, and some members were also directly interacting with their counterparts on the other side. In other words, the crisis can be seen as a complex many-person game with alignments into two broad coalitions. Kennedy and Khrushchev can be regarded as the top-level players in this game, but each was subject to constraints of having to deal with others in his own coalition with divergent views and information, and neither had full control over the actions of these others. We argue that this more subtle game-theoretic perspective is not only a good way to look at the crisis, but also essential in understanding how to practice brinkmanship. We begin with some items of evidence that Allison emphasizes, as well as others that emerge from other writings.

First, there are several indications of divisions of opinion on each side. On the U.S. side, as already noted, there were wide differences within ExComm. In addition, Kennedy found it necessary to consult others such as former President Eisenhower and leading members of Congress. Some of them had very different views; for example, Senator William Fulbright said in a private meeting that the blockade "seems to me the worst alternative" (*Tapes*, 271). The media and the political opposition would not give the President unquestioning support for too long either. Kennedy could not have continued on a moderate course if the opinion among his advisers and the public became decisively hawkish.

Individual people also *shifted* positions in the course of the two weeks. For example, McNamara was at first quite dovish, arguing that the missiles in Cuba were not a significant increase in the Soviet threat (*Tapes*, 89) and favoring blockade and negotiations (*Tapes*, 191), but ended up more hawkish, claiming that Khrushchev's conciliatory letter of Friday, October 26, was "full of holes" (*Tapes*, 495, 585) and urging an invasion (*Tapes*, 537). Most important, the U.S. military chiefs always advocated a far more aggressive response. Even after the crisis was over and everyone thought the United States had won a major round in the cold war, Air Force General Curtis LeMay remained dissatisfied and wanted action: "We lost! We ought to just go in there today and knock 'em off," he said (*Essence*, 206; *Profile*, 425).

Even though Khrushchev was the dictator of the Soviet Union, he was not in full control of the situation. Differences of opinion on the Soviet side are less well documented, but, for what it is worth, later memoirists have claimed that Khrushchev made the decision to install the missiles in Cuba almost unilaterally, and, when he informed the members of the Presidium, they thought it a reckless gamble (*Tapes*, 674; *Gamble*, 180). There were limits to how far he could count on the Presidium to rubber-stamp his decisions. Indeed, two years later, the disastrous Cuban adventure was one of the main charges leveled against Khrushchev when the Presidium dismissed him (*Gamble*, 353–355). It has also been claimed that Khrushchev wanted to defy the U.S. blockade, and only the insistence of First Deputy Premier Anastas Mikoyan led to the cautious response (*War*, 521). Finally, on Saturday, October 27, Castro ordered his antiaircraft forces to fire on all U.S. planes overflying Cuba and refused the Soviet ambassador's request to rescind the order (*War*, 544).

Various parties on the U.S. side had very different information and a very different understanding of the situation, and at times this led to actions that were inconsistent with the intentions of the leadership or even against their explicit orders. The concept of an "air strike" to destroy the missiles is a good example. The nonmilitary people in ExComm thought this would be very narrowly targeted and would not cause significant Cuban or Soviet casualties, but the Air Force intended a much broader attack. Luckily, this difference came out in the open early, leading ExComm to decide against an air strike and the President to turn down an appeal by the Air Force (*Essence*, 123, 209). As for the blockade, the U.S. Navy had set procedures for this action. The political leadership wanted a different and softer process: form the ring closer to Cuba to give the Soviets more time to reconsider, allow the obviously nonmilitary cargo ships to pass unchallenged, and cripple but not sink the ships that defy challenge. Despite McNamara's explicit instructions, however, the Navy mostly followed its standard procedures (*Essence*, 130–132). The U.S. Air Force created even greater dangers. A U-2 plane drifted "accidentally" into Soviet air space and almost caused a serious setback. General Curtis LeMay, acting without the President's knowledge or authorization, ordered the Strategic Air Command's nuclear bombers to fly past their "turnaround" points and some distance toward Soviet air space to positions where they would be detected by Soviet radar. Fortunately, the Soviets responded calmly; Khrushchev merely protested to Kennedy.[4]

There was similar lack of information and communication, as well as weakness of the chain of command and control, on the Soviet side. For example, the construction of the missiles was left to the standard bureaucratic procedures. The Soviets, used to construction of ICBM sites in their own country where they did not face significant risk of air attack, laid out the sites in Cuba in a similar way, where they would have been much more vulnerable. At the height of the crisis, when the Soviet SA-2 troops saw an overflying U.S. U-2 plane on Friday, October 26, Pliyev was temporarily away from his desk and his deputy gave the order to shoot it down; this incident created far more risk than Moscow would have wished (*Gamble*, 277–288). And at numerous other points—for example, when the U.S. Navy was trying to get the freighter *Marucla* to stop and be boarded—the people involved might have set off an incident with alarming consequences by taking some action in fear of the immediate situation. Even more dramatically, it was revealed that a Soviet submarine crew, warned to surface when approaching the quarantine line on October 27, did consider firing a nuclear-tipped torpedo that it carried onboard (unknown to the U.S. Navy). The firing-authorization rule required the approval of three officers, only two of whom agreed; the third officer himself may have prevented all-out nuclear war.[5]

All these factors made the outcome of any decision by the top-level commander on each side somewhat *unpredictable*. This gave rise to a substantial risk of the "threat going wrong." In fact, Kennedy thought that the chances of the blockade leading to war were "between one out of three and even" (*Essence*, 1).

As we pointed out, such uncertainty can make a simple threat too large to be acceptable to the threatener. We will take one particular form of the uncertainty—namely, U.S. lack of knowledge of the Soviets' true motives—and analyze its effect formally, but similar conclusions hold for all other forms of uncertainty.

Reconsider the game shown in Figure 14.1. Suppose the Soviet payoffs from withdrawal and defiance are the opposite of what they were before: -8 for withdrawal and -4 for defiance. In this alternative scenario, the Soviets are hard-liners. They prefer nuclear annihilation to the prospect of a humiliating withdrawal and

[4] Richard Rhodes, *Dark Sun: The Making of the Hydrogen Bomb* (New York: Simon & Schuster, 1995), pp. 573–75. LeMay, renowned for his extreme views and his constant chewing of large unlit cigars, is supposed to be the original inspiration, in the 1963 movie *Dr. Strangelove*, for General Jack D. Ripper, who orders his bomber wing to launch an unprovoked attack on the Soviet Union.

[5] This story became public in a conference held in Havana, Cuba, in October 2002, to mark the 40th anniversary of the missile crisis. See Kevin Sullivan, "40 Years After Missile Crisis, Players Swap Stories in Cuba," *Washington Post*, October 13, 2002, p. A28. Vadim Orlov, who was a member of the Soviet submarine crew, identified the officer who refused to fire the torpedo as Vasili Arkhipov, who died in 1999.

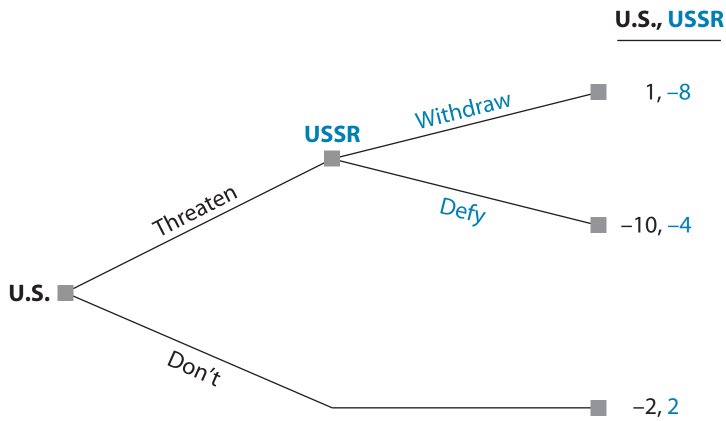

FIGURE 14.2 The Game with Hard-Line Soviets

the prospect of living in a world dominated by the capitalist United States; their slogan is "Better dead than red-white-and-blue." We show the game tree for this case in Figure 14.2. Now, if the United States makes the threat, the Soviets defy it. So the United States stands to get −10 from the threat but only −2 if it makes no threat and accepts the presence of the missiles in Cuba. It takes the lesser of the two evils. In the subgame-perfect equilibrium of this version of the game, the Soviets "win" and the U.S. threat does not work.

In reality, when the United States makes its move, it does not know whether the Soviets are hard-liners, as in Figure 14.2, or softer, as in Figure 14.1. The United States can try to estimate the probabilities of the two scenarios, for example, by studying past Soviet actions and reactions in different situations. We can regard Kennedy's statement that the probability of the blockade leading to war was between one-third and one-half as his estimate of the probability that the Soviets are hard-line. Because the estimate is imprecise over a range, we work with a general symbol, p, for the probability, and examine the consequences of different values of p.

The tree for this more complex game is shown in Figure 14.3. The game starts with an outside force (here labeled "Nature") determining the Soviets' type. Along the upper branch of Nature's choice, the Soviets are hard-line. This leads to the upper node, where the United States makes its decision whether to issue its threat, and the rest of the tree is exactly like the game in Figure 14.2. Along the lower branch of Nature's choice, the Soviets are soft. This leads to the lower node, where the United States makes its decision whether to issue its threat, and the rest of the tree is exactly like the game in Figure 14.1. But the United States does not know from which node it is making its choice. Therefore, the two U.S. nodes are enclosed in an "information set." Its significance is that the United States cannot take different actions at the nodes within the set, such as issuing the threat only if the Soviets are soft. It must take the same action at

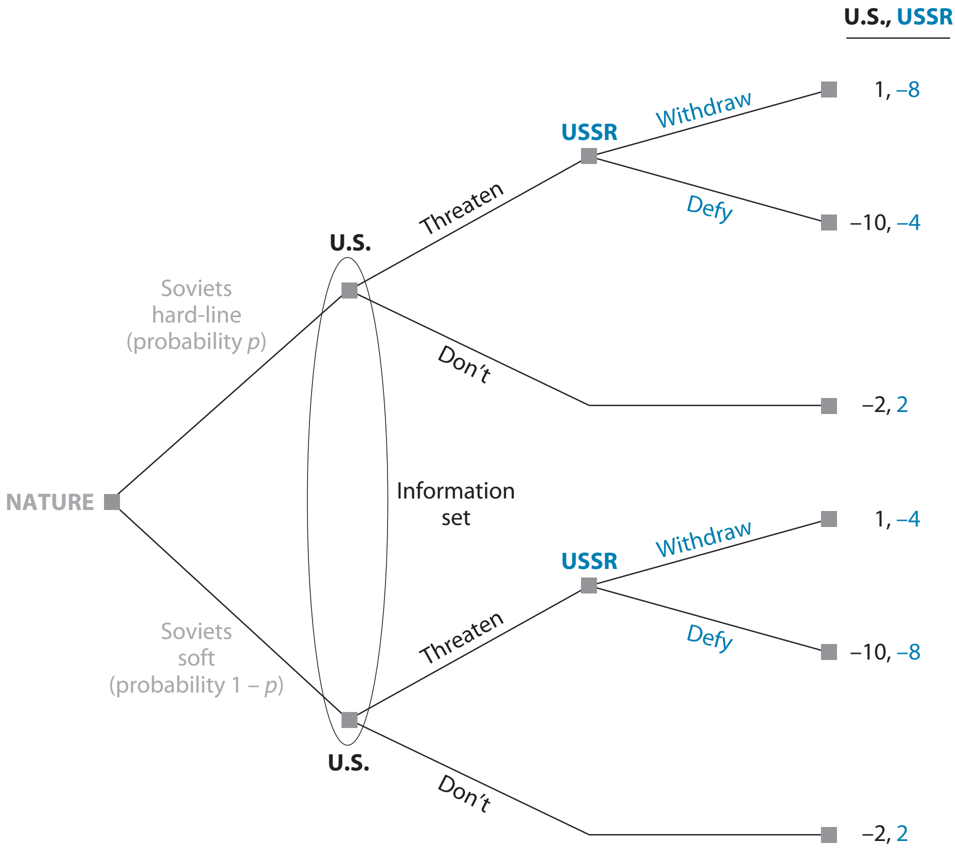

FIGURE 14.3 The Threat with Unknown Soviet Payoffs

both nodes, either threatening at both nodes or not threatening at both. It must make this decision in the light of the probabilities that the game might in truth be "located" at the one node or the other—that is, by calculating the *expected* payoffs of the two actions.

The Soviets themselves know what type they are. So we can do some rollback near the end of the game. Along the upper path, the hard-line Soviets will defy a U.S. threat, and along the lower path, the soft Soviets will withdraw in the face of the threat. Therefore, the United States can look ahead and calculate that a threat will yield a −10 if the game is actually moving along the upper path (a probability of p) and a 1 if it is moving along the lower path (a probability of $1 - p$). The expected U.S. payoff from making the threat is therefore $-10p + (1 - p) = 1 - 11p$.

If the United States does not make the threat, it gets a −2 along either path; so its expected payoff is also −2. Comparing the expected payoffs of the two actions, we see that the United States should make the threat if $1 - 11p > -2$, or $11p < 3$, or $p < 3/11 = 0.27$.

If the threat were sure to work, the United States would not care how bad its payoff could be if the Soviets defied it, whether −10 or even far more negative. But the risk that the Soviets might be hard-liners and thus defy a threat makes the −10 relevant in the U.S. calculations. Only if the probability, p, of the Soviets' being hard-line is small enough will the United States find it acceptable to make the threat. Thus, the upper limit of 3/11 on p is also the upper limit of this U.S. tolerance, given the specific numbers that we have chosen. If we choose different numbers, we will get a different upper limit; for example, if we rate a nuclear war as −100 for the United States, then the upper limit on p will be only 3/101. But the idea of a large threat being "too large to make" if the probability of its going wrong is above a critical limit holds in general.

In this instance, Kennedy's estimate was that p lay somewhere in the range from 1/3 to 1/2. The lower end of this range, 0.33, is unfortunately just above our upper limit 0.27 for the risk that the United States is willing to tolerate. Therefore, the simple bald threat "if you defy us, there will be nuclear war" is too large, too risky, and too costly for the United States to make.

4 A PROBABILISTIC THREAT

If an outright threat of war is too large to be tolerable and if you cannot find another, naturally smaller threat, then you can reduce the threat by creating merely a probability rather than a certainty that the dire consequences for the other side will occur if it does not comply. However, this does not mean that you decide after the fact whether to take the drastic action. If you had that freedom, you would choose to avoid the terrible consequences, and your opponents would know or assume this, so the threat would not be credible in the first place. You must relinquish some freedom of action and make a credible commitment. In this case, you must commit to a probabilistic device.

When making a *simple threat*, one player says to the other player: "If you don't comply, something will *surely* happen that will be very bad for you. By the way, it will also be bad for me, but my threat is credible because of my reputation [or through delegation or other reasons]." With a **probabilistic threat**, one player says to the other, "If you don't comply, there is a *risk* that something very bad for you will happen. By the way, it will also be very bad for me, but later I will be powerless to reduce that risk."

Metaphorically, a probabilistic threat of war is a kind of Russian roulette (an appropriate name in this context). You load a bullet into one chamber of a revolver and spin the barrel. The bullet acts as a "detonator" of the mutually costly war. When you pull the trigger, you do not know whether the chamber in the firing path is loaded. If it is, you may wish you had not pulled the trigger, but by then it will be too late. Before the fact, you would not pull the trigger if you knew that the bullet was in that chamber (that is, if the certainty of the dire action was too costly), but you are willing to pull the trigger knowing that there is only a 1 in 6 chance—in which the threat has been reduced by a factor of 6, to a point where it is now tolerable.

Brinkmanship is the creation and control of a suitable risk of this kind. It requires two apparently inconsistent things. On the one hand, you must let matters get enough out of your control that you will not have full freedom after the fact to refrain from taking the dire action, and so your threat will remain credible. On the other hand, you must retain sufficient control to keep the risk of the action from becoming too large and your threat too costly. Such "controlled lack of control" looks difficult to achieve, and it is. We will consider in Section 5 how the trick can be performed. Just one hint: all the complex differences of judgment, the dispersal of information, and the difficulties of enforcing orders, which made a simple threat too risky, are exactly the forces that make it possible to create a risk of war and therefore make brinkmanship credible. The real difficulty is not how to lose control, but how to do so in a controlled way.

We first focus on the mechanics of brinkmanship. For this purpose, we slightly alter the game of Figure 14.3 to get Figure 14.4. Here, we introduce a different kind of U.S. threat. It consists of choosing and fixing a probability, q, such that if the Soviets defy the United States, war will occur with that probability. With the remaining probability, $(1 - q)$, the United States will give up and agree to accept the Soviet missiles in Cuba. Remember that if the game gets to the point where the Soviets defy the United States, the latter does not have a choice in the matter. The Russian-roulette revolver has been set for the probability, q, and chance determines whether the firing pin hits a loaded chamber (that is, whether nuclear war actually happens).

Thus, nobody knows the precise outcome and payoffs that will result if the Soviets defy this brinkmanship threat, but they know the probability, q, and can calculate expected values. For the United States, the outcome is -10 with the probability q and -2 with the probability $(1 - q)$, so the expected value is

$$-10q - 2(1 - q) = -2 - 8q.$$

For the Soviets, the expected payoff depends on whether they are hard-line or soft (and only they know their own type). If hard-line, they get a -4 from war, which happens with probability q, and a 2 if the United States gives up, which happens with the probability $(1 - q)$. The Soviets' expected payoff is $-4q + 2(1 - q) = 2 - 6q$. If they were to withdraw, they would get a -8, which is clearly worse no matter what value q takes between 0 and 1. Thus, the hard-line Soviets will defy the brinkmanship threat.

The calculation is different if the Soviets are soft. Reasoning as before, we see that they get the expected payoff $-8q + 2(1 - q) = 2 - 10q$ from defiance

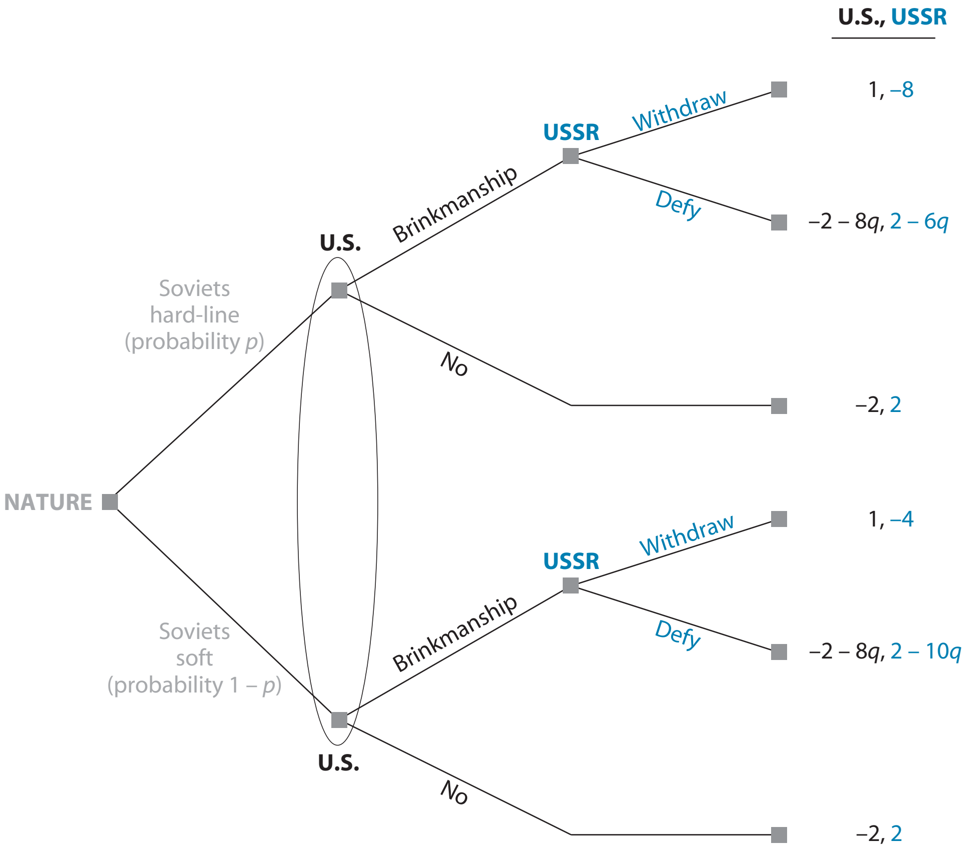

FIGURE 14.4 The Brinkmanship Model of the Crisis

and the sure payoff -4 if they withdraw. For them, withdrawal is better if $-4 > 2 - 10q$, or $10q > 6$, or $q > 0.6$. Thus, U.S. brinkmanship must contain at least a 60% probability of war; otherwise it will not deter the Soviets, even if they are the soft type. We call this lower bound on the probability q the **effectiveness condition**.

Observe how the expected payoffs for U.S. brinkmanship and Soviet defiance shown in Figure 14.4 relate to the simple-threat model of Figure 14.3; the latter can now be thought of as a special case of the general brinkmanship-threat model of Figure 14.4, corresponding to the extreme value $q = 1$.

We can solve the game shown in Figure 14.4 in the usual way. We have already seen that along the upper path the Soviets, being hard-line, will defy the United States and that along the lower path the soft Soviets will comply with U.S. demands if the effectiveness condition is satisfied. If this condition is not satisfied, then both types of Soviets will defy the United States; so the latter would do better never to make this threat at all. So let us proceed by assuming that the soft Soviets will comply; we look at the U.S. choices. Basically, how risky can the U.S. threat be and still remain tolerable to the United States?

If the United States makes the threat, it runs the risk, p, that it will encounter the hard-line Soviets, who will defy the threat. Then the expected U.S. payoff will be $(-2 - 8q)$, as calculated before. The probability is $(1 - p)$ that the United States will encounter the soft-type Soviets. We are assuming that they comply; then the United States gets a 1. Therefore, the expected payoff to the United States from the probabilistic threat, assuming that it is effective against the soft-type Soviets, is

$$(-2 - 8q) \times p + 1 \times (1 - p) = -8pq - 3p + 1.$$

If the United States refrains from making a threat, it gets a -2. Therefore, the condition for the United States to make the threat is

$$-8pq - 3p + 1 > -2 \quad \text{or}$$

$$q < \frac{3}{8}\,\frac{1-p}{p} = \frac{0.375(1-p)}{p}.$$

That is, the probability of war must be small enough to satisfy this expression or the United States will not make the threat at all. We call this upper bound on q the **acceptability condition**. Note that p enters the formula for the maximum value of q that will be acceptable to the United States; the larger the chance that the Soviets will not give in, the smaller the risk of mutual disaster that the United States finds acceptable.

If the probabilistic threat is to work, it should satisfy both the effectiveness condition and the acceptability condition. We can determine the appropriate level of the probability of war by using Figure 14.5. The horizontal axis is the probability, p, that the Soviets are hard-line, and the vertical axis is the probability, q, that war will occur if they defy the U.S. threat. The horizontal line $q = 0.6$ gives the lower limit of the effectiveness condition; the threat should be such that its associated (p, q) combination is above this line if it is to work even against the soft-type Soviets. The curve $q = 0.375(1 - p)/p$ gives the upper limit of the acceptability condition; the threat should be such that (p, q) is below this curve if it is to be tolerable to the United States even with the assumption that it works against the soft-type Soviets. Therefore, an effective and acceptable threat should fall somewhere between these two lines, above and to the left of their point of intersection, at $p = 0.38$ and $q = 0.6$ (shown as a gray "wedge" in Figure 14.5).

The curve reaches $q = 1$ when $p = 0.27$. For values of p less than this value, the dire threat (certainty of war) is acceptable to the United States and is effective against the soft-type Soviets. This just confirms our analysis in Section 3.

For values of p in the range from 0.27 to 0.38, the dire threat with $q = 1$ puts (p, q) to the right of the acceptability condition and is too large to be tolerable to the United States. But a scaled-down threat can be found. For this range of values of p, some values of q are low enough to be acceptable to the United States and yet high enough to compel the soft-type Soviets. Brinkmanship (using a

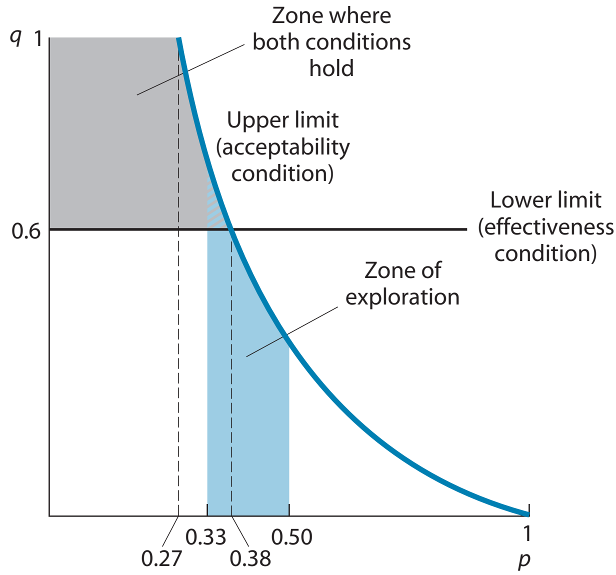

FIGURE 14.5 Conditions of Successful Brinkmanship

probabilistic threat) can do the job in this situation, whereas a simple dire threat would be too risky.

If p exceeds 0.38, then no value of q satisfies both conditions. If the probability that the Soviets will never give in is greater than 0.38, then any threat large enough to work against the soft-type Soviets ($q \geq 0.6$) creates a risk of war too large to be acceptable to the United States. If $p \geq 0.38$, therefore, the United States cannot help itself by using the brinkmanship strategy.

5 PRACTICING BRINKMANSHIP

If Kennedy has a very good estimate of the probability, p, of the Soviets being hard-liners, and is very confident about his ability to control the risk, q, that the blockade will lead to nuclear war, then he can calculate and implement his best strategy. As we saw in Section 3, if $p < 0.27$, the dire threat of a certainty of war is acceptable to Kennedy. (Even then he will prefer to use the smallest effective threat—namely, $q = 0.6$.) If p is between 0.27 and 0.38, then he has to use brinkmanship. Such a threat has to have the risk of disaster $0.6 < q < 0.375(1 - p)/p$, and again Kennedy prefers the smallest of this range—namely, $q = 0.6$. If $p > 0.38$, then he should give in.

In practice, Kennedy does not know p precisely; he only estimates that it lies within the range from 1/3 to 1/2. Similarly, he cannot be confident about the exact location of the critical value of q in the acceptability condition. That depends on the numbers used for the Soviet payoffs in various outcomes—for example, −8 (for war) versus −4 (for compliance)—and Kennedy can only estimate these values. Finally, he may not even be able to control the risk created by his brinkmanship action very precisely. All these ambiguities make it necessary to proceed cautiously.

Suppose Kennedy thinks that $p = 0.35$ and issues a threat backed by an action that carries the risk $q = 0.65$. The risk is greater than what is needed to be effective—namely, 0.6. The limit of acceptability is $0.375 \times (1 - 0.35)/0.35 = 0.7$, and the risk $q = 0.65$ is less than this limit. Thus, according to Kennedy's calculations, the risk satisfies both of the conditions—effectiveness and acceptability. However, suppose Kennedy is mistaken. For example, if he has not realized that LeMay might actually defy orders and take an excessively aggressive action, then q may in reality be higher than Kennedy thinks it is; for example, q may equal 0.8, which Kennedy would regard as too risky. Or suppose p is actually 0.4; then Kennedy would regard even $q = 0.65$ as too risky. Or Kennedy's experts may have misestimated the values of the Soviet payoffs. If they rate the humiliation of withdrawal as −5 instead of −4, then the threshold of the effectiveness condition will actually be $q = 0.7$, and Kennedy's threat with $q = 0.65$ will go wrong.

All that Kennedy knows is that the general shape of the effectiveness and acceptability conditions is like that shown in Figure 14.5. He does not know p for sure. Therefore, he does not know exactly what value of q to choose to fulfill both the effectiveness and the acceptability conditions; indeed, he does not even know if such a range exists for the unknown true value of p: it might be greater than or less than the borderline value of 0.38 that divides the two cases. And he is not able to fix q very precisely; therefore, even if he knew p, he would not be able to act confident of his willingness to tolerate the resulting risk.

With such hazy information, imprecise control, and large risks, what is Kennedy to do? He has to *explore* the boundaries of the Soviets' risk tolerance as well as his own. It would not do to start the exploration with a value of q that might turn out to be too high. Instead, Kennedy must explore the boundaries "from below"; he must start with something quite safe and gradually increase the level of risk to see "who blinks first." That is exactly how brinkmanship is practiced in reality.

We explain this with the aid of Figure 14.5. Observe the color-shaded area. Its left and right boundaries, $p = 1/3$ and $p = 1/2$, correspond to the limits of Kennedy's estimated range of p. The lower boundary is the horizontal axis ($q = 0$). The upper boundary is composed of two segments. For $p < 0.38$, this segment corresponds to the effectiveness condition; for $p > 0.38$, it corresponds to the acceptability condition. Remember that Kennedy does not know the precise positions of these boundaries but must grope toward them from below. Therefore, the color-shaded region is where he must start the process.

Suppose Kennedy starts with a very safe action—say, q equaling approximately 0.01 (1%). In our context of the Cuban missile crisis, we can think of this as his television speech, which announced that a quarantine would soon go into effect. At this juncture, the point with coordinates (p, q) lies somewhere near the bottom edge of the shaded region. Kennedy does not know exactly where, because he does not know p for sure. But the overwhelming likelihood is that at this point the threat is quite safe but also ineffective. Therefore, Kennedy escalates it a little bit. That is, he moves the point (p, q) in a vertically upward direction from wherever it was initially. This could be the actual start of the quarantine. If that proves to be still safe but ineffective, he jacks up the risk one more notch. This could be the leaking of information about bombing plans.

As he proceeds in this way, eventually his exploration will encounter one of the boundaries of the color-shaded area in Figure 14.5, and which boundary this is depends on the value of p. One of two things comes to pass. Either the threat becomes serious enough to deter the Soviets; this happens if the true value of p is less than its true critical value, here 0.38. On the diagram, we see this as a movement out of the color-shaded area and into the area in which the threat is both acceptable *and* effective. Then the Soviets concede and Kennedy has won. Or the threat becomes too risky for the United States; this happens if $p > 0.38$. Kennedy's exploration in this case pushes him above the acceptability condition. Then Kennedy decides to concede, and Khrushchev has won. Again we point out that because Kennedy is not sure of the true value of p, he does not know in advance which of these two outcomes will prevail. As he gradually escalates the risk, he may get some clues from Soviet behavior that enable him to make his estimate of p somewhat more precise. Eventually he will reach sufficient precision to know which part of the boundary he is headed toward and therefore whether the Soviets will concede or the United States must be the player to do so.

Actually, there are two possible outcomes only so long as the ever-present and steadily increasing mutual risk of disaster does not come to pass while Kennedy is groping through the range of ever more risky military options. Therefore, there is a third possibility—namely, that the explosion occurs before either side recognizes that it has reached its limit of tolerance of risk and climbs down. This continuing and rising risk of a very bad outcome is what makes brinkmanship such a delicate and dangerous strategy.

Thus, brinkmanship in practice is the **gradual escalation of the risk of mutual harm**. It can be visualized vividly as **chicken in real time**. In our analysis of chicken in Chapter 4, we gave each player a simple binary choice: either go straight or swerve. In reality, the choice is usually one of timing. The two cars are rushing toward each other, and either player can choose to swerve at any time. When the cars are very far apart, swerving ensures safety. As they get closer together, they face an ever-increasing risk that they will collide anyway, and even swerving will not avoid a collision. As the two players continue to drive toward one another, each is exploring the limit of the other's willingness to take this risk and is perhaps at the same time exploring his own limit. The one who hits that limit first swerves. But there is always the risk that they have left it long enough and are close enough that, even after choosing Swerve, they can no longer avoid the collision.

Now we see why, in the Cuban missile crisis, the very features that make it inaccurate to regard it as a two-person game make it easier to practice such brinkmanship. The blockade was a relatively small action, unlikely to start a nuclear war at once. But once Kennedy set the blockade in motion, its operation, escalation, and other features were not totally under his control. So Kennedy was not saying to Khrushchev, "If you defy me (cross a sharp brink), I will coolly and deliberately launch a nuclear war that will destroy both our peoples." Rather, he was saying, "The wheels of the blockade have started to turn and are gathering their own momentum. The more or longer you defy me, the more likely it is that some operating procedure will slip up, the political pressure on me will rise to a point where I must give in, or some hawk will run amok. If this risk comes to pass, I will be unable to prevent nuclear war, no matter how much I may regret it at that point. Only you can now defuse the tension by complying with my demand to withdraw the missiles."

We believe that this perspective gives a much better and deeper understanding of the crisis than can most analyses based on simple threats. It tells us why the *risk* of war played such an important role in all discussions. It even makes Allison's compelling arguments about bureaucratic procedures and internal divisions on both sides an integral part of the picture: these features allow the top-level players on both sides credibly to lose some control—that is, to practice brinkmanship.

One important condition remains to be discussed. In Chapter 9, we saw that every threat has an associated implicit promise—namely, that the bad consequence will not take place if your opponent complies with your wishes. The same is required for brinkmanship. If, as you are increasing the level of risk, your opponent does comply, you must be able to "go into reverse"—begin reducing the risk immediately and quite quickly remove it from the picture. Otherwise, the opponent would not gain anything by compliance. This may have been a problem in the Cuban missile crisis. If the Soviets feared that Kennedy could not control hawks such as LeMay ("We ought to just go in there today and knock 'em off"), they would gain nothing by giving in.

To reemphasize and sum up, brinkmanship is the strategy of exposing your rival and yourself to a gradually increasing risk of mutual harm. The actual occurrence of the harmful outcome is not totally within the threatener's control.

Viewed in this way, brinkmanship is everywhere. In most confrontations—for example, between a company and a labor union, a husband and a wife, a parent and a child, and the President and Congress—one player cannot be sure of the other party's objectives and capabilities. Therefore, most threats carry a risk of error, and every threat must contain an element of brinkmanship. We hope that we have given you some understanding of this strategy and that we have impressed on you the risks that it carries. Unsuccessful brinkmanship can lead to a labor strike, the dissolution of a marriage, or the down-grading of U.S. bonds as was discovered by President Obama and members of Congress following their 2011 dispute over raising the nation's debt ceiling. You will have to face up to brinkmanship or to conduct it yourself on many occasions in your personal and professional lives. Please do so carefully, with a clear understanding of its potentialities and risks.

To help you do so, we now recapitulate the important lessons learned from the handling of the Cuban missile crisis, reinterpreted as a labor union leadership contemplating a strike in pursuit of its wage demand, unsure whether this action will result in the whole firm's shutting down:

1. Start small and safe. Your first step should not be an immediate walkout; it should be to schedule a membership meeting at a date a few days or weeks hence, while negotiations continue.
2. Raise the risks gradually. Your public and private statements, as well as the stirring up of the sentiments of the membership, should induce management to believe that acceptance of its current low-wage offer is becoming less and less likely. If possible, stage small incidents—for example, a few one-day strikes or local walkouts.
3. As this process continues, read and interpret signals in management's actions to figure out whether the firm has enough profit potential to afford the union's high-wage demand.
4. Retain enough control over the situation; that is, retain the power to induce your membership to ratify the agreement that you will reach with management; otherwise management will think that the risk will not de-escalate even if it concedes to your demands.

SUMMARY

In some game situations, the risk of error in the presence of a threat may call for the use of as small a threat as possible. When a large threat cannot be reduced in other ways, it can be scaled down by making its fulfillment probabilistic. Strategic use of *probabilistic threat*, in which you expose your rival and yourself to an increasing risk of harm, is called brinkmanship.

Brinkmanship requires a player to relinquish control over the outcome of the game without completely losing control. You must create a threat with a risk level that is both large enough to be effective in compelling or deterring your rival and small enough to be acceptable to you. To do so, you must determine the levels of risk tolerance of both players through a *gradual escalation of the risk of mutual harm.*

The Cuban missile crisis of 1962 serves as a case study in the use of brinkmanship on the part of President Kennedy. Analyzing the crisis as an example of a simple threat, with the U.S. blockade of Cuba establishing credibility, is inadequate. A better analysis accounts for the many complexities and uncertainties inherent in the situation and the likelihood that a simple threat was too risky. Because the actual crisis included numerous political and military players, Kennedy was able to achieve "controlled loss of control" by ordering the blockade and gradually letting incidents and tension escalate, until Khrushchev yielded in the face of the rising risk of nuclear war.

KEY TERMS

acceptability condition (578)
chicken in real time (581)
effectiveness condition (577)
gradual escalation of the risk of mutual harm (581)
probabilistic threat (575)

SOLVED EXERCISES

S1. Consider a game between a union and the company that employs the union membership. The union can threaten to strike (or not) to get the company to meet its wage and benefits demands. When faced with a threatened strike, the company can choose to concede to the demands of the union or to defy its threat of a strike. The union, however, does not know the company's profit position when it decides whether to make its threat; it does not know whether the company is sufficiently profitable to meet its demands—and the company's assertions in this matter cannot be believed. Nature determines whether the company is profitable; the probability that the firm is unprofitable is p.

The payoff structure is as follows: (i) When the union makes no threat, the union gets a payoff of 0 (regardless of the profitability of the company). The company gets a payoff of 100 if it is profitable but a payoff of 10 if it is unprofitable. A passive union leaves more profit for the company if there is any profit to be made. (ii) When the union threatens to strike and the company concedes, the union gets 50 (regardless of the profitability of the company) and the company gets 50 if it is profitable but −40 if it is not. (iii) When the union threatens to strike and the company defies the union's threat, the union must strike and gets −100 (regardless of the profitability of the company). The company gets −100 if it is profitable and −10 if it is not. Defiance is very costly for a profitable company but not so costly for an unprofitable one.

(a) What happens when the union uses the pure threat to strike unless the company concedes to the union's demands?

(b) Suppose that the union sets up a situation in which there is some risk, with probability $q < 1$, that it will strike after the company defies its threat. This risk may arise from the union leadership's imperfect ability to keep the membership in line. Draw a game tree similar to Figure 14.4. for this game.

(c) What happens when the union uses brinkmanship, threatening to strike with some probability q unless the company accedes to its demands?

(d) Derive the effectiveness and acceptability conditions for this game, and determine the values for p and q for which the union can use a pure threat, brinkmanship, or no threat at all.

S2. Scenes from many movies illustrate the concept of brinkmanship. Analyze the following descriptions from this perspective. What are the risks the two sides face? How do those risks increase during the course of the execution of the brinkmanship threat?

(a) In the 1980 film *The Gods Must Be Crazy*, the only survivor of a rebel team that tried to assassinate the president of an African country has been captured and is being interrogated. He stands blindfolded with his back to the open door of a helicopter. Above the noise of the helicopter rotors, an officer asks him, "Who is your leader? Where is your hideout?" The man does not answer, and the officer pushes him out of the door. In the next scene, we see that although its engine is running, the helicopter is actually on the ground, and the man has fallen 6 feet on his back. The officer appears at the door and says, laughing, "Next time it will be a little higher."

(b) In the 1998 film *A Simple Plan*, two brothers remove some of a \$4.4 million ransom payment that they find in a crashed airplane. After many intriguing twists of fate, the remaining looter, Hank, finds himself in conference with an FBI agent. The agent, who suspects but cannot prove that Hank has some of the missing money, fills Hank in on the story of the money's origins and tells him that the FBI possesses the serial numbers of about 1 of every 10 of the bills in that original ransom payment. The agent's final words to Hank are, "Now it's simply a matter of waiting for the numbers to turn up. You can't go around passing \$100 bills without eventually sticking in someone's memory."

S3. In this exercise, we provide a couple examples of the successful use of brinkmanship, where "success" is indicative of the two sides' reaching a mutually acceptable deal. For each example, (i) identify the interests of the parties; (ii) describe the nature of the uncertainty inherent in the situation; (iii) give the strategies the parties used to escalate the risk of disaster; (iv) discuss whether the strategies were good ones; and (v) **(Optional)** if you can, set up a small mathematical model of the kind presented in this chapter. In each case, we provide a few readings to get you started; you should locate more by using the resources of your library and resources on the World Wide Web such as Lexis-Nexis.

(a) The Uruguay Round of international trade negotiations that started in 1986 and led to the formation of the World Trade Organization in 1994. *Reading*: John H. Jackson, *The World Trading System*, 2nd ed. (Cambridge, Mass.: MIT Press, 1997), pp. 44–49 and ch. 12 and 13.

(b) The Camp David Accords between Israel and Egypt in 1978. *Reading*: William B. Quandt, *Camp David: Peacemaking and Politics* (Washington, D.C.: Brookings Institution, 1986).

S4. The following examples illustrate the unsuccessful use of brinkmanship, where brinkmanship is considered "unsuccessful" when the mutually bad outcome (disaster) occurs. Answer the questions outlined in Exercise S3 for the following situations:

(a) The confrontation between the regime and the student prodemocracy demonstrators in Beijing in June 1989. *Readings*: Donald Morrison, ed., *Massacre in Beijing: China's Struggle for Democracy* (New York: Time Magazine Publications, 1989); Suzanne Ogden, Kathleen Hartford, L. Sullivan, and D. Zweig, eds., *China's Search for Democracy: The Student and Mass Movement of 1989* (Armonk, N.Y.: M.E. Sharpe, 1992).

(b) The Caterpillar strike, from 1991 to 1998. *Readings*: "The Caterpillar Strike: Not Over Till It's Over," *Economist*, February 28, 1998; "Caterpillar's Comeback," *Economist*, June 20, 1998; Aaron Bernstein, "Why Workers Still Hold a Weak Hand," *BusinessWeek*, March 2, 1998.

S5. Answer the questions listed in Exercise S3 for these potential cases for brinkmanship in the future:

(a) A Taiwanese declaration of independence from the People's Republic of China. *Reading*: Ian Williams, "Taiwan's Independence," *Foreign Policy in Focus*, December 20, 2006. Available at www.fpif.org/fpiftxt/3815.

(b) The militarization of space, for example, the positioning of weapons in space or the shooting down of satellites. *Reading*: "Disharmony in the Spheres," *Economist*, January 17, 2008. Available at www.economist.com/node/10533205.

UNSOLVED EXERCISES

U1. In the chapter, we argue that the payoff to the United States is −10 when (either type) Soviets defy the U.S. threat; these payoffs are illustrated in Figure 14.3. Suppose now that this payoff is in fact −12 rather than −10.

(a) Incorporate this change in payoff into a game tree similar to the one in Figure 14.4.

(b) Using the payoffs from your game tree in part (a), find the effectiveness condition for this version of the U.S.–USSR brinkmanship game.

(c) Using the payoffs from part (a), find the acceptability condition for this game.

(d) Draw a diagram similar to that in Figure 14.5, illustrating the effectiveness and acceptability conditions found in parts (b) and (c).

(e) For what values of p, the probability that the Soviets are hard-line, is the pure threat ($q = 1$) acceptable? For what values of p is the pure threat unacceptable but brinkmanship still possible?

(f) If Kennedy was correct in believing that p lay between 1/3 and 1/2, does your analysis of this version of the game suggest that an effective *and* acceptable probabilistic threat existed? Use this example to explain how a game theorist's assumptions about player payoffs can have a major effect on the predictions that arise from the theoretical model.

U2. Answer the questions from Exercise S2 for the following movies:

(a) In the 1941 movie classic *The Maltese Falcon*, the hero, Sam Spade (Humphrey Bogart), is the only person who knows the location of the immensely valuable gem-studded falcon figure, and the villain, Caspar Gutman (Sydney Greenstreet), is threatening to torture him for that information. Spade points out that torture is useless unless the threat of death lies behind it, and Gutman cannot afford to kill Spade, because then the information dies with him. Therefore, he may as well not bother with the threat of torture. Gutman replies, "That is an attitude, sir, that calls for the most delicate judgment on both sides, because, as you know, sir, men are likely to forget in the heat of action where their best interests lie and let their emotions carry them away."

(b) The 1925 Soviet classic *The Battleship Potemkin* (set in the summer of 1905) closes with a squadron of ships from the tsar's Black Sea fleet chasing the mutinous and rebellious crew of the *Potemkin*. The tension mounts as the ships draw ever closer. Men on each side race to their battle stations, load and aim the huge guns, and wait nervously for the order to fire on their countrymen. Neither side wants to attack the other, but neither wants to back down or to die without defending itself. The tsar's ships have orders to take the *Potemkin* by any means necessary, and the crew knows it will be tried for treason if it surrenders.

U3. Answer the questions in Exercise S3 for these examples of successful brinkmanship:

(a) The negotiations between the South African apartheid regime and the African National Congress to establish a new constitution with majority rule, 1989 to 1994. *Reading*: Allister Sparks, *Tomorrow Is Another Country* (New York: Hill and Wang, 1995).

(b) Peace in Northern Ireland: disarmament of the IRA in July 2005, the St. Andrews Agreement of October 2006, the elections of March 2007, and the power-sharing government of Ian Paisley and Martin McGuinness. *Reading*: "The Thorny Path to Peace and Power Sharing," CBC News, March 26, 2007. Available at www.cbc.ca/news2/background/northern-ireland/timeline.html.

U4. Answer the questions in Exercise S3 for these examples of unsuccessful brinkmanship:

(a) The U.S. budget confrontation between President Clinton and the Republican-controlled Congress in 1995. *Readings*: Sheldon Wolin, "Democracy and Counterrevolution," *Nation*, April 22, 1996; David Bowermaster, "Meet the Mavericks," *U.S. News and World Report*, December 25, 1995–January 1, 1996; "A Flight that Never Seems to End," *Economist*, December 16, 1995.

(b) The television writers' strike of 2007–2008. *Readings*: "Writers Guild of America," online archive of the *New York Times* on the Writers Guild and the strike. Available at http://topics.nytimes.com/top/reference/timestopics/organizations/w/writers_guild_of_america/index.html; "Writers Strike: A Punch from the Picket Line." Available at http://writers-strike.blogspot.com.

U5. Answer the questions in Exercise S3 for these potential cases of future brinkmanship:

(a) The stationing of an American antiballistic missile launch site in Poland with an accompanying radar site in the Czech Republic, ostensibly intended to intercept missiles from Iran but angering Russia. *Reading*: "Q&A: US Missile Defence," *BBC News*, August 20, 2008. Available at http://news.bbc.co.uk/2/hi/europe/6720153.stm.

(b) Deterring Iran from obtaining nuclear weapons. *Readings*: James Fallows, "The Nuclear Power Beside Iraq," *Atlantic*, May 2006. Available at www.theatlantic.com/doc/200605/fallows-iran; James Fallows, "Will Iran Be Next?" *Atlantic*, December 2004. Available at www.theatlantic.com/magazine/archive/2006/05/the-nuclear-power-beside-iraq/304819.

17

Bargaining

PEOPLE ENGAGE IN BARGAINING throughout their lives. Children start by negotiating to share toys and to play games with other children. Couples bargain about matters of housing, child rearing, and the adjustments that each must make for the other's career. Buyers and sellers bargain over price, workers and bosses over wages. Countries bargain over policies of mutual trade liberalization; superpowers negotiate mutual arms reduction. And the two original authors of this book had to bargain with one another—generally very amicably—about what to include or exclude, how to structure the exposition, and so forth. To get a good result from such bargaining, the participants must devise good strategies. In this chapter, we raise and explicate some of these basic ideas and strategies.

All bargaining situations have two things in common. First, the total payoff that the parties to the negotiation are capable of creating and enjoying as a result of reaching an agreement should be greater than the sum of the individual payoffs that they could achieve separately—the whole must be greater than the sum of the parts. Without the possibility of this excess value, or "surplus," the negotiation would be pointless. If two children considering whether to play together cannot see a net gain from having access to a larger total stock of toys or from one another's company in play, then it is better for each to "take his toys and play by himself." The world is full of uncertainty, and the expected benefits may not materialize. But when engaged in bargaining, the parties must at least perceive some gain therefrom: when he agreed to sell his soul to the Devil, Faust thought the benefits of knowledge and power that he gained were worth the price that he would eventually have to pay.

The second important general point about bargaining follows from the first: it is not a zero-sum game. When a surplus exists, the negotiation is about how to divide it up. Each bargainer tries to get more for himself and leave less for the others. This may appear to be zero sum, but behind it lies the danger that if the agreement is not reached, no one will get any surplus at all. This mutually harmful alternative, as well as *both* parties' desire to avoid it, is what creates the potential for the threats—explicit and implicit—that make bargaining such a strategic matter.

Before the advent of game theory, one-on-one bargaining was generally thought to be a difficult and even indeterminate problem. Observation of widely different outcomes in otherwise similar-looking situations lent support to this view. Theorists were not able to achieve any systematic understanding of why one party gets more than another and attributed this result to vague and inexplicable differences in "bargaining power."

Even the simple theory of Nash equilibrium does not take us any further. Suppose two people are to split $1. Let us construct a game in which each is asked to announce what he would want. The moves are simultaneous. If their announcements x and y add up to 1 or less, each gets what he announced. If they add up to more than 1, neither gets anything. Then *any* pair (x, y) adding to 1 constitutes a Nash equilibrium in this game: *given* the announcement of the other, each player cannot do better than to stick to his own announcement.[1]

Further advances in game theory have brought progress along two quite different lines, each using a distinct mode of game-theoretic reasoning. In Chapter 2, we distinguished between cooperative-game theory, in which the players decide and implement their actions jointly, and noncooperative-game theory, in which the players decide and take their actions separately. Each of the two lines of advance in bargaining theory uses one of these two approaches. One approach views bargaining as a *cooperative* game, in which the parties find and implement a solution jointly, perhaps using a neutral third party such as an arbitrator for enforcement. The other approach views bargaining as a *noncooperative* game, in which the parties choose strategies separately and we look for an equilibrium. However, unlike our earlier simple game of simultaneous announcements, whose equilibrium was indeterminate, here we impose more structure and specify a sequential-move game

[1] As we saw in Chapter 5, Section 3.B, this type of game can be used as an example to bolster the critique that the Nash-equilibrium concept is too imprecise. In the bargaining context, we might say that the multiplicity of equilibria is just a formal way of showing the indeterminacy that previous analysts had claimed.

of offers and counteroffers, which leads to a determinate equilibrium. As in Chapter 2, we emphasize that the labels "cooperative" and "noncooperative" refer to joint versus separate actions, not to nice versus nasty behavior or to compromise versus breakdown. The equilibria of noncooperative bargaining games can entail a lot of compromise.

1 NASH'S COOPERATIVE SOLUTION

In this section, we present Nash's cooperative-game approach to bargaining. First we present the idea in a simple numerical example; then we develop the more general algebra.[2]

A. Numerical Example

Imagine two Silicon Valley entrepreneurs, Andy and Bill. Andy produces a microchip set that he can sell to any computer manufacturer for $900. Bill has a software package that can retail for $100. The two meet and realize that their products are ideally suited to each other and that, with a bit of trivial tinkering, they can produce a combined system of hardware and software worth $3,000 in each computer. Thus, together they can produce an extra value of $2,000 per unit, and they expect to sell millions of these units each year. The only obstacle that remains on this path to fortune is to agree to a division of the spoils. Of the $3,000 revenue from each unit, how much should go to Andy and how much to Bill?

Bill's starting position is that without his software, Andy's chip set is just so much metal and sand, so Andy should get only the $900 and Bill himself should get $2,100. Andy counters that without his hardware, Bill's programs are just symbols on paper or magnetic signals on a diskette, so Bill should get only $100, and $2,900 should go to him, Andy.

Watching them argue, you might suggest they "split the difference." But that is not an unambiguous recipe for agreement. Bill might offer to split the profit on each unit equally with Andy. Under this scheme, each will get a profit of $1,000, meaning that $1,100 of the revenue goes to Bill and $1,900 to Andy. Andy's response might be that they should have an equal percentage of profit on their contribution to the joint enterprise. Thus, Andy should get $2,700 and Bill $300.

The final agreement depends on their stubbornness or patience if they negotiate directly with one another. If they try to have the dispute arbitrated by a third party, the arbitrator's decision depends on his sense of the relative value

[2] John F. Nash Jr., "The Bargaining Problem," *Econometrica*, vol. 18, no. 2 (1950), pp. 155–62.

of hardware and software and on the rhetorical skills of the two principals as they present their arguments before the arbitrator. For the sake of definiteness, suppose the arbitrator decides that the division of the profit should be 4:1 in favor of Andy; that is, Andy should get four-fifths of the surplus while Bill gets one-fifth, or Andy should get four times as much as Bill. What is the actual division of revenue under this scheme? Suppose Andy gets a total of x and Bill gets a total of y; thus Andy's profit is $(x - 900)$ and Bill's is $(y - 100)$. The arbitrator's decision implies that Andy's profit should be four times as large as Bill's; so $x - 900 = 4(y - 100)$, or $x = 4y + 500$. The total revenue available to both is \$3,000; so it must also be true that $x + y = 3{,}000$, or $x = 3{,}000 - y$. Then $x = 4y + 500 = 3{,}000 - y$, or $5y = 2{,}500$, or $y = 500$, and thus $x = 2{,}500$. This division mechanism leaves Andy with a profit of $2{,}500 - 900 = \$1{,}600$ and Bill with $500 - 100 = \$400$, which is the 4:1 split in favor of Andy that the arbitrator wants.

We now develop this simple data into a general algebraic formula that you will find useful in many practical applications. Then we go on to examine more specifics of what determines the ratio in which the profits in a bargaining game get split.

B. General Theory

Suppose two bargainers, A and B, seek to split a total value v, which they can achieve if and only if they agree on a specific division. If no agreement is reached, A will get a and B will get b, each by acting alone or in some other way acting outside of this relationship. Call these their *backstop* payoffs or, in the jargon of the Harvard Negotiation Project, their **BATNAs (best alternative to a negotiated agreement)**.[3] Often a and b are both zero, but, more generally, we only need to assume that $a + b < v$, so that there is a positive **surplus** $(v - a - b)$ from agreement; if this were not the case, the whole bargaining would be moot because each side would just take up its outside opportunity and get its BATNA.

Consider the following rule: each player is to be given his BATNA plus a share of the surplus, a fraction h of the surplus for A and a fraction k for B, such that $h + k = 1$. Writing x for the amount that A finally ends up with, and similarly y for B, we translate these statements as

$$x = a + h(v - a - b) = a(1 - h) + h(v - b)$$
$$x - a = h(v - a - b)$$

and

$$y = b + k(v - a - b) = b(1 - k) + k(v - a)$$
$$y - b = k(v - a - b).$$

[3] See Roger Fisher and William Ury, *Getting to Yes*, 2nd ed. (New York: Houghton Mifflin, 1991).

We call these expressions the Nash formulas. Another way of looking at them is to say that the surplus $(v - a - b)$ gets divided between the two bargainers in the proportions of h: k, or

$$\frac{y - b}{x - a} = \frac{k}{h}$$

or, in slope-intercept form,

$$y = b + \frac{k}{h}(x - a) = \left(b - \frac{ak}{h}\right) + \frac{k}{h}x.$$

To use up the whole surplus, x and y must also satisfy $x + y = v$. The Nash formulas for x and y are actually the solutions to these last two simultaneous equations.

A geometric representation of the **Nash cooperative solution** is shown in Figure 17.1. The backstop, or BATNA, is the point P, with coordinates (a, b). All points (x, y) that divide the gains in proportions h: k between the two players lie along the straight line passing through P and having slope k/h; this slope is just the line $y = b + (k/h)(x - a)$ that we derived earlier. All points (x, y) that use up the whole surplus lie along the straight line joining $(v, 0)$ and $(0, v)$; this line is the second equation that we derived—namely, $x + y = v$. The Nash solution is at the intersection of the lines, at the point Q. The coordinates of this point are the parties' payoffs after the agreement.

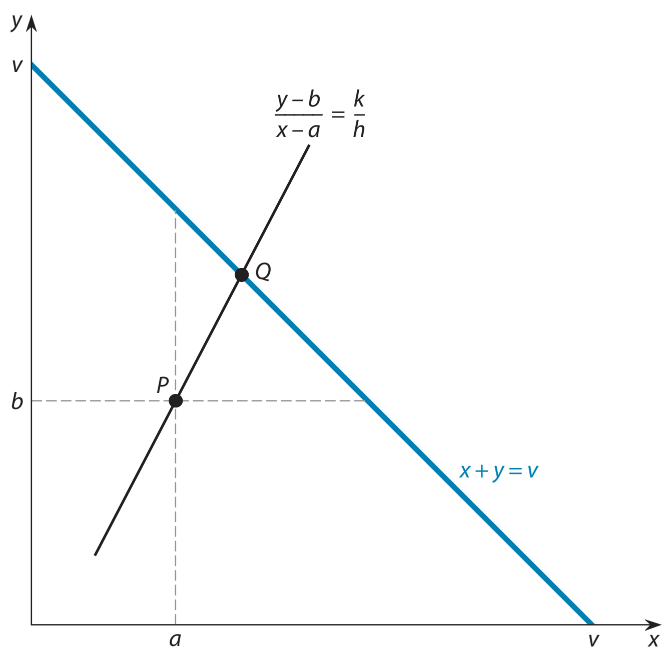

FIGURE 17.1 The Nash Bargaining Solution in the Simplest Case

The Nash formula says nothing about how or why such a solution might come about. And this vagueness is its merit—it can be used to encapsulate the results of many different theories taking many different perspectives.

At the simplest, you might think of the Nash formula as a shorthand description of the outcome of a bargaining process that we have not specified in detail. Then h and k can stand for the two parties' relative bargaining strengths. This shorthand description is a cop-out; a more complete theory should explain where these bargaining strengths come from and why one party might have more than the other. We do so in a particular context later in the chapter. In the meantime, by summarizing any and all of the sources of bargaining strength in these numbers h and k, the formula has given us a good tool.

Nash's own approach was quite different—and indeed different from the whole approach to game theory that we have taken thus far in this book. Therefore, it deserves more careful explanation. In all the games that we have studied so far, the players chose and played their strategies separately from each other. We have looked for equilibria in which each player's strategy was in his own best interest, given the strategies of the others. Some such outcomes were very bad for some or even all of the players, the prisoners' dilemma being the most prominent example. In such situations, there was scope for the players to get together and agree that all would follow some particular strategy. But in our framework, there was no way in which they could be sure that the agreement would hold. After reaching an agreement, the players would disperse, and, when it was each player's turn to act, he would actually take the action that served his own best interest. The agreement on joint action would unravel in the face of such separate temptations. True, in considering repeated games in Chapter 10, we found that the implicit threat of the collapse of an ongoing relationship might sustain an agreement, and, in Chapter 8, we did allow for communication by signals. But individual action was of the essence, and any mutual benefit could be achieved only if it did not fall prey to the selfishness of separate individual actions. In Chapter 2, we called this approach to game theory *noncooperative*, emphasizing that the term signified how actions are taken, not whether outcomes are jointly good. The important point, again, is that any joint good has to be an equilibrium outcome of separate action in such games.

What if joint action *is* possible? For example, the players might take all their actions immediately after the agreement is reached, in one another's presence. Or they might delegate the implementation of their joint agreement to a neutral third party, or to an arbitrator. In other words, the game might be *cooperative* (again in the sense of joint action). Nash modeled bargaining as a cooperative game.

The thinking of a collective group that is going to implement a joint agreement by joint action can be quite different from that of a set of individual people who know that they are *interacting* strategically but are *acting* noncooperatively. Whereas the latter set will think in terms of an equilibrium and then delight or grieve, depending on whether they like the results, the former can think first of what is a good outcome and then see how to implement it. In other words, the theory defines the outcome of a cooperative game in terms of some general principles or properties that seem reasonable to the theorist.

Nash formulated a set of such principles for bargaining and proved that they implied a unique outcome. His principles are roughly as follows: (1) the outcome should be invariant if the scale in which the payoffs are measured changes linearly; (2) the outcome should be **efficient**; and (3) if the set of possibilities is reduced by removing some that are irrelevant in the sense that they would not be chosen anyway, then the outcome should not be affected.

The first of these principles conforms to the theory of expected utility, which we discussed briefly in the appendix to Chapter 8. We saw there that a nonlinear rescaling of payoffs represents a change in a player's attitude toward risk and a real change in behavior; a concave rescaling implies risk aversion, and a convex rescaling implies risk preference. A linear rescaling, being the intermediate case between these two, represents no change in the attitude toward risk. Therefore, it should have no effect on expected payoff calculations and no effect on outcomes.

The second principle simply means that no available mutual gain should go unexploited. In our simple example of A and B splitting a total value of v, it would mean that x and y has to sum to the full amount of v available, and not to any smaller amount; in other words, the solution has to lie on the $x + y = v$ line in Figure 17.1. More generally, the complete set of logically conceivable agreements to a bargaining game, when plotted on a graph as in Figure 17.1, will be bounded above and to the right by the subset of agreements that leave no mutual gain unexploited. This subset need not lie along a straight line such as $x + y = v$ (or $y = v - x$); it could lie along any curve of the form $y = f(x)$.

In Figure 17.2, all of the points on and below (that is, "south" and to the "west" of) the thick blue curve labeled $y = f(x)$ constitute the complete set of conceivable outcomes. The curve itself consists of the efficient outcomes; there are no conceivable outcomes that include more of both x and y than the outcomes on $y = f(x)$; so there are no unexploited mutual gains left. Therefore, we call the curve $y = f(x)$ the **efficient frontier** of the bargaining problem.

We can illustrate a curved efficient frontier using the example of efficient risk allocation from Chapter 8, Section1.A. Two farmers, each with a square root utility function, face the risk that equally likely good or bad weather would make their incomes either \$160,000 or \$40,000, yielding each an expected utility of

$$1/2 \times \sqrt{160{,}000} + 1/2 \times \sqrt{40{,}000} = 1/2 \times 400 + 1/2 \times 200 = 300.$$

But their risks are perfectly negatively correlated. One gets good weather only when the other gets bad, so their combined income is \$200,000 no matter which of them gets the good weather. If they negotiate so that the first of them

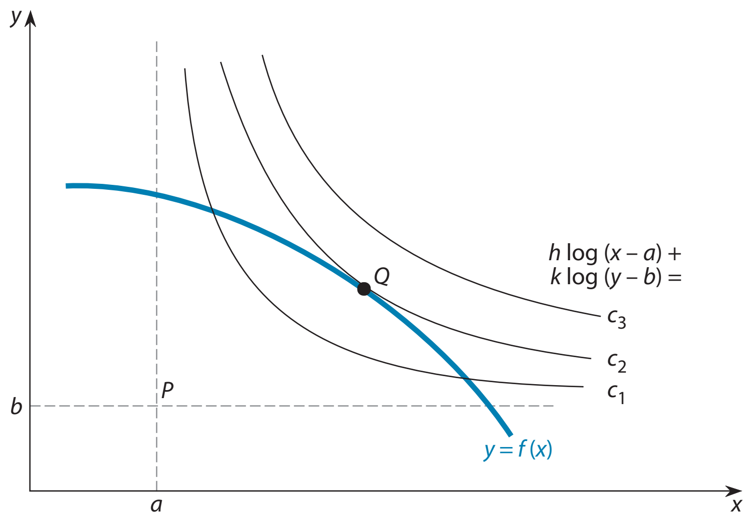

FIGURE 17.2 The General Form of the Nash Bargaining Solution

gets z of the combined income and the other gets the remaining $(200{,}000 - z)$, their respective utilities x and y will be

$$x = \sqrt{z} \quad \text{and} \quad y = \sqrt{200{,}000 - z}.$$

Therefore, we can describe the set of possible risk-sharing outcomes by the equation

$$x^2 + y^2 = z + (200{,}000 - z) = 200{,}000.$$

This equation defines a quarter-circle in the positive quadrant and represents the efficient frontier of the farmers' bargaining problem. The BATNA of each farmer is the expected utility 300 he would get if the two are not able to come to any risk-sharing agreement. Substituting this value into the equation above yields $300^2 + 300^2 = 90{,}000 + 90{,}000 = 180{,}000 < 200{,}000$. So the farmers' BATNA point lies inside the quarter-circle efficient frontier.

The third principle also seems appealing. If an outcome that a bargainer wouldn't have chosen anyway drops out of the picture, what should it matter? This assumption is closely connected to the "independence of irrelevant alternatives" assumption of Arrow's impossibility theorem, which we met in Chapter 15, Section 3, but we must leave the development of this connection to more advanced treatments of the subject.

Nash proved that the cooperative outcome that satisfied all three of these assumptions could be characterized by the mathematical maximization problem: choose x and y to

$$\text{maximize } (x - a)^h (y - b)^k \quad \text{subject to } y = f(x).$$

Here x and y are the outcomes, a and b the backstops, and h and k two positive numbers summing to 1, which are like the bargaining strengths of the Nash formula. The values for h and k cannot be determined by Nash's three assumptions alone; thus they leave a degree of freedom in the theory and in the outcome. Nash actually imposed a fourth assumption on the problem—that of symmetry between the two players; this additional assumption led to the outcome $h = k = 1/2$ and fixed a unique solution. We have given the more general formulation that subsequently became common in game theory and economics.

Figure 17.2 also gives a geometric representation of the objective of the maximization. The black curves labeled c_1, c_2, and c_3 are the level curves, or contours, of the function being maximized; along each such curve, $(x - a)^h(y - b)^k$ is constant and equals c_1, c_2, or c_3 (with $c_1 < c_2 < c_3$) as indicated. The whole space could be filled with such curves, each with its own value of the constant, and curves farther to the "northeast" would have higher values of the constant.

It is immediately apparent that the highest possible value of the function is at that point of tangency, Q, between the efficient frontier and one of the level curves.[4] The location of Q is defined by the property that the contour passing through Q is tangent to the efficient frontier. This tangency is the usual way to illustrate the Nash cooperative solution geometrically.[5]

In our example of Figure 17.1, we can also derive the Nash solution mathematically; to do so requires calculus, but the ends here are more important—at least to the study of games of strategy—than the means. For the solution, it helps to write $X = x - a$ and $Y = y - b$. Thus, X is the amount of the surplus that goes to A, and Y is the amount of the surplus that goes to B. The efficiency of the outcome guarantees that $X + Y = x + y - a - b = v - a - b$, which is just the total surplus and which we will write as S. Then $Y = S - X$, and

$$(x - a)^h(y - b)^k = X^hY^k = X^h(S - X)^k.$$

In the Nash solution, X takes on the value that maximizes this function. Elementary calculus tells us that the way to find X is to take the derivative of this expression with respect to X and set it equal to zero. Using the rules of calculus for taking the derivatives of powers of X and of the product of two functions of X, we get

$$hX^{h-1}(S - X)^k - X^hk(S - X)^{k-1} = 0.$$

[4] One and only one of the (convex) level curves can be tangential to the (concave) efficient frontier; in Figure 17.2, this level curve is labeled c_2. All lower-level curves (such as c_1) cut the frontier in two points; all higher-level curves (such as c_3) do not meet the frontier at all.

[5] If you have taken an elementary microeconomics course, you will have encountered the concept of social optimality, illustrated graphically by the tangent point between the production possibility frontier of an economy and a social indifference curve. Our Figure 17.2 is similar in spirit; the efficient frontier in bargaining is like the production possibility frontier, and the level curves of the objective in cooperative bargaining are like social indifference curves.

When we cancel the common factor $X^{h-1}(S - X)^{k-1}$, this equation becomes

$$h(S - X) - kX = 0$$
$$hY - kX = 0$$
$$kX = hY$$
$$\frac{X}{h} = \frac{Y}{k}.$$

Finally, expressing the equation in terms of the original variables x and y, we have $(x - a)/h = (y - b)/k$, which is just the Nash formula. The punch line: Nash's three conditions lead to the formula we originally stated as a simple way of splitting the bargaining surplus.

The three principles, or desired properties, that determine the Nash cooperative-bargaining solution are simple and even appealing. But in the absence of a good mechanism to make sure that the parties take the actions stipulated by the agreement, these principles may come to nothing. A player who can do better by strategizing on his own than by using the Nash solution may simply reject the principles. If an arbitrator can enforce a solution, the player may simply refuse to go to arbitration. Therefore, Nash's cooperative solution will seem more compelling if it can be given an alternative interpretation—namely, as the Nash equilibrium of a noncooperative game played by the bargainers. This can indeed be done, and we will develop an important special case of it in Section 5.

2 VARIABLE-THREAT BARGAINING

In this section, we embed the Nash cooperative solution within a specific game—namely, as the second stage of a sequential-play game. We assumed in Section 1 that the players' backstops (BATNAs) a and b were fixed. But suppose there is a first stage to the bargaining game in which the players can make strategic moves to manipulate their BATNAs within certain limits. After they have done so, the Nash cooperative outcome starting from those BATNAs will emerge in a second stage of the game. This type of game is called **variable-threat bargaining**. What kind of manipulation of the BATNAs is in a player's interest in this type of game?

We show the possible outcomes from a process of manipulating BATNAs in Figure 17.3. The originally given backstops (a and b) are the coordinates for the game's backstop point P; the Nash solution to a bargaining game with these backstops is at the outcome Q. If player A can increase his BATNA to move the game's backstop point to P_1, then the Nash solution starting there leads to the outcome Q', which is better for player A (and worse for B). Thus, a strategic move that improves one's own BATNA is desirable. For example, if you have a

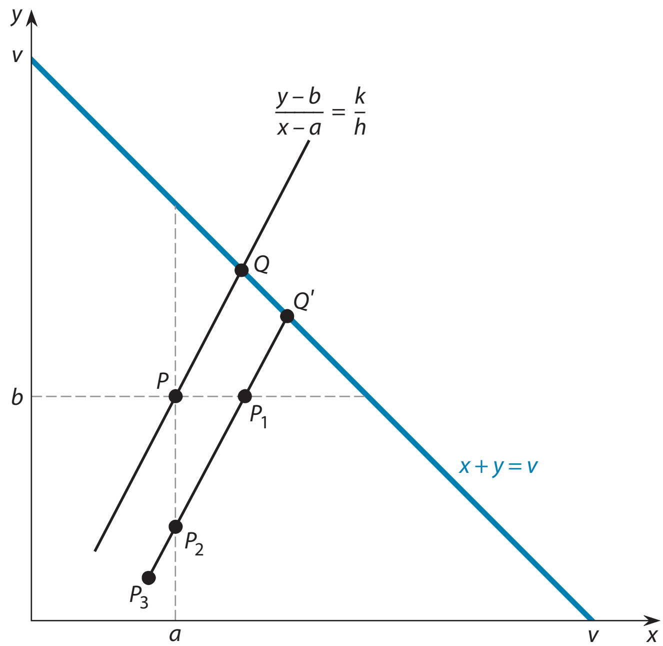

FIGURE 17.3 Bargaining Game of Manipulating BATNAs

good job offer in your pocket—a higher BATNA—when you go for an interview at another company, you are likely to get a better offer from that employer than you would if you did not have the first alternative.

The result that improving your own BATNA can improve your ultimate outcome is quite obvious, but the next step in the analysis is less so. It turns out that if player A can make a strategic move that *reduces* player B's BATNA and moves the game's backstop point to P_2, the Nash solution starting there leads to the *same* outcome Q' that was achieved after A increased his own BATNA to get to the backstop point P_1. Therefore, this alternative kind of manipulation is equally in player A's interest. As an example of decreasing your opponent's BATNA, think of a situation in which you are already working and want to get a raise. Your chances are better if you can make yourself indispensable to your employer so that without you his business has much worse prospects; his low outcome in the absence of an agreement—not offering you a raise and your leaving the firm—may make him more likely to accede to your wishes.

Finally and even more dramatically, if player A can make a strategic move that lowers *both players' BATNAs* so that the game's backstop point moves to P_3, that again has the same result as each of the preceding manipulations. This particular move is like using a threat that says, "This will hurt you more than it hurts me."

In general, the key for player A is to shift the game's BATNA point to somewhere below the line *PQ*. The farther southeast the BATNA point is moved, the better it is for player A in the eventual outcome. As is usual with threats, the idea is not actually to suffer the low payoff but merely to use its prospect as a lever to get a better outcome.

The possibility of manipulating BATNAs in this way depends on the context. We offer just one illustration. In 1980 there was a baseball players' strike. It took a very complicated form. The players struck in spring training, then resumed working (playing, really) when the regular season began in April, and went on strike again starting on Memorial Day. A strike is costly to both sides, employers and employees, but the costs differ. During spring training the players do not have salaries, but the owners make some money from vacationing spectators. At the start of the regular season, in April and May, the players get salaries but the weather is cold and the season is not yet exciting; therefore the crowds are small, and so the cost of a strike to the owners is low. The crowds start to build up from Memorial Day onward, which raises the cost of a strike to the owners, but the salaries that the players stand to lose stay the same. So we see that the two-piece strike was very cleverly designed to lower the BATNA of the owners *relative to* that of the players as much as possible.[6]

One puzzle remains: Why did the strike occur at all? According to the theory, everyone should have seen what was coming; a settlement more favorable to the players should have been reached so that the strike would have been unnecessary. A strike that actually happens is a threat that has "gone wrong." Some kind of uncertainty—asymmetric information or brinkmanship—must be responsible.

3 ALTERNATING-OFFERS MODEL I: TOTAL VALUE DECAYS

Here we move back to the more realistic noncooperative-game theory and think about the process of individual strategizing that may produce an equilibrium in a bargaining game. Our standard picture of this process is one of **alternating offers**. One player—say, A—makes an offer. The other—say, B—either accepts it or makes a counteroffer. If he does the latter, then A can either accept it or come back with another offer of his own. And so on. Thus, we have a sequential-move game and look for its rollback equilibrium.

To find a rollback equilibrium, we must start at the end and work backward. But where is the end point? Why should the process of offers and counteroffers ever terminate? Perhaps more drastic, why would it ever start? Why would

[6] See Larry DeBrock and Alvin Roth, "Strike Two: Labor-Management Negotiations in Major League Baseball," *Bell Journal of Economics*, vol. 12, no. 2 (Autumn 1981), pp. 413–25.

the two bargainers not stick to their original positions and refuse to budge? It is costly to both if they fail to agree at all, but the benefit of an agreement is likely to be smaller to the one who makes the first or the larger concession. The reason that anyone concedes must be that continuing to stand firm would cause an even greater loss of benefit. This loss takes one of two broad forms. The available pie, or surplus, may **decay** (shrink) with each offer, a possibility that we consider in this section. The alternative possibility is that time has value and **impatience** is important, and so a delayed agreement is worth less; we examine this possibility in Section 5.

Consider the following story of bargaining over a shrinking pie. A fan arrives at a professional football (or basketball) game without a ticket. He is willing to pay as much as $25 to watch each quarter of the game. He finds a scalper who states a price. If the fan is not willing to pay this price, he goes to a nearby bar to watch the first quarter on the big-screen TV. At the end of the quarter, he comes out, finds the scalper still there, and makes a counteroffer for the ticket. If the scalper does not agree, the fan goes back to the bar. He comes out again at the end of the second quarter, when the scalper makes him yet another offer. If that offer is not acceptable to the fan, he goes back into the bar, emerging at the end of the third quarter to make yet another counteroffer. The value of watching the rest of the game is declining as the quarters go by.[7]

Rollback analysis enables us to predict the outcome of this alternating-offers bargaining process. At the end of the third quarter, the fan knows that, if he does not buy the ticket then, the scalper will be left with a small piece of paper of no value. So the fan will be able to make a very small offer that, for the scalper, will still be better than nothing. Thus, on his last offer, the fan can get the ticket almost for free. Backing up one period, we see that, at the end of the second quarter, the scalper has the initiative in making the offer. But he must look ahead and recognize that he cannot hope to extract the whole of the remaining two quarters' value from the fan. If the scalper asks for more than $25—the value of the *third* quarter to the fan—the fan will turn down the offer because he knows that he can get the fourth quarter later for almost nothing, so the scalper can ask for $25 at most. Now consider the situation at the end of the first quarter. The fan knows that if he does not buy the ticket now, the scalper can expect to get only $25 later, and so $25 is all that the fan needs to offer now to secure the ticket. Finally, before the game even begins, the scalper can look ahead and ask for $50; this $50 includes the $25 value of the *first* quarter to the fan plus the $25 for which the fan can get the remaining three quarters' worth. Thus, the two will strike an immediate

[7] Just to keep the argument simple, we imagine this process as one-on-one bargaining. Actually, there may be several fans and several scalpers, turning the situation into a *market*. You can access our supplemental chapter on interactions in markets on the textbook Web site.

agreement, and the ticket will change hands for $50, but the price is determined by the full forward-looking rollback reasoning.[8]

This story can be easily turned into a more general argument for two bargainers, A and B. Suppose A makes the first offer to split the total surplus, which we call v (in some currency—say, dollars). If B refuses the offer, the total available drops by x_1 to $(v - x_1)$; B offers a split of this amount. If A refuses B's offer, the total drops by a further amount x_2 to $(v - x_1 - x_2)$; A offers a split of this amount. This offer and counteroffer process continues until finally, say, after 10 rounds, $v - x_1 - x_2 - \cdots - x_{10} = 0$, so the game ends. As usual with sequential-play games, we begin our analysis at the end.

If the game has gone to the point where only x_{10} is left, B can make a final offer whereby he gets to keep "almost all" of the surplus, leaving a measly cent or so to A. Left with the choice of that or absolutely nothing, A should accept the offer. To avoid the finicky complexity of keeping track of tiny cents, let us call this outcome "x_{10} to B, 0 to A." We will do the same in the other (earlier) rounds.

Knowing what is going to happen in round 10, we turn to round 9. Here A is to make the offer, and $(x_9 + x_{10})$ is left. A knows that he must offer at least x_{10} to B or else B will refuse the offer and take the game to round 10, where he can get that much. Bargainer A does not want to offer any more to B. So, on round 9, A will offer a split where he keeps x_9 and leaves x_{10} to B.

Then on the round before, when $x_8 + x_9 + x_{10}$ is left, B will offer a split where he gives x_9 to A and keeps $(x_8 + x_{10})$. Working backward, on the very first round, A will offer a split where he keeps $(x_1 + x_3 + x_5 + x_7 + x_9)$ and gives $(x_2 + x_4 + x_6 + x_8 + x_{10})$ to B. This offer will be accepted.

You can remember these formulas by means of a simple trick. *Hypothesize* a sequence in which all offers are refused. (This sequence is *not* what actually happens.) Then add up the amounts that would be destroyed by the refusals of one player. This total is what the other player gets in the actual equilibrium. For example, when B refused A's first offer, the total available surplus dropped by x_1, and x_1 became part of what went to A in the equilibrium of the game.

If each player has a positive BATNA, the analysis must be modified somewhat to take them into account. At the last round, B must offer A at least the BATNA a. If x_{10} is greater than a, B is left with $(x_{10} - a)$; if not, the game must terminate before this round is reached. Now at round 9, A must offer B the larger of the two amounts—the $(x_{10} - a)$ that B can get in round 10 or the BATNA b that B can get outside this agreement. The analysis can proceed all the way back to round 1 in this way; we leave it to you to complete the rollback reasoning for this case.

[8] To keep the analysis simple, we omitted the possibility that the game might get exciting, and so the value of the ticket might actually increase as the quarters go by. The uncertainty makes the problem much more complex but also more interesting. The ability to deal with such possibilities should inspire you to go beyond this book or course to study more advanced game theory.

We have found the rollback equilibrium of the alternating-offers bargaining game, and in the process of deriving the outcome, we have also described the full strategies (complete contingent plans of action) behind the equilibrium—namely, what each player *would* do if the game reached some later stage. In fact, actual agreement is immediate on the very first offer. The later stages are not reached; they are off-equilibrium nodes and paths. But as usual with rollback reasoning, the foresight about what rational players would do at those nodes if they were reached is what informs the initial action.

The other important point to note is that *gradual decay* (several potential rounds of offers) leads to a more even or fairer split of the total than does *sudden decay* (only one round of bargaining permitted). In the latter, no agreement would result if B turned down A's very first offer; then, in a rollback equilibrium, A would get to keep (almost) the whole surplus, giving B an "ultimatum" to accept a measly cent or else get nothing at all. The subsequent rounds give B the credible ability to refuse a very uneven first offer.

4 EXPERIMENTAL EVIDENCE

The theory of this particular type of bargaining process is fairly simple, and many people have staged laboratory or classroom experiments that create such conditions of decaying totals, to observe what the experimental subjects actually do. We mentioned some of them briefly in Chapter 3 when considering the validity of rollback reasoning; now we examine them in more detail in the context of bargaining.[9]

The simplest bargaining experiment is the **ultimatum game**, in which there is only one round: player A makes an offer and, if B does not accept it, the bargaining ends and both get nothing. The general structure of these games is as follows. A pool of players is brought together, either in the same room or at computer terminals in a network. They are paired; one person in the pair is designated to be the *proposer* (the one who makes the offer or is the seller who posts a price) and the other to be the *chooser* (the one who accepts or refuses the offer or is the customer who decides whether to buy at that price). The pair is given a fixed surplus, usually $1 or some other sum of money, to split.

Rollback reasoning suggests that A should offer B the minimal unit—say, 1 cent out of a dollar—and that B should accept such an offer. Actual results are dramatically different. In the case in which the subjects are together in a

[9] For more details, see Douglas D. Davis and Charles A. Holt, *Experimental Economics* (Princeton: Princeton University Press, 1993), pp. 263–69, and *The Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth (Princeton: Princeton University Press, 1995), pp. 255–74.

room and the assignment of the role of proposer is made randomly, the most common offer is a 50 : 50 split. Very few offers worse than 75 : 25 are made (with the proposer to keep 75% and the chooser to get 25%), and if made, they are often rejected.

This finding can be interpreted in one of two ways. Either the players cannot or do not perform the calculation required for rollback or the payoffs of the players include something other than what they get out of this round of bargaining. Surely the calculation in the ultimatum game is simple enough that anyone should be able to do it, and the subjects in most of these experiments are college students. A more likely explanation is the one that we put forth in Chapter 3, Section 6, and Chapter 5, Section 3—that the theory, which assumed payoffs to consist only of the sum earned in this one round of bargaining, is too simplistic.

Participants can have payoffs that include other things. They may have self-esteem or pride that prevents them from accepting a very unequal split. Even if the proposer A does not include this consideration in his own payoff, if he thinks that B might, then it is a good strategy for A to offer enough to make it likely that B will accept. Proposer A balances his higher payoff with a smaller offer to B against the risk of getting nothing if B rejects an offer deemed too unequal.

A second possibility is that, when the participants in the experiment are gathered in a room, the anonymity of pairing cannot be guaranteed. If the participants come from a group such as college students or villagers who have ongoing relationships outside this game, they may value those relationships. Then the proposers fear that, if they offer too unequal a split in this game, those relationships may suffer. Therefore, they would be more generous in their offers than the simplistic theory suggests. If this is the explanation, then ensuring greater anonymity should enable the proposers to make more unequal offers, and experiments do find this to be the case.

Finally, people may have a sense of fairness drilled into them during their nurture and education. This sense of fairness may have evolutionary value for society as a whole and may therefore have become a social norm. Whatever its origin, it may lead the proposers to be relatively generous in their offers, quite irrespective of the fear of rejection. One of us (Skeath) has conducted classroom experiments of several different ultimatum games. Students who had partners previously known to them with whom to bargain were noticeably “fairer” in their split of the pie. In addition, several students cited specific cultural backgrounds as explanations for behavior that was inconsistent with theoretical predictions.

Experimenters have tried variants of the basic game to differentiate between these explanations. The point about ongoing relationships can be handled by stricter procedures that visibly guarantee anonymity. Doing so by itself has some effect on the outcomes but still does not produce offers as extreme as those predicted by the purely selfish rollback argument of the theory. The remaining explanations—namely, "fear of rejection" and the "ingrained sense of fairness"—remain to be sorted out.

The fear of rejection can be removed by considering a variant called the *dictator game*. Again, the participants are matched in pairs. One person (say, A) is designated to determine the split, and the other (say, B) is simply a passive recipient of what A decides. Now the split becomes decidedly more uneven, but even here a majority of As choose to keep no more than 70%. This result suggests a role for an ingrained sense of fairness.

But such a sense has its limits. In some experiments, a sense of fairness was created simply when the experimenter randomly assigned roles of proposer and chooser. In one variant, the participants were given a simple quiz, and those who performed best were made proposers. This created a sense that the role of proposer had been earned, and the outcomes did show more unequal splits. When the dictator game was played with earned rights and with stricter anonymity conditions, most As kept everything, but some (about 5%) still offered a 50 : 50 split.

One of us (Dixit) carried out a classroom experiment in which students in groups of 20 were gathered together in a computer cluster. They were matched randomly and anonymously in pairs, and each pair tried to agree on how to split 100 points. Roles of proposer and chooser were not assigned; either could make an offer or accept the other's offer. Offers could be made and changed at any time. The pairs could exchange messages instantly with their matched opponent on their computer screens. The bargaining round ended at a random time between 3 and 5 minutes; if agreement was not reached in time by a pair, both got zero. There were 10 such rounds with different random opponents each time. Thus, the game itself offered no scope for cooperation through repetition. In a classroom context, the students had ongoing relationships outside the game, but they did not generally know or guess with whom they were playing in any round, even though no great attempt was made to enforce anonymity. Each student's score for the whole game was the sum of his point score for the 10 rounds. The stakes were quite high, because the score accounted for 5% of the course grade!

The highest total of points achieved was 515. Those who quickly agreed on 50 : 50 splits did the best, and those who tried to hold out for very uneven scores or who refused to split a difference of 10 points or so between the offers and ran the risk of time running out on them did poorly.[10] It seems that moderation and fairness do get rewarded, even as measured in terms of one's own payoff.

[10] Those who were best at the mathematical aspects of game theory, such as problem sets, did a little worse than the average, probably because they tried too hard to eke out an extra advantage and met resistance. And women did slightly better than men.

5 ALTERNATING-OFFERS MODEL II: IMPATIENCE

Now we consider a different kind of cost of delay in reaching an agreement. Suppose the actual monetary value of the total available for splitting does not decay, but players have a "time value of money" and therefore prefer early agreement to later agreement. They make offers alternately as described in Section 3, but their time preferences are such that money now is better than money later. For concreteness, we will say that both bargainers believe that having only 95 cents right now is as good as having \$1 one round later.

A player who prefers having something right away to having the same thing later is impatient; he attaches less importance to the future relative to the present. We came across this idea in Chapter 10, Section 2, and saw two reasons for it. First, player A may be able to invest his money—say, \$1—now and get his principal back along with interest and capital gains at a rate of return r, for a total of $(1 + r)$ in the next period (tomorrow, next week, next year, or whatever is the length of the period). Second, there may be some risk that the game will end between now and the next offer (as in the sudden end at a time between 3 and 5 minutes in the classroom game described earlier). If p is the probability that the game continues, then the chance of getting a dollar next period has an expected value of only p now.

Suppose we consider a bargaining process between two players with zero BATNAs. Let us start the process with one of the two bargainers—say, A—making an offer to split \$1. If the other player, B, rejects A's offer, then B will have an opportunity to make his own offer one round later. The two bargainers are in identical situations when each makes his offer, because the amount to be split is always \$1. Thus, in equilibrium the amount that goes to the person currently in charge of making the offer (call it x) is the same, regardless of whether that person is A or B. We can use rollback reasoning to find an equation that can be solved for x.

Suppose A starts the alternating offer process. He knows that B can get x in the next round when it is B's turn to make the offer. Therefore, A must give B at least an amount that is equivalent, in B's eyes, to getting x in the next round; A must give B at least $0.95x$ now. (Remember that, for B, 95 cents received now is equivalent to \$1 received in the next round; so $0.95x$ now is as good as x in the next round.) Player A will not give B any more than is required to induce B's acceptance. Thus, A offers B exactly $0.95x$ and is left with $(1 - 0.95x)$. But the amount that A gets when making the offer is just what we called x. Therefore, $x = 1 - 0.95x$, or $(1 + 0.95)x = 1$, or $x = 1/1.95 = 0.512$.

Two things about this calculation should be noted. First, even though the process allows for an unlimited sequence of alternating offers and counteroffers, in the equilibrium the very first offer A makes gets accepted and the bargaining ends. Because time has value, this outcome is efficient. The cost of delay governs how much A must offer B to induce acceptance; it thus affects A's rollback reasoning. Second, the player who makes the first offer gets more than half of the pie—namely, 0.512 rather than 0.488. Thus, each player gets more when he makes the first offer than when the other player makes the first offer. But this advantage is far smaller than that in an ultimatum game with no future rounds of counteroffers.

Now suppose the two players are not equally patient (or impatient, as the case may be). Player B still regards \$1 in the next round as being equivalent to 95 cents now, but A regards it as being equivalent to only 90 cents now. Thus, A is willing to accept a smaller amount to get it sooner; in other words, A is more impatient. This inequality in rates of impatience can translate into unequal equilibrium payoffs from the bargaining process. To find the equilibrium for this example, we write x for the amount that A gets when he starts the process and y for what B gets when he starts the process.

Player A knows that he must give B at least $0.95y$; otherwise B will reject the offer in favor of the y that he knows he can get when it becomes his turn to make the offer. Thus, the amount that A gets, x, must be $1 - 0.95y$; $x = 1 - 0.95y$. Similarly, when B starts the process, he knows that he must offer A at least $0.90x$, and then $y = 1 - 0.90x$. These two equations can be solved for x and y:

$$\begin{aligned} x &= 1 - 0.95(1 - 0.9x) \\ [1 - 0.95(0.9)]x &= 1 - 0.95 \\ 0.145x &= 0.05 \\ x &= 0.345 \end{aligned} \qquad \text{and} \qquad \begin{aligned} y &= 1 - 0.9(1 - 0.95y) \\ [1 - 0.9(0.95)]y &= 1 - 0.9 \\ 0.145y &= 0.10 \\ y &= 0.690 \end{aligned}$$

Note that x and y do not add up to 1, because each of these amounts is the payoff to a given player when he makes the first offer. Thus, when A makes the first offer, A gets 0.345 and B gets 0.655; when B makes the first offer, B gets 0.69 and A gets 0.31. Once again, each player does better when he makes the first offer than when the other player makes the first offer, and once again the difference is small.

The outcome of this case with unequal rates of impatience differs from that of the preceding case with equal rates of impatience in a major way. With unequal rates of impatience, the more impatient player, A, gets a lot less than B even when he is able to make the first offer. We expect that the person who is willing to accept less to get it sooner ends up getting less, but the difference is very dramatic. The proportion of A's shares and B's shares is almost 1 : 2.

As usual, we can now build on these examples to develop the more general algebra. Suppose A regards \$1 immediately as being equivalent to \$$(1 + r)$ one offer later or, equivalently, A regards \$$1/(1 + r)$ immediately as being equivalent to \$1 one offer later. For brevity, we substitute a for $1/(1 + r)$ in the calculations that follow. Likewise, suppose player B regards \$1 today as being equivalent to \$$(1 + s)$ one offer later; we use b for $1/(1 + s)$. If r is high (or equivalently, if a is low), then player A is very impatient. Similarly, B is impatient if s is high (or if b is low).

Here we look at bargaining that takes place in alternating rounds, with a total of \$1 to be divided between two players, both of whom have zero BATNAs. (You can do the even more general case easily once you understand this one.) What is the rollback equilibrium?

We can find the payoffs in such an equilibrium by extending the simple argument used earlier. Suppose A's payoff in the rollback equilibrium is x when he makes the first offer; B's payoff in the rollback equilibrium is y when he makes the first offer. We look for a pair of equations linking the values x and y and then solve these equations to determine the equilibrium payoffs.[11]

When A is making the offer, he knows that he must give B an amount that B regards as being equivalent to y one period later. This amount is $by = y/(1 + s)$. Then, after making the offer to B, A can keep only what is left: $x = 1 - by$.

Similarly, when B is making the offer, he must give A the equivalent of x one period later—namely, ax. Therefore $y = 1 - ax$. Solving these two equations is now a simple matter. We have $x = 1 - b(1 - ax)$, or $(1 - ab)x = 1 - b$. Expressed in terms of r and s, this equation becomes

$$x = \frac{1-b}{1-ab} = \frac{s+rs}{r+s+rs}.$$

Similarly, $y = 1 - a(1 - by)$, or $(1 - ab)y = 1 - a$. This equation becomes

$$y = \frac{1-a}{1-ab} = \frac{r+rs}{r+s+rs}.$$

Although this quick solution might seem a sleight of hand, it follows the same steps used earlier, and we soon give a different reasoning yielding exactly the same answer. First, let us examine some features of the answer.

First note that, as in our simple unequal-impatience example, the two magnitudes x and y add up to more than 1:

$$x + y = \frac{r+s+2rs}{r+s+rs} > 1.$$

Remember that x is what A gets when he has the right to make the first offer, and y is what B gets when he has the right to make the first offer. When A makes the first offer, B gets $(1 - x)$, which is less than y; this just shows A's advantage from

[11] We are taking a shortcut; we have simply assumed that such an equilibrium exists and that the payoffs are uniquely determined. More rigorous theory proves these conditions. For a step in this direction, see John Sutton, "Non-Cooperative Bargaining: An Introduction," *Review of Economic Studies*, vol. 53, no. 5 (October 1986), pp. 709–24. The fully rigorous (and quite difficult) theory is given in Ariel Rubinstein, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, vol. 50, no. 1 (January 1982), pp. 97–109.

being the first proposer. Similarly, when B makes the first offer, B gets y and A gets $(1 - y)$, which is less than x.

However, usually r and s are small numbers. When offers can be made at short intervals such as a week or a day or an hour, the interest that your money can earn from one offer to the next or the probability that the game ends precisely within the next interval is quite small. For example, if r is 1% (0.01) and s is 2% (0.02), then the formulas yield $x = 0.668$ and $y = 0.337$; so the advantage of making the first offer is only 0.005. (A gets 0.668 when making the first offer, but $1 - 0.337 = 0.663$ when B makes the first offer; the difference is only 0.005.) More formally, when r and s are each small compared with 1, then their product rs is very small indeed; thus we can ignore rs to write an approximate solution for the split that does not depend on which player makes the first offer:

$$x = \frac{s}{r+s} \qquad \text{and} \qquad y = \frac{r}{r+s}\,.$$

Now $x + y$ is approximately equal to 1.

Most important, x and y in the approximate solution are the shares of the surplus that go to the two players, and $y/x = r/s$; that is, the shares of the players are inversely proportional to their rates of impatience as measured by r and s. If B is twice as impatient as A, then A gets twice as much as B; so the shares are 1/3 and 2/3, or 0.333 and 0.667, respectively. Thus, we see that patience is an important advantage in bargaining. Our formal analysis supports the intuition that, if you are very impatient, the other player can offer you a quick but poor deal, knowing that you will accept it.

This effect of impatience hurts the United States in numerous negotiations that our government agencies and diplomats conduct with other countries. The American political process puts a great premium on speed. The media, interest groups, and rival politicians all demand results and are quick to criticize the administration or the diplomats for any delay. Under this pressure to deliver, the negotiators are always tempted to come back with results of any kind. Such results are often poor from the long-run U.S. perspective; the other countries' concessions often have loopholes, and their promises are less than credible. The U.S. administration hails the deals as great victories, but they usually unravel after a few years. The financial crisis of 2008 offers another and dramatic example. When the housing boom collapsed, some major financial institutions that held mortgage-backed assets faced bankruptcy. That led them to curtail credit, which in turn threatened to throw the U.S. economy into a severe recession. The crisis exploded in September, in the midst of a presidential election campaign. The Treasury, the Federal Reserve, and political leaders in Congress all wanted to act fast. This impatience led them to offer far more generous terms of rescue to many financial insitutions, when a slower process would have yielded an outcome that cost the taxpayers much less and offered them much better prospects of sharing in future gains on the assets being rescued.

Individuals who suffer losses are in a much weaker position when they negotiate with insurance companies on coverage. The companies often make lowball offers of settlement to people who have suffered a major loss, knowing that their clients urgently want to make a fresh start and are therefore very impatient.

As a conceptual matter, the formula $y/x = r/s$ ties our noncooperative game approach to bargaining to the cooperative approach of the Nash solution discussed in Section 1. The formula for shares of the available surplus that we derived there becomes, with zero BATNAs, $y/x = k/h$. In the cooperative approach, the shares of the two players stood in the same proportions as their bargaining strengths, but these strengths were assumed to be given somehow from the outside. Now we have an explanation for the bargaining strengths in terms of some more basic characteristics for the players—h and k are inversely proportional to the players' rates of impatience r and s. In other words, Nash's cooperative solution can also be given an alternative and perhaps more satisfactory interpretation as the rollback equilibrium of a noncooperative game of offers and counteroffers, if we interpret the abstract bargaining-strength parameters in the cooperative solution correctly in terms of the players' characteristics, such as impatience.

Finally, note that agreement is once again immediate—the very first offer is accepted. As usual, the whole rollback analysis disciplines by making the first proposer recognize that the other would credibly reject a less adequate offer.

To conclude this section, we offer an alternative derivation of the same (precise) formula for the equilibrium offers that we derived earlier. Suppose this time that there are 100 rounds; A is the first proposer and B the last. Start the backward induction in the 100th round; B will keep the whole dollar. Therefore in the 99th round, A will have to offer B the equivalent of $1 in the 100th round—namely, b, and A will keep $(1 - b)$. Then proceed backward:

In round 98, B offers $a(1 - b)$ to A and keeps
$1 - a(1 - b) = 1 - a + ab$.
In round 97, A offers $b(1 - a + ab)$ to B and keeps
$1 - b(1 - a + ab) = 1 - b + ab - ab^2$.
In round 96, B offers $a(1 - b + ab - ab^2)$ to A and keeps
$1 - a + ab - a^2b + a^2b^2$.
In round 95, A offers $b(1 - a + ab - a^2b + a^2b^2)$ to B and keeps
$1 - b + ab - ab^2 + a^2b^2 - a^2b^3$.

Proceeding in this way and following the established pattern, we see that, in round 1, A gets to keep

$$\begin{aligned} &1 - b + ab - ab^2 + a^2b^2 - a^2b^3 + \cdots + a^{49}b^{49} - a^{49}b^{50} \\ &= (1 - b)[1 + ab + (ab)^2 + \cdots + (ab)^{49}] \end{aligned}$$

The consequence of allowing more and more rounds is now clear. We just get more and more of these terms, growing geometrically by the factor *ab* for every two offers. To find A's payoff when he is the first proposer in an infinitely long sequence of offers and counteroffers, we have to find the limit of the infinite geometric sum. In the appendix to Chapter 10 we saw how to sum such series. Using the formula obtained there, we get the answer

$$(1 - b)\,[1 + ab + (ab) + (ab)^2 + \cdots + (ab)^{49} + \cdots] = \frac{1 - b}{1 - ab}.$$

This is exactly the solution for x that we obtained before. By a similar argument, you can find B's payoff when he is the proposer and, in doing so, improve your understanding and technical skills at the same time.

6 MANIPULATING INFORMATION IN BARGAINING

We have seen that the outcomes of a bargain depend crucially on various characteristics of the parties to the bargain, most important their BATNAs and their impatience. We have proceeded thus far by assuming that the players knew each other's characteristics as well as their own. In fact, we have assumed that each player knew that the other knew, and so on; that is, the characteristics were common knowledge. In reality, we often engage in bargaining without knowing the other side's BATNA or degree of impatience; sometimes we do not even know our own BATNA very precisely.

As we saw in Chapter 8, a game with such uncertainty or informational asymmetry has associated with it an important game of signaling and screening of strategies for manipulating information. Bargaining is replete with such strategies. A player with a good BATNA or a high degree of patience wants to signal this fact to the other. However, because someone without these good attributes will want to imitate them, the other party will be skeptical and will examine the signals critically for their credibility. And each side will also try screening, by using strategies that induce the other to take actions that will reveal its characteristics truthfully.

In this section, we look at some such strategies used by buyers and sellers in the housing market. Most Americans are active in the housing market several times in their lives, and many people are professional real-estate agents or brokers who have even more extensive experience in the matter. Moreover, housing is one of the few markets in the United States where haggling or bargaining over price is accepted and even expected. Therefore, considerable experience of strategies is available. We draw on this experience for many of our examples and interpret it in the light of our game-theoretic ideas and insights.[12]

When you contemplate buying a house in a new neighborhood, you are unlikely to know the general range of prices for the particular type of house in which you are interested. Your first step should be to find this range so that you can then determine your BATNA. And that does not mean looking at newspaper ads or realtors' listings, which indicate only asking prices. Local newspapers and some Internet sites list recent actual transactions and the actual prices; you should check them against the asking prices of the same houses to get an idea of the state of the market and the range of bargaining that might be possible.

Next comes finding out (screening) the other side's BATNA and level of impatience. If you are a buyer, you can find out why the house is being sold and how long it has been on the market. If it is empty, why? And how long has it been that way? If the owners are getting divorced or have moved elsewhere and are financing another house on an expensive bridge loan, it is likely that they have a low BATNA or are rather impatient.

You should also find out other relevant things about the other side's preferences, even though these preferences may seem irrational to you. For example, some people consider an offer too far below the asking price an insult and will not sell at any price to someone who makes such an offer. Norms of this kind vary across regions and times. It pays to find out what the common practices are.

Most important, the *acceptance* of an offer more accurately reveals a player's true willingness to pay than anything else and therefore is open to exploitation by the other player. A brilliant game-theorist friend of ours tried just such a ploy. He was bargaining for a floor lamp. Starting with the seller's asking price of $100, the negotiation proceeded to a point where our friend made an offer to buy the lamp for $60. The seller said yes, at which point our friend thought: "This guy is willing to sell it for $60, so his true rock-bottom price must be even lower. Let me try to find out whether it is." So our friend said, "How about $55?" The seller got very upset, refused to sell for any price, and asked our friend to leave the store and never come back.

The seller's behavior conformed to the norm that it is utmost bad faith in bargaining to renege on an offer once it is accepted. It makes good sense as a norm in the whole context of all bargaining games that take place in society. If an offer on the table cannot be accepted in good faith by the other player without fear of the kind of exploitation attempted by our friend, then each bargainer will wait to get the other to accept an offer, thereby revealing the limit of his true rock-bottom acceptance level, and the whole process of bargains will grind to a

[12] We have taken the insights of practitioners from Andrée Brooks, "Honing Haggling Skills," *New York Times*, December 5, 1993.

halt. Therefore, such behavior has to be disallowed. Making it a social norm to which people adhere instinctively, as the seller in the example did, is a good way for society to achieve this aim.

The offer may explicitly say that it is open only for a specified and limited time; this stipulation can be part of the offer itself. Job offers usually specify a deadline for acceptance; stores have sales for limited periods. But in that case the offer is truly a *package* of price and time, and reneging on either dimension provokes a similar instinctive anger. For example, customers get quite angry if they arrive at a store in the sale period and find an advertised item unavailable. The store must offer a rain check, which allows the customer to buy the item at its sale price when next available at the regular price; even this offer causes some inconvenience to the customer and risks some loss of goodwill. The store can specify “limited quantities, no rain checks” very clearly in its advertising of the sale; even then, many customers get upset if they find that the store has run out of the item.

Next on our list of strategies to use in one-on-one bargaining, as in the housing market, comes signaling your own high BATNA or patience. The best way to signal patience is to *be* patient. Do not come back with counteroffers too quickly, “let the sellers think they’ve lost you.” This signal is credible because someone not genuinely patient would find it too costly to mimic the leisurely approach. Similarly, you can signal a high BATNA by starting to walk away, a tactic that is common in negotiations at bazaars in other countries and some flea markets and tag sales in the United States.

Even if your BATNA is low, you may commit yourself to not accepting an offer below a certain level. This constraint acts just like a high BATNA, because the other side cannot hope to get you to accept anything less. In the housing context, you can claim your inability to concede any further by inventing (or creating) a tightwad parent who is providing the down payment or a spouse who does not really like the house and will not let you offer any more. Sellers can try similar tactics. A parallel in wage negotiations is the *mandate.* A meeting is convened of all the workers who pass a resolution—the mandate—authorizing the union leaders to represent them at the negotiation but with the constraint that the negotiators must not accept an offer below a certain level specified in the resolution. Then, at the meeting with the management, the union leaders can say that their hands are tied; there is no time to go back to the membership to get their approval for any lower offer.

Most of these strategies entail some risk. While you are signaling patience by waiting, the seller of the house may find another willing buyer. As employer and union wait for one another to concede, tensions may mount so high that a strike that is costly to both sides nevertheless cannot be prevented. In other words, many strategies of information manipulation are instances of brinkmanship. We saw in Chapter 14 how such games can have an outcome that is bad for both parties. The same is true in bargaining. *Threats* of breakdown of negotiations or of strikes are strategic moves intended to achieve quicker agreement or a better deal for the player making the move; however, an *actual* breakdown or strike is an instance of the threat "gone wrong." The player making the threat—initiating the brinkmanship—must assess the risk and the potential rewards when deciding whether and how far to proceed down this path.

7 BARGAINING WITH MANY PARTIES AND ISSUES

Our discussion thus far has been confined to the classic situation where two parties are bargaining about the split of a given total surplus. But many real-life negotiations include several parties or several issues simultaneously. Although the games get more complicated, often the enlargement of the group or the set of issues actually makes it easier to arrive at a mutually satisfactory agreement. In this section, we take a brief look at such matters.[13]

A. Multi-Issue Bargaining

In a sense, we have already considered multi-issue bargaining. The negotiation over price between a seller and a buyer always comprises *two* things: (1) the object offered for sale or considered for purchase and (2) money. The potential for mutual benefit arises when the buyer values the object more than the seller does—that is, when the buyer is willing to give up more money in return for getting the object than the seller is willing to accept in return for giving up the object. Both players can be better off as a result of their bargaining agreement.

The same principle applies more generally. International trade is the classic example. Consider two hypothetical countries, Freedonia and Ilyria. If Freedonia can convert 1 loaf of bread into 2 bottles of wine (by using less of its resources such as labor and land in the production of bread and using them to produce more wine instead) and Ilyria can convert 1 bottle of wine into 1 loaf of bread (by switching its resources in the opposite direction), then between them they can create more goods "out of nothing." For example, suppose that Freedonia can produce 200 more bottles of wine if it produces 100 fewer loaves of bread and that Ilyria can produce 150 more loaves of bread if it produces 150 fewer bottles of wine. These switches in resource utilization create an extra 50 loaves of bread and 50 bottles of wine relative to what the two countries produced originally. This extra bread and wine is the "surplus" that they can create if they can agree

[13] For a more thorough treatment, see Howard Raiffa, *The Art and Science of Negotiation* (Cambridge, Mass.: Harvard University Press, 1982), parts III and IV.

on how to divide it between them. For example, suppose Freedonia gives 175 bottles of wine to Ilyria and gets 125 loaves of bread. Then each country will have 25 more loaves of bread and 25 more bottles of wine than it did before. But there is a whole range of possible exchanges corresponding to different divisions of the gain. At one extreme, Freedonia may give up all the 200 extra bottles of wine that it has produced in exchange for 101 loaves of bread from Ilyria, in which case Ilyria gets almost all the gain from trade. At the other extreme, Freedonia may give up only 151 bottles of wine in exchange for 150 loaves of bread from Ilyria, and so Freedonia gets almost all the gain from trade.[14] Between these limits lies the frontier where the two can bargain over the division of the gains from trade.

The general principle should now be clear. When two or more issues are on the bargaining table at the same time and the two parties are willing to trade more of one against less of the other at different rates, then a mutually beneficial deal exists. The mutual benefit can be realized by trading at a rate somewhere between the two parties' different rates of willingness to trade. The division of gains depends on the choice of the rate of trade. The closer it is to one side's willingness ratio, the less that side gains from the deal.

Now you can also see how the possibilities for mutually beneficial deals can be expanded by bringing more issues to the table at the same time. With more issues, you are more likely to find divergences in the ratios of valuation between the two parties and are thereby more likely to locate possibilities for mutual gain. In regard to a house, for example, many of the fittings or furnishings may be of little use to the seller in the new house to which he is moving, but they may be of sufficiently good fit and taste that the buyer values having them. Then if the seller cannot be induced to lower the price, he may be amenable to including these items in the original price to close the deal.

However, the expansion of issues is not an unmixed blessing. If you value something greatly, you may fear putting it on the bargaining table; you may worry that the other side will extract big concessions from you, knowing that you want to protect that one item of great value. At the worst, a new issue on the table may make it possible for one side to deploy threats that lower the other side's BATNA. For example, a country engaged in diplomatic negotiations may be vulnerable to an economic embargo; then it would much prefer to keep the political and economic issues distinct.

[14] Economics uses the concept *ratio of exchange*, or price, which here is expressed as the number of bottles of wine that trade for each loaf of bread. The crucial point is that the possibility of gain for both countries exists with any ratio that lies between the 2 : 1 at which Freedonia can just convert bread into wine and the 1 : 1 at which Ilyria can do so. At a ratio close to 2 : 1, Freedonia gives up almost all of its 200 extra bottles of wine and gets little more than the 100 loaves of bread that it sacrificed to produce the extra wine; thus Ilyria has almost all of the gain. Conversely, at a ratio close to 1 : 1, Freedonia has almost all of the gain. The issue in the bargaining is the division of gain and therefore the ratio or the price at which the two should trade.

B. Multiparty Bargaining

Having many parties simultaneously engaged in bargaining also may facilitate agreement, because instead of having to look for pairwise deals, the parties can seek a circle of concessions. International trade is again the prime example. Suppose the United States can produce wheat very efficiently but is less productive in cars, Japan is very good at producing cars but has no oil, and Saudi Arabia has a lot of oil but cannot grow wheat. In pairs, they can achieve little, but the three together have the potential for a mutually beneficial deal.

As with multiple issues, expanding the bargaining to multiple parties is not simple. In our example, the deal would be that the United States would send an agreed amount of wheat to Saudi Arabia, which would send its agreed amount of oil to Japan, which would in turn ship its agreed number of cars to the United States. But suppose that Japan reneges on its part of the deal. Saudi Arabia cannot retaliate against the United States, because, in this scenario, it is not offering anything to the United States that it can potentially withhold. Saudi Arabia can only break its deal to send oil to Japan, an important party. Thus, enforcement of multilateral agreements may be problematic. The General Agreement on Tariffs and Trade (GATT) between 1946 and 1994, as well as the World Trade Organization (WTO) since then, have indeed found it difficult to enforce their agreements and to levy punishments on countries that violate the rules.

SUMMARY

Bargaining negotiations attempt to divide the *surplus* (excess value) that is available to the parties if an agreement can be reached. Bargaining can be analyzed as a *cooperative* game in which parties find and implement a solution jointly or as a (structured) *noncooperative* game in which parties choose strategies separately and attempt to reach an equilibrium.

Nash's cooperative solution is based on three principles of the outcomes' invariance to linear changes in the payoff scale, *efficiency*, and invariance to removal of irrelevant outcomes. The solution is a rule that states the proportions of division of surplus, beyond the backstop payoff levels (also called *BATNAs* or *best alternatives to a negotiated agreement*) available to each party, based on relative bargaining strengths. Strategic manipulation of the backstops can be used to increase a party's payoff.

In a noncooperative setting of *alternating offer and counteroffer*, rollback reasoning is used to find an equilibrium; this reasoning generally includes a first-round offer that is immediately accepted. If the surplus value *decays* with refusals, the sum of the (hypothetical) amounts destroyed owing to the refusals of a single player is the payoff to the other player in equilibrium. If delay in agreement is costly owing to *impatience*, the equilibrium offer shares the surplus roughly in inverse proportion to the parties' rates of *impatience*. Experimental evidence indicates that players often offer more than is necessary to reach an agreement in such games; this behavior is thought to be related to player anonymity as well as beliefs about fairness.

The presence of information asymmetries in bargaining games makes signaling and screening important. Some parties will wish to signal their high BATNA levels or extreme patience; others will want to screen to obtain truthful revelation of such characteristics. When more issues are on the table or more parties are participating, agreements may be easier to reach, but bargaining may be riskier or the agreements more difficult to enforce.

KEY TERMS

alternating offers (674)
best alternative to a negotiated agreement (BATNA) (666)
decay (675)
efficient frontier (669)
efficient outcome (669)
impatience (675)
Nash cooperative solution (667)
surplus (666)
ultimatum game (677)
variable-threat bargaining (672)

SOLVED EXERCISES

S1. Consider the bargaining situation between Compaq Computer Corporation and the California businessman who owned the Internet address www.altavista.com.[15] Compaq, which had recently taken over Digital Equipment Corporation, wanted to use this man's Web address for Digital's Internet search engine, which at that time had the address www.altavista.digital.com. Compaq and the businessman apparently negotiated long and hard during the summer of 1998 over a selling price for the latter's address.

Although the businessman was the "smaller" player in this game, the final agreement appeared to entail a \$3.35 *million* price tag for the Web address in question. Compaq confirmed the purchase in August and began using the address in September but refused to divulge any of the financial details of the settlement. Given this information, comment on

[15] Details regarding this bargaining game were reported in "A Web Site by Any Other Name Would Probably Be Cheaper," *Boston Globe*, July 29, 1998, and in Hiawatha Bray's "Compaq Acknowledges Purchase of Web Site," *Boston Globe*, August 12, 1998.

the likely values of the BATNAs for these two players, their bargaining strengths or levels of impatience, and whether a cooperative outcome appears to have been attained in this game.

S2. Ali and Baba are bargaining to split a total that starts out at $100. Ali makes the first offer, stating how the $100 will be divided between them. If Baba accepts this offer, the game is over. If Baba rejects it, a dollar is withdrawn from the total, so it is now only $99. Then Baba gets the second turn to make an offer of a division. The turns alternate in this way, a dollar being removed from the total after each rejection. Ali's BATNA is $2.25 and Baba's BATNA is $3.50. What is the rollback-equilibrium outcome of the game?

S3. Two hypothetical countries, Euphoria and Militia, are holding negotiations to settle a dispute. They meet once a month, starting in January, and take turns making offers. Suppose the total at stake is 100 points. The government of Euphoria is facing reelection in November. Unless the government produces an agreement at the October meeting, it will lose the election, which it regards as being just as bad as getting zero points from an agreement. The government of Militia does not really care about reaching an agreement; it is just as happy to prolong the negotiations or even to fight, because it would be settling for anything significantly less than 100.

(a) What will be the outcome of the negotiations? What difference will the identity of the first mover make?

(b) In light of your answer to part (a), discuss why actual negotiations often continue right down to the deadline.

UNSOLVED EXERCISES

U1. Recall the variant of the pizza pricing game in Exercise U2, part (b), in Chapter 10, in which one store (Donna's Deep Dish) was much larger than the other (Pierce's Pizza Pies). The payoff table for that version of the game is:

		PIERCE'S PIZZA PIES	
		High	Low
DONNA'S DEEP DISH	High	156, 60	132, 70
	Low	150, 36	130, 50

The noncooperative dominant-strategy equilibrium is (High, Low), yielding profits of 132 to Donna's and 70 to Pierce's, for a total of 202. If the two could achieve (High, High), their total profit would be 156 + 60 = 216, but Pierce's would not agree to this pricing.

Suppose the two stores can reach an enforceable agreement whereby both charge High and Donna's pays Pierce's a sum of money. The alternative to this agreement is simply the noncooperative dominant-strategy equilibrium. They bargain over this agreement, and Donna's has 2.5 times as much bargaining power as Pierce's. In the resulting agreement, what sum will Donna's pay to Pierce's?

U2. Consider two players who bargain over a surplus initially equal to a whole-number amount V, using alternating offers. That is, Player 1 makes an offer in round 1; if Player 2 rejects this offer, she makes an offer in round 2; if Player 1 rejects this offer, she makes an offer in round 3; and so on. Suppose that the available surplus decays by a constant value of $c = 1$ each period. For example, if the players reach agreement in round 2, they divide a surplus of $V - 1$; if they reach agreement in round 5, they divide a surplus of $V - 4$. This means that the game will be over after V rounds, because at that point there will be nothing left to bargain over. (For comparison, remember the football-ticket example, in which the value of the ticket to the fan started at $100 and declined by $25 per quarter over the four quarters of the game.) In this problem, we will first solve for the rollback equilibrium to this game, and then solve for the equilibrium to a generalized version of this game in which the two players can have BATNAs.

(a) Let's start with a simple version. What is the rollback equilibrium when $V = 4$? In which period will they reach agreement? What payoff x will Player 1 receive, and what payoff y will Player 2 receive?

(b) What is the rollback equilibrium when $V = 5$?

(c) What is the rollback equilibrium when $V = 10$?

(d) What is the rollback equilibrium when $V = 11$?

(e) Now we're ready to generalize. What is the rollback equilibrium for any whole-number value of V? (Hint: You may want to consider even values of V separately from odd values.)

Now consider BATNAs. Suppose that if no agreement is reached by the end of round V, Player A gets a payoff of a and Player B gets a payoff of b. Assume that a and b are whole numbers satisfying the inequality $a + b < V$, so that the players can get higher payoffs from reaching agreement than they can by not reaching agreement.

(f) Suppose that $V = 4$. What is the rollback equilibrium for any possible values of a and b? (Hint: You may need to write down more than one formula, just as you did in part (e). If you get stuck, try assuming specific values for a and b, and then change those values to see what happens. In order to roll back, you'll need to figure out the turn at which the value of V has declined to the point where a negotiated agreement would no longer be profitable for the two bargainers.)

(g) Suppose that $V = 5$. What is the rollback equilibrium for any possible values of a and b?

(h) For any whole-number values of a, b, and V, what is the rollback equilibrium?

(i) Relax the assumption that a, b, and V are whole numbers: let them be any nonnegative numbers such that $a + b < V$. Also relax the assumption that the value of V decays by exactly 1 each period: let the value decay each period by some constant amount $c > 0$. What is the rollback equilibrium to this general problem?

U3. Let x be the amount that player A asks for, and let y be the amount that B asks for, when making the first offer in an alternating-offers bargaining game with impatience. Their rates of impatience are r and s, respectively.

(a) If we use the approximate formulas $x = s/(r + s)$ for x and $y = r/(r + s)$ for y, and if B is twice as impatient as A, then A gets two-thirds of the surplus and B gets one-third. Verify that this result is correct.

(b) Let $r = 0.01$ and $s = 0.02$, and compare the x and y values found by using the approximation method with the more exact solutions for x and y found by using the formulas $x = (s + rs)/(r + s + rs)$ and $y = (r + rs)/(r + s + rs)$ derived in the chapter.