

Getting to Stalemate

How Much Is Enough?

Nuclear weapons are uniquely effective tools of deterrence because they tend to produce stalemate, the condition in which military victory is impossible. If nuclear-armed countries cannot be conquered, and are largely immune to any major military attacks, then they have little reason to worry about their security. The strategies that have defined international politics for centuries—for example, engaging in arms races, building alliances, and securing strategic territory—no longer seem relevant for countries locked in nuclear stalemate.

But behind this straightforward logic are important theoretical and empirical questions. For example, how much nuclear capability must countries build to create stalemate? Are even small and potentially vulnerable arsenals enough to reliably deter? That is, will the mere possibility of nuclear retaliation, with even a few nuclear weapons, be enough to paralyze potential attackers? Or are large and truly survivable forces required to deter enemies, including highly aggressive ones? In short, when it comes to building a nuclear arsenal for deterrence, how much is enough? The answers to these questions have important implications for international politics in the nuclear age.

If even small, potentially vulnerable arsenals are enough to create stalemate, then nuclear weapons should dampen a wide range of competitive behavior, just as many analysts expect. After all, in a world governed by the logic of “existential” or “minimum” deterrence, once a country acquires a few nuclear weapons it can relax—because even aggressive enemies will be deterred by the mere possibility of nuclear retaliation. In such a world there is no need for arms racing; once a country has a small nuclear arsenal, the race is over. There should also be little concern about relative gains (that is, shifts in the economic balance of power), because in a world of nuclear stalemate advantages in wealth cannot be easily translated into useful military advantage. The value of allies will be diminished, too, since pooling resources is unnecessary when countries have no need to fear attack.

On the other hand, if creating stalemate requires that nuclear-armed countries build truly survivable arsenals—that is, capable of confronting adversaries with “assured retaliation”—then the nuclear age will remain highly competitive. In this case, the process of establishing stalemate will be drawn-out and dangerous. New nuclear states will race to construct truly survivable arsenals; their adversaries will be tempted to enhance counterforce capabilities to delay the onset of stalemate, and those adversaries may be tempted to launch preventive disarming attacks. New nuclear states will seek alliances to share costs

or provide protection during their period of nuclear vulnerability. In most cases, nuclear armed countries will be able to build truly survivable arsenals if they invest time and resources, but the path to stalemate will be marked by arms racing, relative-gains concerns, and alliance competition.

Many studies conducted during the Cold War sought to answer the question, how much is enough? This chapter differs in two important ways. First, whereas most previous studies explored the effect of U.S. nuclear capabilities on Soviet behavior,

we flip the analysis and track the impact of Soviet capabilities on U.S. behavior. Our approach is advantageous because it was the Soviet Union that was playing catch-up for the first fifteen years of the nuclear age, fielding an existential deterrent in the early 1950s, a minimum deterrent in the late 1950s, and an assured retaliatory capability starting in the early 1960s. In order to determine how much is enough, one needs to track how the growth of the Soviet arsenal through those different phases affected U.S. behavior. That is the task we undertake here.

This chapter diverges from other studies in another way. Whereas many nuclear analyses look at the deterrent effects of nuclear weapons in peacetime, we are interested in how they deter aggressive actions in both peacetime and war. In peacetime, countries wish to deter adversaries from initiating conflict or pursuing aggressive foreign policies. In wartime, countries hope

to constrain adversary military operations, by compelling the enemy to seek limited objectives and restricting the scope of adversary military strikes, for example. A nuclear arsenal that reduced the likelihood of conflict but also invited a devastating disarming strike if war occurred might not be enough to allow a nuclear power to feel secure, and thus might invite arms racing and other competitive behaviors. Therefore, the key question in this chapter is, how much nuclear capability is enough to satisfy both peacetime and wartime deterrent needs? Do small, simple arsenals meet those objectives? Or is a larger, more sophisticated arsenal required?

Our analysis leads to two principal findings. First, even a small and vulnerable Soviet nuclear arsenal likely bolstered peacetime deterrence during the first decade of the Cold War. Even when the Soviet nuclear force was highly vulnerable to a U.S. disarming strike during the 1950s, the United States demonstrated notable restraint toward its rival. U.S. and Soviet leaders continued to frame their competition in Manichean terms—as a clash between pure good and evil—but the United States took pains to conduct its anti-Soviet policies in a manner designed to avoid war. This foreign policy approach had many

3

causes, but the Soviet possession of nuclear weapons likely reinforced that cautious orientation.

The second finding is crucial: the Soviet Union's initial deployment of a nuclear arsenal did not make it safe and secure. In fact, by fielding a vulnerable nuclear force, the Soviets energized U.S. efforts to build weapons and war plans to rapidly destroy the Soviet arsenal if war erupted. The small, vulnerable Soviet nuclear force was a double-edged sword: it benefited the Soviets in peacetime by making U.S. aggression less likely, but it vastly increased the damage the Soviets would suffer in the event of war. The United States remained committed to winning World War III through a massive nuclear disarming strike as soon as war erupted, at least until the Soviets developed a true assured retaliation force.

The findings in this chapter help explain the general puzzle of why international competition remains so intense in the nuclear age. A small, vulnerable nuclear force did not give the Soviet Union adequate protection, which explains why the Soviets continued to build up their nuclear capabilities. The United States valued its ability to win a nuclear war if conflict with the Soviets erupted, and therefore worked hard to retain its disarming capabilities as long as possible. The result was an intense arms race, concerns over relative gains, and competition for key allies and the use of their territory for staging bases for nuclear strikes. The first decades of the nuclear age saw intense competition because small nuclear forces were not enough to create real stalemate.

How Much Is Enough? Four Views

How much nuclear capability do states need to deter aggressive behavior during peacetime and limit enemy military operations during war? If even small and potentially vulnerable arsenals are enough to reliably deter, then the nuclear age should be characterized by greatly reduced security competition. In such a world, the presence of even a few nuclear weapons will thwart ambitions and assuage fears—two primary causes of conflict in anarchy. Countries could build modest nuclear forces and then relax. But if small and vulnerable arsenals do not make countries truly secure—because risk acceptant enemies will initiate conflicts or attempt disarming strikes during wars—then competition should remain intense in the nuclear age. If large and truly secure arsenals are required for robust deterrence, the development of nuclear weapons will trigger extended arms races and perhaps preventive wars.

Scholars have developed four major schools of thought about the requirements of deterrence. One view, that of *existential deterrence*, contends that fear of nuclear war is so great that the mere possibility of nuclear retaliation is enough to deter any act that could trigger escalation. The implication of this view is that a nuclear arsenal need not be large, survivable, or even easily deliverable to terrify and thus deter; its mere existence is enough. The *minimum deterrence* school argues that states must do more than join the nuclear club to establish robust deterrence, but not much more. Deterrence will be robust whenever retaliation in the wake of an attack is plausible, instead of merely possible. To create the plausibility of retaliation, a country needs to be capable of delivering nuclear weapons against an adversary. A small deliverable nuclear arsenal—even if it is not certain to survive a disarming strike—will still deter attack because the chance of retaliation will be sufficiently terrifying.

The *assured retaliation* school is rooted in a more pessimistic view about potential aggressors, and judges a deterrent capability against a higher bar than either of the preceding schools. This school views deterrence as robust only when the defender's ability to retaliate after suffering a disarming strike is nearly certain. The threat of retaliation need not be society ending, however; the prospect of absorbing "just a few" nuclear strikes will deter any aggressor. The fourth school, *assured*

4

destruction, contends that truly robust deterrence requires that retaliation be both assured and massive.

These four views lie along a continuum, without concrete boundaries between them. But each school is based upon a distinct understanding of the requirements of deterrence, and each posits a different level of nuclear capability required for a robust deterrent. We next examine each school's basic claims in greater detail, the causal logic behind those claims, and the kind of retaliatory posture indicative of each approach. Then the remainder of the chapter evaluates these views against the historical record from the Cold War.

The first school of thought contends that the mere existence of nuclear weapons is enough to reliably deter attack. “Everything about the atomic bomb,” Bernard Brodie wrote at the dawn of the nuclear age, “is overshadowed by the twin facts⁵

that it exists and its destructive power is fantastically great.” According to the existential deterrence view, details such as the relative size of a country’s nuclear arsenal compared to that of its adversary or the vulnerability of a country’s arsenal to attack are essentially irrelevant for generating deterrence. What matters instead is fear—the brooding shadow of destruction cast by the mere presence of these weapons. In this sense, deterrence is “a simple philosophical consequence of the existence of⁶⁷

thermonuclear bombs.” Nuclear weapons “have only to exist and deterrence is the law of their existence.”

McGeorge Bundy, perhaps the most famous public figure associated with this school of thought, wrote: “The terrible and unavoidable uncertainties in any recourse to nuclear war create what could be called ‘existential’ deterrence, where the⁸

function of the adjective is to distinguish this phenomenon from anything based on strategic theories or declared policies.” Elsewhere, Bundy argued that even one chance in a hundred that a country will use nuclear weapons to retaliate is enough to⁹

deter an aggressor. Bundy noted, “We must remember that at the upper levels of force the two greatest powers have been extraordinarily cautious with each other. This is not the result of estimates of each other’s first- or second-strike counter-force capability, or a consequence of the possession or absence of escalation dominance.... They do not dare get close to war with¹⁰

each other because of their fear of what would happen if it turned nuclear, as it always might.”

The logic of the existential deterrence view rests on the potent combination of uncertainty, fear, and the sheer destructive power of nuclear weapons. As James Lebovic writes, “The existential deterrent acquires its power from the non-rational world¹¹

of fear, psychological bias, and uncertainty and not from the rational world of deduction and mathematical precision.” Similarly, according to Hedley Bull, “A potential attacker is deterred when his leaders are in a certain state of mind. Their state of mind, even when advised that the opposing retaliatory forces can be eliminated with near certainty, is still likely to include feelings of uncertainty about weapons that have not been tried in battle.... The essential conditions of mutual deterrence are subjective or psychological, and these conditions may in principle be satisfied even in the absence of totally invulnerable¹²

retaliatory forces.”

An existential deterrent does not require much actual nuclear capability. Even a rudimentary arsenal will dissuade potential attackers, since a disarming strike must be virtually assured of success before a rational leader would undertake such a momentous gamble. Relative nuclear force levels matter little when it comes to creating a sliver of doubt about success—and generating fear. Sophisticated retaliatory delivery systems, robust command-and-control systems, and serious training for nuclear retaliatory operations are irrelevant. Simply joining the nuclear club should be sufficient, although analysts can differ on whether this entails just testing a nuclear device or actually building a few weapons. Some even argue that states with the technological know-how to readily build nuclear weapons (such as Japan today) could be considered to have an existential¹³

deterrent.

In sum, existential deterrence holds that nuclear deterrence is easy to achieve. As soon as a country develops nuclear weapons, other states will be very unlikely to initiate major attacks, even if the new nuclear power has no clear ability to deliver its warheads. And if conventional war does occur between nuclear-armed states, even a rudimentary nuclear force will ensure that the adversaries will conduct military operations with great caution. As Michael Howard aptly summarizes the existential deterrence view, “It is the prospect of nuclear war *as such*, not any calculation of the balance of probable losses or¹⁴

gains, that now deters statesmen from taking risks.”

The implications of existential deterrence theory for international politics are straightforward. If the existential deterrence view is correct, then the development of nuclear weapons should have eliminated most of the factors that have driven states throughout history to compete for security and power. For advocates of existential deterrence, the development of nuclear weapons and the creation of a robust deterrent are synonymous.

A second school of thought posits that deterrence will be robust whenever nuclear retaliation is *plausible*. Retaliation does not need to be assured in order to ward off a disarming attack. Instead, the mere chance of triggering a catastrophic retaliatory response will deter any rational potential aggressor. As Kenneth Waltz argued, “An adversary need only believe that some

15

warheads *may* survive its attack and be visited on it. That belief is not hard to create.” For advocates of minimum deterrence, a sufficient nuclear retaliatory force is simple to build, cheap to maintain, easy to hide, and mostly unaffected by adversary nuclear force levels, deployment policies, and strategies.

The minimum deterrence view is built upon a simple and powerful logic. A rational person does not need to be *certain* that an attempted disarming strike will trigger apocalyptic retaliation to conclude that such a strike is a terrible idea—the mere

16

chance of triggering a catastrophe will deter anyone rational. Waltz explains, “Contemplating war when the use of nuclear weapons is possible focuses one’s attention not on the probability of victory but on the possibility of annihilation. Because catastrophic outcomes of nuclear exchanges are easy to imagine, leaders of states will shrink in horror from initiating them.... Anyone—political leader or man in the street—can see that catastrophe lurks if events spiral out of control and nuclear

17

warheads begin to fly.” The prospect of nuclear retaliation is sufficient to deter any rational attacker, but that prospect, according to the minimum deterrence view, must be realistic, not just hypothetical. William Kaufmann differentiates minimum from existential deterrence when he writes, “The enemy must be persuaded not only that the instrument exists but also that its

18

power is operational.” As long as a country’s deterrent force makes nuclear retaliation plausible, its adversaries will be extremely reluctant to initiate war—and highly restrained in their military operations if war occurs.

Avery Goldstein explains the logic of minimum deterrence in clear, rationalist terms. “The standard of first-strike uncertainty,” he writes, “assumes that a potential attacker facing a nuclear-armed state will find almost any slippage from 100 percent certainty in successful preemption excruciatingly inhibiting.” In other words, “nuclear-armed states do not need to

19

convince a potential aggressor that retaliation is certain, or even likely, only that it is possible.” Even a small chance of retaliation will deter because of the catastrophic consequences that retaliation would bring.

Capabilities matter in the minimal deterrence view, but this school does not place much emphasis on the size of arsenals or the strategic interaction of capabilities among adversaries. A minimum deterrent force is more than a “bomb in the basement,” but less than the large and diversified arsenals deployed by the superpowers in the Cold War. A sufficient retaliatory arsenal need not be large, since “even with numbers immensely disproportionate, a small force strongly inhibits the use of a large

20

one,” and it need not be guaranteed to survive an enemy first strike. What is crucial is the plausibility that a portion of the force might survive and retaliate.

The precise dividing line between an existential deterrence and minimum deterrence posture is somewhat arbitrary. We thus distinguish between the two views in a manner that holds true to their underlying logic, yet permits coding of arsenals and testing of their competing claims. In our conceptualization, the deployment of operational weapons—with sufficient range to

21

reach a potential attacker’s territory—is enough to move beyond *existential* and qualify as a *minimum* deterrence posture.

If a minimum deterrence posture were enough to create robust deterrence, one would expect the nuclear age to be much less competitive than previous eras. If countries require only small, deliverable arsenals to deter their enemies, then moving from an initial nuclear test to a robust deterrent would be a fairly modest step. If leaders reason the way minimum deterrence theorists believe they do, then the nuclear age may see brief periods of intense competition—after new nuclear states test their initial devices and before they successfully deploy operational delivery systems. In those periods, their adversaries may contemplate preventive strikes. But those periods of anxiety will be short and competition will die down. After all, once a country has a deliverable force, its most essential security needs will be fulfilled. In a world characterized by minimum deterrence reasoning, gaining any kind of military advantage over an enemy with a deliverable nuclear arsenal provides no real strategic benefits.

ASSURED RETALIATION

The assured retaliation school concludes that deterrence is robust when retaliation is *assured*, not when it is merely possible or plausible. A state that aims to deploy a robust deterrent needs to ensure that some portion of the force would almost

certainly survive an enemy first strike. The goal is to convince potential attackers that even a flawlessly executed disarming strike, carried out under ideal circumstances, would still fail—and result in the destruction of several of the attacker’s cities.

22

Robert Jervis is one of the most prominent scholars of the assured retaliation school. He and other analysts writing in the 1970s and 1980s made the case for the sufficiency of assured retaliation because official U.S. nuclear policy reflected, in their view, an erroneous belief that only a big and sophisticated arsenal could reliably deter the Soviet Union. The “illogic” of the superpower arms race, according to Jervis, stemmed from a failure to grasp that building more and better nuclear weapons did

23

not matter for deterrence and security. He wrote: “The vulnerability of population centers in both the United States and the Soviet Union that comes with mutual second-strike capability has transformed strategy. Because a military advantage no longer assures a decisive victory, old ways of thinking are no longer appropriate. The healthy fear of devastation, which cannot

24

be exorcised short of the attainment of a first-strike capability, makes deterrence relatively easy.”

The assured retaliation school takes a more pessimistic view of international politics than either of the first two schools, noting that most major wars in history began with leaders taking huge risks. Thus, proponents of assured retaliation worry that the mere possibility of nuclear retaliation may not be enough to deter such risk-acceptant leaders. Furthermore, proponents of assured retaliation may worry that a minimum deterrent force can become less capable over time, since first-rate militaries work hard to identify enemy weaknesses. For example, adversaries spy on each other to learn about the deployments and capabilities of enemy weapon systems; they study enemy command-and-control systems; and they devise doctrines and technologies to exploit the weaknesses they discover. In a crisis, or during a conventional war, a clever war plan designed to

25

destroy the enemy’s “existential” or “minimum” nuclear force may be attractive to decision makers. To ensure robust deterrence—that is, to deter war and compel military restraint if conflict erupts—states must field nuclear retaliatory forces that cannot be destroyed in any conceivable disarming strike, regardless of the malevolence and creativity of potential enemies. In Jervis’s words, a reliable nuclear deterrent should ensure that “even an emotional, short-sighted, and dim-witted opponent

26

[would] see that to start a war would be the worst alternative.” An assured retaliatory force is required to ensure that all

27

leaders exercise restraint in times of calm or crisis, peace or war.

There is no single recipe for creating an assured retaliation arsenal. The ingredients of a survivable nuclear force typically include some combination of the following: moderate (or large) force size; diverse delivery systems; survivable basing modes; redundant command and control systems; and well-trained personnel. The survivability of an arsenal—meaning its ability to absorb a ruthless disarming strike and still retaliate—is not determined by a single feature of the force, such as a numerical threshold or the use of a particular delivery system. Instead, the condition of assured retaliation is the result of both the attributes of an arsenal and the capabilities of potential attackers. Unlike existential deterrence and minimal deterrence, which are theories of deterrence that are based on what the deterring party can do (possess nuclear weapons in the former case, and deliver them in the latter), the assured retaliation mission, by definition, depends on a calculation of both the attacking force’s capabilities and the defending force’s resilience.

If the condition of assured retaliation is required to produce stalemate, then one can begin to understand why international politics remains highly competitive in the nuclear age. We should expect that a truly survivable nuclear arsenal may require substantial time to develop, but also that a force’s survivability will be the result of a competitive process—the net result of a country’s efforts to secure its arsenal and its enemy’s ability to target it. The competitive nature of “survivability” implies that the development of nuclear weapons may invigorate competition rather than inhibit it; new nuclear states will struggle to develop survivable arsenals while their enemies press to maintain superiority and keep such arsenals vulnerable. In fact, if “assured retaliation” is a key threshold for stalemate, then arms racing, alliance competition, and relative gains concerns may not wither away even after secure forces are deployed, since each side will face incentives to gain (or regain) superiority and

28

negate the other side’s assurances.

The assured retaliation school might be correct about the capability requirements for robust deterrence. If so, that may help shed light on why international politics has remained so competitive in the nuclear age, as countries race to establish an assured retaliatory capability for themselves and are tempted to compete to deny that security to their adversaries.

ASSURED DESTRUCTION

The final school holds that retaliation must be *assured and massive* in order for deterrence to be truly robust. Potential

attackers need to face the near certainty of retaliation, and the consequences of that retaliation must be mind-bogglingly bad. Attackers are deterred when there is no prospect of meaningfully limiting the amount of damage a target country can inflict after an attacker launches a first strike. When the potential victim possesses a very large and invulnerable retaliatory capability, a nuclear attack would be tantamount to national suicide—and no remotely rational leader would invite that outcome.

The basic logic of assured destruction is that nuclear arsenals need to do more than deter typical leaders in normal circumstances. A robust nuclear posture needs to terrify and frustrate the goals of even highly risk-acceptant leaders during high-stakes crises. Major wars, after all, are never initiated by the faint of heart; they occur when leaders accept enormous risks

29

in pursuit of their goals. Neither World War I nor World War II would have occurred had leaders been afraid to risk huge numbers of military and civilian deaths to advance their geopolitical goals. To assured destruction advocates, robust deterrence requires convincing highly aggressive potential attackers that their actions will result in a much worse outcome than the loss of a few cities. Only the guaranteed and utter destruction of the attacker's population, industry, and leadership suffices against the most motivated foes. Designing a deterrent policy to deter any leader less threatening, determined, and formidable would be foolhardy.

The indicators of an assured destruction arsenal are similar to those of an assured retaliation force, but with more of everything. The best example of such a force can be found in either of the superpower arsenals in the latter half of the Cold War. The United States kept roughly two-thirds of its ballistic-missile submarine fleet at sea in normal peacetime conditions, as well as a quick-response nuclear bomber force ready to take off with little warning. Launch control officers were trained to fire silo-based missiles moments after receiving orders to do so. These redundant precautions were designed to ensure that even if the Soviet Union completely surprised the United States with a bolt-from-the-blue nuclear-disarming strike, U.S. forces would still be able to deliver vast numbers of high-yield thermonuclear weapons against the Soviet Union in retaliation.

If leaders reasoned the way that the assured destruction school predicts, then the nuclear age would be characterized by significant competition. Nuclear-armed countries would compete to build larger and more potent arsenals. Some countries would be tempted to strive for damage-limiting capabilities in the hope of escaping stalemate, while many others would justly fear such efforts by their adversaries and thus feel compelled to respond with steps to make their own arsenals more secure and more lethal. The large force requirements and dispersed basing needs of an assured destruction force may require allies around the world. If robust deterrence might require truly assured destruction capabilities, there is little reason to expect that a stable world of plentiful security would follow.

Each of the four schools described above can be used to identify the threshold of nuclear capability needed to deter a wide range of behavior, including limited conventional strikes against nuclear-armed states, major conventional invasions, bolt-from-the-blue surprise attacks, and nuclear-disarming strikes in the midst of war. In each of these circumstances, the existential and minimum deterrence schools claim that a little nuclear retaliatory capability goes a long way; even a small, vulnerable nuclear arsenal will generate great caution among enemies contemplating any of those types of attack. Scholars who identify assured retaliation or destruction as the key threshold, on the other hand, will expect the benefits of deterrence to grow substantially when a country deploys a truly survivable force.

There are reasons to be cautious about generalizing these concepts, however. Perhaps none of these schools of deterrence is correct across the board. For example, one posture might be sufficient to induce caution during peacetime, yet a higher level of capability may be necessary in darker times, such as during an intense crisis or during a conventional war. Relatedly, perhaps not all national leaders reason about nuclear weapons in the same way. For example, the policies of the Eisenhower administration in the 1950s might not be a good guide to the nuclear employment policies of other leaders around the world today.

We believe the analysis below reveals some general truths about deterrence, despite such concerns. Whatever historians may think about Eisenhower, to stick with the example, few would describe his administration as more aggressive, irrational, or insensitive to costs than those of the risk takers who initiated major wars throughout history. This is relevant because if an existential or minimum deterrent force is insufficient to deter a "normal" leader (like Eisenhower) in both peacetime and wartime, then one would expect such deterrent strategies to be inadequate for deterring highly aggressive foes.

Furthermore, unlike other areas of social science, where scholars seek to understand averages and trends (rather than extremes and outliers), the goal of nuclear deterrence is different. Nuclear arsenals are designed to deter not a median enemy but rather the most ruthless plausible adversary—and they must do so all the time, including in the darkest days of a crisis or war. Therefore, a deterrent posture that does *not* deter seemingly normal leaders is clearly inadequate.

The remainder of this chapter evaluates the four schools of deterrence in the canonical case—the Cold War superpower standoff. As the Soviet nuclear arsenal grew from an existential posture to a minimum deterrent force, and eventually to a true assured retaliation arsenal, how did the United States react? What level of Soviet nuclear capability, if any, made the United States accept stalemate? Which of those postures made the Soviet Union safe in peacetime and in war?

The Evolving Strategic Balance

The United States emerged from World War II as the world's most powerful country. In the years immediately following the

30

conflict, the United States accounted for roughly half of total world economic output. Unlike other great powers, the United States was neither heavily bombed nor occupied during the war; as a result, the country lost 0.32 percent of its population in the fighting, whereas Japan lost 4 percent, Germany 9 percent, and the Soviet Union 14 percent—a staggering forty-five times

31

higher fatality rate than the United States suffered. For every major combatant country except the United States, the war was a calamity.

Not only did the United States emerge from war with an intact industrial base and population, but Washington suddenly had access to a network of military facilities in friendly countries around the world. No longer was the United States geographically isolated; for the first time in history the United States could project substantial military power across the oceans on a day-to-day basis, rather than for just short periods during wars. Overall, on the basis of the aggregate measures that scholars often use to assess power in international politics—for example, gross domestic product (GDP), population, alliances—the immediate post–World War II period seems like a time of triumph and supremacy for the United States.

It did not feel like supremacy at the time, however. The data on GDP and other abstract measures of power fail to capture the dilemmas that the United States faced as the Cold War began. For one thing, U.S. leaders feared that economic despair in Europe and Japan would allow local communist parties to gain political power and then usher their countries into the Soviet orbit. U.S. leaders viewed that possibility as a strategic disaster for the United States, as it would create a massive sphere of

32

Soviet dominance across two of the world's most important industrial regions.

The military situation was even more dire. Specifically, the imbalance in conventional military power in Europe in the late 1940s was staggering. In 1948 the U.S. Joint Chiefs of Staff (JCS) estimated that the Soviet Union had 135 army divisions in Europe, with 100 more divisions available from Soviet allies. Standing between the Soviet Red Army and all of Western Europe were just two U.S. divisions, and those divisions were not combat-ready because they were engaged in postwar occupation duties. Nor could the United States rush major reinforcements to Europe in a crisis; the United States had only ten

33

other divisions worldwide. America's principal military partner, the United Kingdom, had twelve divisions (including forces from the British Dominions), but they were scattered across the world and could not be deployed quickly to the European

34

continent. In retrospect, the JCS estimates probably overstated the size and readiness of Soviet forces in Europe, but the imbalance of conventional military power was so stark that it hardly mattered.

One crucial factor that mitigated the sense of vulnerability in the immediate postwar period was the U.S. monopoly on

35

atomic bombs. How did those weapons affect the balance of power, and how did the balance shift in ensuing years as the Soviet Union developed and deployed its own nuclear arsenal?

1945–1949: U.S. NUCLEAR MONOPOLY

The United States monopoly on atomic weapons lasted from August 1945 until August 1949. The arsenal grew from eleven bombs in 1946 to three hundred in 1950, but this period was no golden age of American supremacy. The main problem was that the U.S. Air Force initially had only a rudimentary capability to deliver these bombs against the Soviet Union. In the late 1940s, a U.S. atomic assault on the Soviet Union would have been a slow and uncertain operation. Few U.S. bombers could carry atomic bombs—approximately thirty-five such aircraft were available from 1946 to 1948—and they lacked the range to strike the Soviet Union directly from the United States, meaning that they would need to launch attacks from bases close to the

36

Soviet periphery. To make matters worse, U.S. fighter aircraft at the time had even shorter ranges, so they could not escort the bombers on their way to Soviet targets, even when flying from forward bases. Thus, a U.S. atomic attack on the Soviet Union in the late 1940s would have required its small fleet of bombers to fly unescorted through Soviet airspace to find and strike Soviet cities. If the U.S. atomic offensive were launched in response to a Soviet invasion of Western Europe—or during another acute crisis—the Soviet air defenses would have been on alert, reducing the odds that the heavily laden B-29s would

37

ever make it to their targets.

The lack of effective fighter escorts was only one of many major problems facing the U.S. bomber force in the era of nuclear monopoly, all of which meant that the air force was unable to launch a single coordinated strike to surprise or

overwhelm Soviet air defenses. The United States stored all of its bombs unassembled; a thirty-nine-man crew needed two full days to assemble a single weapon. Since the United States had only four such crews in 1948, at best, it could prepare two

38

39

atomic bombs per day. Even then, an assembled bomb had only a forty-eight-hour shelf life before its battery died. An atomic offensive would have unfolded slowly, using a few bombs and bombers at a time, giving the Soviet Union ample opportunity to organize its air defenses.

The slow pace of attack might not have been the worst problem facing the U.S. bomber force. War plans called for bombers to conduct operations at night in order to increase the odds of surviving against enemy air defenses, but for safety reasons, the bomber crews trained only in daylight. Furthermore, although U.S. aircrews relied on onboard radar to navigate in the dark, they lacked reliable maps of Soviet territory that would have offered them suitable geographic reference points for radar navigation. In some cases, U.S. bombers were assigned to strike Soviet cities whose locations were not precisely known because U.S. planners were relying on Russian tsarist-era maps. (The Soviet Union exacerbated U.S. intelligence problems by restricting the circulation of maps and introducing errors into them to complicate U.S. targeting.) Finally, although atomic bombs in this era had an estimated destructive radius of approximately 1 mile against a city, the history of U.S. bombing in World War II suggests that at least half of the weapons that American pilots released would fall more than a mile from the

40

intended target.

With all this taken together, the balance of power looked bleak for the United States and its allies in the late 1940s. Given the great conventional military imbalance, the shortcomings in U.S. atomic capabilities were glaring. If war erupted, small numbers of U.S. bombers would be sent unescorted to penetrate an alerted Soviet air defense system at night in search of

41

Russian cities they had never flown over, using old and unreliable maps. U.S. leaders were justifiably worried about the military balance during the age of their atomic monopoly.

1950–1955: SOVIET EXISTENTIAL DETERRENT

The Soviet Union tested its first atomic bomb in 1949, and produced 150 atomic bombs over the next five years. But these weapons were not deliverable against the United States, and the Soviet arsenal was not particularly survivable in the face of a U.S. attack. While the Soviets scrambled to create a bare-bones nuclear deterrent, the United States vastly upgraded its offensive nuclear strike capabilities. The Soviet nuclear arsenal deployed during this period was an “existential” deterrent force.

Although Moscow tested a thermonuclear weapon in November 1955, paving the way for an arsenal filled with vastly more

42

powerful fusion bombs, the country still had no bombers that could reach the United States on a round-trip mission. Soviet bombers could target European cities, but unless they flew one-way suicide missions, they could not reach the U.S. homeland. The problem was as much about the lack of allies and geography as it was about aircraft technology. The only Soviet aircraft capable of delivering atomic bombs was the Tu-4 Bull, which was a copy of the U.S. B-29. But while the United States could launch its nuclear-armed B-29s from bases around the Soviet periphery (including airfields in the United Kingdom, Egypt, Turkey, Pakistan, Japan, and Guam), the Soviet Union had no allied bases close enough to the U.S. homeland that were suitable for the Tu-4. The Soviets sought to extend the bomber’s range by experimenting with midair refueling and placing external fuel tanks on a few Tu-4s, but these efforts failed to give the plane sufficient range to strike targets in the United

43

States and return to Soviet territory.

While the Soviets struggled to develop a rudimentary capability to strike the United States, U.S. offensive nuclear

44

capabilities soared. From 1947 to 1955, the U.S. stockpile of strategic weapons doubled almost every year, from 32 to approximately 1,750. The lethality of U.S. weapons also grew exponentially. The United States tested a fusion bomb in 1952,

45

and quickly began to integrate the new “hydrogen bombs” into the force. The leap in destructiveness from the “simple” fission bombs of the 1940s to the hydrogen bombs of the early 1950s is staggering; for example, the bomb dropped on Nagasaki had an explosive yield of approximately 22 kilotons, whereas the first U.S. thermonuclear weapon had a yield of

46

about 12 *megatons*—540 times the explosive yield of the Nagasaki bomb.

Not only did the number and yield of U.S. bombs soar, but also the United States greatly expanded its strategic bomber

force. The number of aircraft assigned to conduct nuclear strikes grew from roughly two hundred bombers in 1950 to over twelve hundred in 1955, including long-range bombers and medium-range aircraft that could attack Soviet targets from Europe.

By 1955, the nuclear balance was anything but balanced. Soviet bombers still could not reach the U.S. homeland, while the United States could deliver more than fifteen hundred atomic and thermonuclear weapons against a wide range of Soviet

47

military, industrial, and leadership targets in a few days.

To label the Soviet arsenal of 1950 to 1955 as “existential” does not mean it had zero chance of responding to a U.S. nuclear attack with retaliation against the U.S. homeland. Indeed, the core idea behind existential deterrence is that *some possibility* of retaliation exists as long as nuclear weapons exist. In the early 1950s, the Soviets could have tried to retaliate after a U.S. nuclear strike by cobbling together a one-way, suicide mission using any surviving bombers. Nevertheless, the Soviet armed forces had no units trained to carry out such an attack; their bomber crews had not planned or trained to conduct intercontinental missions. Their mission in war was to strike targets throughout Europe, not the U.S. homeland.

1956–1961: SOVIET MINIMUM DETERRENT

The Soviets finally progressed from an “existential” to a “minimum” deterrent force around 1956, when their first intercontinental bombers entered service. Their new, long-range bomber, the Tu-95 Bear, could fly round-trip missions to the United States—though they still needed to refuel at Arctic air bases along the northern edges of the Soviet Union before they

48

headed over the pole. In 1956 there were only twenty Tu-95 bombers in the Soviet inventory, but by 1961 the fleet had grown to 150 aircraft. Although the Soviet Union finally had the capability to strike the United States, its nuclear force

49

remained vulnerable to a disarming attack—that is, it was still not yet a survivable deterrent force.

A bomber force like the one the Soviets fielded could be deployed in a manner that would make it survivable, by developing sophisticated warning systems, creating a nimble command-and-control system, and keeping a portion of the force on alert. The Soviets took a step in that direction by deploying a radar network around their periphery to detect U.S. bombers and cue Soviet fighter planes to intercept them. The radars could also allow Soviet commanders to warn their bomber bases of an incoming U.S. strike. However, the Soviet efforts to enhance nuclear force survivability did not yet bear real fruit for several reasons. First, the United States extensively probed the Soviet periphery with bombers and reconnaissance aircraft to identify holes in the Soviet warning network—specifically, routes and altitudes that U.S. bombers could fly without being detected by Soviet radar. Evidence suggests that the United States found several holes and expected to slip through the Soviet radar screen

50

without much difficulty. More important, even if U.S. aircraft were observed by radar, the Soviets’ new bombers were not prepared to disperse quickly to escape the attack; Soviet planes required six to eight hours of preparation before takeoff, and

51

their nuclear warheads were not stored on the same air bases as the aircraft. Finally, although the Soviets could have raised the readiness of their bombers during a crisis, and moved nuclear weapons to the bomber bases, there was yet another major Soviet vulnerability: the Arctic refueling bases required by Soviet bombers to reach U.S. cities were easy targets for a U.S. first strike. In fact, U.S. bombers would fly over those bases (and presumably bomb them) on their way over the pole to strike airfields and other targets further south in the Soviet Union. Moscow’s new long-range bomber fleet was a step in the right direction for the Soviets, but the result was far from a survivable deterrent.

In 1958 to 1959, the Soviets took two additional steps that did not end their vulnerability immediately, but which laid the groundwork for a truly survivable deterrent force: they began to deploy nuclear-armed submarines and intercontinental ballistic missiles (ICBMs). Initially, the submarine force and the ICBMs were at least as vulnerable as the Soviet bombers. Early Soviet submarines were loud and hence relatively easy for the U.S. Navy to track. Furthermore, the early Soviet submarines had to come within approximately 75 miles of the U.S. coast to fire their missiles, and unlike modern ballistic missile submarines (SSBNs), they had to surface before firing. Most importantly, the early Soviet submarines spent most of

52

their time in port and did not bring their nuclear warheads with them when they went on patrol; the warheads were kept in

53

weapons storage sites on land.

Similarly, the first-generation Soviet ICBMs were highly vulnerable. The R-7 (SS-6 as designated by NATO) had to be fired from a fixed launchpad out in the open rather than from a hardened underground protective silo, as with a modern missile.

Even worse from the Soviet perspective, the missile required twenty-four hours of preparation to launch.

Although the Soviet nuclear arsenal was still highly vulnerable throughout the 1950s, the window for a successful U.S. strike was beginning to close. The Soviet submarine force, flawed as it was, continued to grow, from fifteen submarines in

55

1960 to forty by 1961. Additionally, in 1961 the Soviets began sending submarines on deterrent patrols with their nuclear warheads onboard. At any given time in 1961 only a fraction of the forty Soviet submarines were at sea, and U.S. naval forces could probably destroy many of those, but the challenge of destroying all the Soviet nuclear weapons that could reach the United States was growing rapidly. Another problem for U.S. war planners was the gradual deployment of new Soviet ICBMs. By 1961 the Soviets were fielding newer missiles that were more reliable than the first generation and could be launched in a

56

few hours or less.

Of course, some of the Soviet gains in survivability were countered by growing U.S. offensive capabilities. By 1961, the

57

United States had a robust triad: five hundred B-52 long-range bombers, five Polaris submarines, and fifty ICBMs. U.S. target intelligence was improving as well. In August 1960, the United States acquired its first spy satellite photographs of the interior of the Soviet Union. But these improvements could not prevent U.S. nuclear superiority from slipping away. The Soviet arsenal was becoming too large and too diversified to completely destroy. A successful attack would require striking more than a few hundred Soviet airfields; now ICBM sites, submarine bases, and submarines at sea would all have to be located and destroyed. Moreover, these targets would need to be hit nearly simultaneously. By late 1961, the Soviet “minimum” deterrent was becoming an “assured retaliation” force.

1962–1964: SOVIET ASSURED RETALIATION

The Soviet Union’s vulnerability to a U.S. disarming attack disappeared as the Soviet arsenal grew larger, more sophisticated, and more diverse. In 1958, a U.S. nuclear strike against Soviet strategic nuclear forces likely would have worked unless several things went wrong; by late 1961, it could have worked only if everything went right. And just a few years later, even that would have left the Soviet Union with nuclear retaliatory forces that could be delivered against the U.S. homeland. In short, the balance of power shifted from a condition of Soviet minimum deterrence to one of Soviet assured retaliation.

A simple numerical comparison of U.S. and Soviet strategic nuclear forces in 1962 makes it appear that the United States still enjoyed nuclear supremacy. The United States could have used as many as 2,000 nuclear weapons in a surprise attack

58

against the Soviet Union, with most being delivered by B-52 bombers stationed in the United States. The Soviets had 160 bombers, 38 ICBMs, and 48 nuclear-missile armed submarines. Attacking these forces would require destroying up to 140

59

major Soviet airfields, plus 10 to 25 ICBM launch sites and up to 30 submarine bases. A U.S. first strike would have targeted other sites, too—including an enormous number of additional airfields—but the critical targets numbered only about 200.

But simply counting the number of U.S. warheads and Soviet targets does not give the full picture of the strategic balance at the time. Coordinating a nuclear attack on a diversified nuclear arsenal raised tremendous complications for U.S. war planners. One major problem was the expanded deployment of Soviet nuclear-armed submarines. By 1962, only a fraction of the Soviet fleet was at sea at any given time; however, with forty-eight boats in the force (each with three nuclear warheads), the probability that a few missiles would survive and be launched at U.S. cities had grown dramatically. American antisubmarine warfare (ASW) capabilities were good, but as the United States learned during the Cuban Missile Crisis, it could not always

60

track Soviet submarines in the open ocean.

A second problem for U.S. war planners stemmed from the deployment of newer Soviet ICBMs, which required less time to launch after warning of an attack. Incoming U.S. bombers might be detected as they approached Soviet airspace, about an hour

61

or two before reaching Soviet missile sites. The United States could target Soviet ICBMs with American missiles (ICBMs

62

and SLBMs), but in 1962 U.S. ballistic missiles were relatively inaccurate and unreliable. Even worse, a quick U.S. missile strike on Soviet ICBMs would arrive many hours before U.S. bombers reached some of their targets, giving the Soviets time to get their bombers in the air and submarines out to sea. A third approach would launch U.S. bombers first and delay the missile

attack on Soviet ICBMs until American bombers approached Soviet air defense radars. However, this plan ran the risk that U.S. bombers would be detected early (for example, by human spotters near U.S. bomber bases), giving the Soviets hours to

63

push submarines out to sea and launch their ICBMs in retaliation. In sum, once the Soviets built a nontrivial number of ICBMs and submarines, coordinating an effective nuclear strike became much more complicated.

The odds of the United States executing a successful disarming strike against the Soviet Union had dropped substantially by 1962. There was no precise moment at which a U.S. first strike became impossible; instead, it melted away as Soviet capabilities grew. The age of mutual assured retaliation had arrived. Mutual assured destruction would soon follow.

1965–1990: MUTUAL ASSURED DESTRUCTION

Between 1964 and 1967, the Soviet ICBM force grew rapidly, to over eight hundred missiles. American planners contemplating a disarming first strike in 1967 thus faced the prospect of needing to locate and destroy more than one thousand Soviet long-range nuclear-delivery vehicles (including missiles, submarines, and bombers), which were equipped with more than fifteen hundred warheads. And U.S. planners would need to do so quickly, nearly simultaneously, and with complete strategic surprise. Such a first strike almost certainly would have resulted in dozens of surviving Soviet forces able to inflict massive nuclear retaliation on the United States. In short, at some point in the mid-1960s, the Soviet Union acquired an assured destruction capability, thus ushering in the condition of mutual assured destruction (MAD) that would endure through the end of the Cold War.

PERCEPTIONS OF THE SHIFTING STRATEGIC BALANCE

The strategic balance of power shifted with the gradual growth of Soviet nuclear capabilities. But how did the nuclear

64

balance appear to U.S. leaders? Did they perceive these shifts correctly?

Conventional wisdom about frequent misunderstandings during the Cold War—especially popular notions about a “bomber gap” and “missile gap” in the 1950s and early 1960s—might lead one to believe that U.S. leaders were in the dark about the balance of power, but they were not. At various points in the 1950s there were fears, which we now know to be baseless, about Soviet nuclear superiority. Yet these misperceptions were not about the contemporary balance of power, but about trends in the balance of power—that is, fears about the future. Ironically, even during periods of anxiety about a bomber and missile gap, U.S. leaders had accurate intelligence assessments of existing Soviet strategic nuclear forces.

Given the limits of early Cold War technology, U.S. government intelligence estimates of Soviet strategic nuclear forces

65

were remarkably good. U.S. and NATO officials received intelligence about the number of Soviet nuclear forces operational limitations and vulnerabilities. For example, as early as 1956, intelligence reports concluded that Soviet long-range bombers could strike targets in North America only if they were staged through Arctic refueling bases. This information meant that the United States and its allies could look for the movement of Soviet bombers to Arctic bases as a critical early warning signal of an impending Soviet attack, and it led U.S. commanders to target the handful of such bases with nuclear strikes as a means of

66

neutralizing the backbone of the Soviet strategic nuclear force.

Western intelligence also learned important facts about the operational limitations of Soviet submarines and ICBMs. Intelligence reports noted that the nuclear missiles on early Soviet submarines were short-range, so they could be fired at the United States only if the submarines came very close to the American coast. Intelligence also learned that Soviet submarines

67

had to surface to fire their missiles, and that they spent most of their time in port. Regarding ICBMs, intelligence correctly inferred that early Soviet ICBMs were so big that the launch sites had to be located on railroad lines, which helped U.S.

68

planners identify likely locations of Soviet launchpads even before they were confirmed by satellite photos. Excellent intelligence about the operational limitations of Soviet bombers, submarines, and ICBMs gave U.S. leaders an accurate picture of America’s nuclear advantages.

69

American leaders correctly perceived their nuclear supremacy, and they also understood that it was slipping away. During the last years of the Eisenhower administration, key foreign policy advisors—including National Security Advisor Robert Cutler and Secretary of State John Foster Dulles—focused intently on the question of how U.S. and NATO security strategy

would need to change as the U.S. nuclear edge disappeared. In May 1958 the National Security Council (NSC) debated how
71

policy should change to account for nuclear parity, which Dulles expected would occur “in two or three years.” President Eisenhower understood the impending challenge, but he also recognized that the Soviets were not building nuclear weapons as quickly as some forecasts claimed, and he privately expressed confidence that if it ever came time to fight a nuclear war with the Soviet Union—to push America’s “whole stack of chips into the pot,” as he put it—the United States could win with
72

acceptable casualties. In 1959, Eisenhower told then senator Lyndon Johnson, “If we were to release our nuclear stockpile on
73

the Soviet Union, the main danger would arise not from retaliation but from fallout in the earth’s atmosphere.”

Even the Kennedy administration, which was skeptical about the wisdom of the previous administration’s nuclear-centric war plans, understood that the United States still had viable nuclear employment options as late as 1961. During the 1961 Berlin Crisis, Kennedy’s national security advisor, McGeorge Bundy, authorized a high-level group of senior civilian and military planners to create plans for a specially tailored nuclear-disarming strike against Soviet long-range nuclear forces in case the crisis escalated. The group created an attack plan that they thought might work: it would probably destroy all Soviet intercontinental nuclear forces on the ground, though there was a danger that one or two bombers or submarines might get
74

through, and hence that a few million Americans would be killed. The report concluded with a mixture of optimism and uncertainty that captures the conflicted mindset of leaders facing the dusk of U.S. nuclear supremacy. According to Kaysen’s report: “While a wide range of outcomes is possible,” the United States has “a fair probability of achieving a substantial
75

measure of success.” In other words, there was also a “fair probability” that the United States would *not* achieve substantial success and would, presumably, be hit by a Soviet retaliatory strike. The window for a U.S. first strike may have still been open, but it seemed barely big enough to squeeze through.

U.S. political leaders and military planners understood that nuclear stalemate lay just around the corner. A December 1960 study by the Pentagon’s Weapon Systems Evaluation Group (WSEG) observed that nuclear supremacy would be long gone by the mid-1960s; by then, according to WSEG calculations, a U.S. disarming strike against the Soviet Union would still leave
76

the Soviets with the capacity to retaliate and kill 70 to 80 percent of the U.S. population. Another 1961 study explicitly questioned the feasibility of U.S. national security policy: “In a condition of nuclear stalemate we can no longer rely on the
77

assumption that if hostilities get above a relatively low threshold we will use nuclear weapons.” In sum, the United States still enjoyed strategic superiority in 1961—and its top political leaders knew it—but the margin of superiority was closing.

By mid-1962, senior U.S. leaders understood that nuclear superiority had essentially disappeared. The United States could no longer strike first at Soviet nuclear forces without suffering a substantial Soviet nuclear counterattack. Whereas in 1961 the chairman of the JCS had told President Kennedy that if war erupted the United States would “prevail,” by September 1963
78

Kennedy was briefed that such an outcome was no longer plausible. That same month, McGeorge Bundy prepared a memo for Kennedy on the results of a report by the NSC’s Net Evaluation Subcommittee (NESC), which had modeled the consequences of nuclear war in the 1964 to 1968 time frame based on whether the United States or the Soviet Union struck first. In either scenario, according to Bundy, “the fundamental conclusion is that these wars are unacceptably destructive for
79

both sides.” On September 12, 1963, Kennedy received a briefing on the nuclear balance of power. According to the minutes of the meeting, “the President asked whether, even if we attack the USSR first, the loss to the U.S. would be unacceptable to political leaders. General Johnson replied that it would be, i.e., even if we preempt, surviving Soviet capability is sufficient to produce an unacceptable loss in the U.S. The President asked whether then in fact we are in a period of nuclear stalemate.
80

General Johnson replied that we are.”

The military was not alone in its skepticism about the United States’ continued ability to conduct a successful disarming nuclear strike. Former secretary of defense Robert McNamara argued (in 1989) that “there was no reasonable chance that we could get away with a first strike unscathed [in 1962]. We simply didn’t know where all the Soviet warheads were.” According

to McNamara, launching a nuclear war in 1962 “would have destroyed us as well as the Soviets,” and “if we’d tried a first strike, the Soviets might have had 25 percent of their original three hundred [strategic nuclear warheads] left.” One should be highly skeptical about statements by former government officials offered decades after the pertinent events, but in this case McNamara’s views about the military balance were documented in a private memo he wrote to Kennedy in 1962, just a few weeks after the Cuban Missile Crisis. McNamara wrote, “I am convinced that we would not be able to achieve tactical surprise, especially in the kinds of crisis circumstances in which a first-strike capability might be relevant. Thus, the Soviets would be able to launch some of their retaliatory forces before we had destroyed their bases.”

81

It appears that senior U.S. policymakers believed, at least by the fall of 1962, that the Soviet Union had an assured retaliatory capability. This does not mean that a preemptive strike against Soviet nuclear forces was impossible. As Marc Trachtenberg shows, as late as mid-1963, Kennedy believed he might have to attempt a disarming first strike if certain extreme circumstances arose, such as the imminent conquest of Europe by the Soviets. But leaders appreciated that the odds of the United States executing a successful disarming strike against the Soviet Union had dropped substantially by 1962.

82

Tracking Changes in U.S. Strategic Posture

As discussed above, the first fifteen years of the Cold War witnessed several significant changes in the nuclear balance of power. But how did those changes affect U.S. behavior? In this section we track U.S. behavior—defined as “strategic posture”—along two key dimensions: peacetime foreign policy and war plans. The first dimension focuses on the relative level of caution or aggressiveness that the United States demonstrated in its strategic orientation toward the Soviet Union: What foreign policy goals did the United States pursue, and how much risk were U.S. leaders willing to take in pursuit of those goals in peacetime conditions and during crises? Specifically, what level of deployed Soviet nuclear capability was enough to significantly moderate U.S. behavior? Did the emergence of a Soviet existential or minimum deterrent arsenal induce much greater U.S. caution? Or did more cautious foreign policies emerge only after the Soviet Union deployed an assured retaliatory force?

83

The second dimension of strategic posture consists of U.S. plans for war against the Soviet Union. We track three details about those plans: the centrality of conventional versus nuclear operations; the timing of nuclear escalation (immediate or delayed); and the war aims. What level of nuclear capability did the Soviets have to build to compel the United States to restrain itself during war? Was a small, vulnerable Soviet deterrent frightening enough to force Washington to downgrade the role of nuclear forces in U.S. plans? Was it enough to encourage the United States to restrain its war aims during a conflict? Or did those changes only happen at higher levels of Soviet nuclear capability?

1945–1949: THE ERA OF NUCLEAR MONOPOLY

For more than four years the United States was the only country in the world armed with nuclear weapons. Since there was no Soviet nuclear arsenal to deter the United States, U.S. fears about war with the Soviet Union were based primarily on the expected costs of conventional conflict. This period of U.S. nuclear monopoly, therefore, offers a useful baseline for evaluating the effects of Soviet nuclear deployments in subsequent years. What was the U.S. foreign policy stance toward the Soviet Union during this period? Did the United States pursue ambitious objectives and adopt uncompromising positions during early Cold War crises? If war had erupted during the period of U.S. nuclear monopoly, how would the United States have fought?

U.S. foreign policy toward the Soviet Union during the era of U.S. nuclear monopoly was characterized by relative prudence and caution. To be sure, by 1947, just two years after the end of World War II, the former wartime allies were already engaged in serious geopolitical competition. Both the United States and Soviet Union sought to extend their influence across Europe, and thus limit the influence of their main rival. Tensions flared over the status of divided Germany (and divided Berlin). The United States was not meek during the early years of the Cold War—for example, Washington played hardball to ensure that the people of Western Europe elected anticommunist leaders—but the United States did not maximize its potential leverage from nuclear monopoly.

The generally restrained approach adopted by the United States in this period is reflected in its response to two policy challenges. In 1948 the Soviet Union and its East German partners blockaded the access routes that connected West Germany with West Berlin. Confronting the difficulties of supplying a city located inside hostile territory, and faced with a highly unfavorable conventional military balance, U.S. leaders could have reasonably acquiesced to Soviet pressure and abandoned its position in West Berlin. Alternatively, buoyed by their monopoly on nuclear weapons, U.S. leaders could have dispatched

ground forces down the access routes toward Berlin, daring the Soviet Union to start a war. The United States instead chose a middle course in aiming to break the Berlin blockade with an airlift. This strategy was designed to avoid directly challenging

84

Soviet or East German military forces, while placing the onus on Moscow to escalate the conflict.

The United States also faced a decision in the late 1940s that would define the course of its foreign policy for decades. As its relationship with the Soviet Union deteriorated, U.S. leaders debated how to respond to increasing Soviet hostility. Some advocated a return to isolationism, to avoid being dragged into another costly European war. Others advocated “rollback”: an effort to use political and military means to expel the Soviet Union from Eastern Europe. Instead, the United States adopted containment—a strategy for limiting the spread of Soviet influence without seeking to overthrow East European communist regimes.

Although one might have expected the period of U.S. nuclear monopoly to be marked by bold and aggressive American policies toward the Soviet Union, the United States opted for a relatively restrained approach for confronting the Soviet Union.

No such restraint was evident in U.S. war plans at the time. If a major conflict had erupted in Europe during the period of nuclear monopoly, the United States planned to employ its advantage in atomic weapons to defeat the Soviet Union and win the war. Three aspects of U.S. planning are notable. First, in terms of the conventional-nuclear balance of operations, atomic weapons played the central role in U.S. war plans. The details of those plans changed from year to year, but the overall military concept remained unchanged: U.S. conventional forces would be rapidly defeated or retreat from the continent; the United States would initiate an atomic bombing campaign against Soviet cities and industry, which would weaken the Red

85

Army; after destroying Soviet industry, the United States would reconquer Europe. The point is that during the period of nuclear monopoly, U.S. atomic bombs were not weapons of last resort; they were the primary means for wearing down Soviet power, allowing the United States to defeat the Soviet army.

Second, in terms of the timing of escalation, atomic attacks on Soviet cities would commence at the outset of a major war. The bombing campaign would unfold slowly, because the United States could assemble only two atomic weapons per day, and had few bombers that could deliver them. It would, therefore, take many weeks to attack all twenty Soviet cities that comprised the target set in 1948, or the seventy cities targeted in the 1949 plan. But the methodical pace of the strategic campaign—like the strategic bombing of Germany and Japan a few years earlier—merely reflected the time it would take to

86

execute the plan.

Third, in terms of war aims, throughout the 1940s the U.S. goal in any war against the Soviet Union was the same as it was

87

during World War II: victory. There was no concept in the plans for limited attacks or midwar negotiations; the goal was to

88

destroy Russia’s major cities and then—months later—mop up the remnants of the Red Army.

In sum, throughout the 1940s, U.S. foreign policy and war plans seemed little changed by the atomic age. The foreign policy was designed to contain the Soviet Union without triggering war. U.S. war plans in particular were strikingly similar to U.S. strategies employed during World War II. Next we turn to the question of what, if anything, changed in the U.S. foreign policy approach and war plans after the Soviet Union tested its first atomic weapon in 1949.

1950–1961: EXISTENTIAL AND MINIMUM DETERRENCE

If the existential deterrence school is correct, 1949 should have been a watershed in U.S. foreign policy. The deployment of even a small nuclear arsenal should have made the United States extremely cautious in its dealings with the Soviet Union, to avoid any risk of war. Furthermore, if conventional war did occur, the mere possibility that one or more Soviet nuclear weapons would detonate in Western Europe or the United States should have injected great caution into U.S. war plans. Those plans should have reflected a high degree of military restraint, including steps to terminate the war on acceptable terms as soon as possible. If the minimum deterrence school is correct, on the other hand, each of those changes should have occurred in the mid-1950s, after the Soviet Union deployed its first long-range delivery systems. So, how did U.S. foreign policy change as the Soviets deployed an existential and then minimal deterrence arsenal? And how did U.S. war plans evolve?

The early 1950s were a period of enormous U.S. military advantage. If war had erupted between the superpowers, the Soviet military could have occupied all of continental Europe, and possibly dropped a few nuclear weapons on U.S. allies in Europe and Asia. Ultimately, however, the Soviet Union would have been utterly destroyed by a massive U.S. thermonuclear bombing campaign. Yet, as in the preceding age of U.S. nuclear monopoly, American leaders sought to merely contain communist expansion and avoid war with the Soviet Union, rather than attempting to roll back Soviet control of strategically important regions or provoking military conflict.

The U.S. goal of avoiding a costly war with the Soviet Union was a reasonable one. Given the enormous imbalance of

nuclear power, however, it is surprising that the United States was willing to make political concessions and moderate its behavior in the face of Soviet provocations. Armed with far superior firepower, why didn't U.S. leaders adopt hard-line positions on virtually every point of dispute, and dare the Soviets to start a war that the United States would surely win?

Advocates of existential and minimum deterrence have a ready-made answer to this puzzle. From the U.S. perspective, even a small chance of war—and thus a small chance of triggering Soviet nuclear attack—was too great a risk to accept. Therefore, U.S. leaders opposed Soviet expansion but moderated their policies to minimize the risk of war.

Two occasions during the early 1950s lend credence to the existential and minimum deterrence view. First, U.S. nuclear restraint during the Korean War is difficult to explain without reference to deterrence logic. In late 1950, U.S.-led forces repelled a North Korean invasion of South Korea and pushed Pyongyang's retreating army all the way to the Chinese border. But China intervened, sending thousands of troops into North Korea and inflicting heavy losses on U.S. forces. In the following months, the United States suffered tens of thousands of casualties in brutal fighting against the Chinese army, with the conflict bogging down into a battle of attrition for two more years. Why did the United States accept such a crushing military defeat and stalemate? Why didn't Washington employ nuclear weapons against massed Chinese forces in North Korea, or against government or military targets in China? After all, China had no nuclear weapons, and its Soviet partner should have been too terrified by U.S. military advantage to either retaliate in kind or launch an invasion of Western Europe. With its enemies outgunned, why did the United States keep its most powerful weapon holstered?

More broadly, in 1953 the new Eisenhower administration conducted a review of U.S. strategy toward the Soviet Union. The basic options on the table were to stick with containment or adopt a more uncompromising stance of forcing leaders in Moscow to choose between accommodating the West and fighting it. The Eisenhower administration opted for the former—containment. Clearly the Eisenhower administration did not want to pay the costs of war, but why did it believe that a hard-line position would lead to violence? Why wasn't Eisenhower confident that superior U.S. military power would force the Soviets to choose peace?

The logic of existential and minimum deterrence may have played a role in U.S. restraint, although historians continue to debate the causes in both cases. For example, in terms of the nonuse of nuclear weapons in Korea, some scholars argue that

89

normative constraints or reputational concerns were decisive; others claim there were few meaningful targets in North Korea,

90

or highlight concerns that ineffective nuclear use would undermine U.S. geopolitical leverage. Nevertheless, the military balance was far more favorable to the United States in the early 1950s than it had been in the late 1940s, so why were U.S. leaders reluctant to employ the coercive leverage that their military superiority should have provided, to force China to halt and deter Russia from intervening? It is at least plausible that the mere possibility of provoking a Soviet nuclear strike—the logic of existential and minimum deterrence—was sufficiently terrifying to moderate U.S. policies during this period.

Although the Soviet existential and minimum deterrent posture *may* have restrained U.S. foreign policy, it did nothing to moderate U.S. military plans. In fact, the growth of Soviet nuclear capabilities had the opposite effect. The United States continued to depend on nuclear threats to deter attacks on Europe, and to rely on nuclear war as the preferred means of fighting if major conflict erupted. U.S. war plans were not designed to prevent or delay escalation, but to conduct nuclear strikes against Soviet targets as quickly as possible. The plans did not limit the targets to be destroyed; they aimed to disarm the Soviet nuclear force, kill the Soviet leadership, and destroy Soviet urban-industrial centers. The goal of U.S. war plans in the 1950s was simple: decisive victory.

U.S. war plans remained as nuclear-centric as ever throughout the 1950s. What changed were the size of the planned U.S. attack and the prioritization of targets. Immediately after the first Soviet atomic test, a Pentagon study concluded that the coming deployment of Soviet nuclear weapons required the United States to shift its war plans; rather than focus on Soviet

91

cities, U.S. operations would need to prioritize Soviet nuclear capabilities.

The focus on counterforce targets required a vast increase in U.S. nuclear forces. Some analysts have derided the expansion of the U.S. nuclear arsenal in the 1950s as the beginning of the era of "overkill," when the number of weapons far exceeded what was necessary to destroy every meaningful Soviet city. But destroying Soviet cities was no longer the main purpose of the arsenal. The goal was, first and foremost, to destroy Soviet nuclear forces. Back in 1949, before the Soviets had tested an atomic bomb, the U.S. war plan called for destroying seventy Soviet cities. In 1955, the plan still included attacks on cities—118 of them—but the focus of the attack had become the destruction of 645 airfields located across the Sino-Soviet

92

bloc, plus Soviet air defense sites (whose destruction was necessary for U.S. bombers to reach their main targets). Destroying those airfields was the key to victory, as it would prevent the Soviet Union from launching its own atomic raids against NATO

cities and forces. Whereas the plans in the 1940s could afford to be methodical, the strikes in the 1950s had to hit the Soviets quickly, before Soviet bombers could begin their own operations. (By the late 1950s, the U.S. attack plan also required that

93

Soviet ballistic missiles needed to be hit before they could launch.)

The development of a Soviet existential deterrent—and the subsequent transition to a minimum deterrent force—therefore changed some aspects of U.S. war plans, such as increasing the speed and ferocity of a U.S. strike. But a key element remained constant: the objectives. The goal of U.S. war plans during the age of existential and minimum deterrence was neither coercion (to compel war termination) nor vengeance, but simply victory. The goal was to win World War III by destroying the Soviet Union's nuclear force, as well as the industry that could support Soviet conventional military operations.

Other scholars have highlighted, with good reason, the ways that U.S. war plans were shaped by internal organizational or bureaucratic politics (for example, a mission rivalry between the U.S. Air Force and Navy), or a military-industrial complex. But the strategic logic underpinning those plans was clarified as the age of minimum deterrence neared its end and the Soviet arsenal became more difficult to destroy. In 1961, the United States found itself embroiled in a crisis over Berlin. President Kennedy's national security advisor established a small group within the White House to evaluate U.S. nuclear plans and suggest ways to improve U.S. counterforce capabilities. The White House group, led by Carl Kaysen, proposed a smaller strike option comprised of only forty-one bombers, to maximize the odds of them slipping through holes in the Soviet radar network undetected. The downside of that strike plan, though, was that it would be too small to eliminate the medium- and short-range Soviet nuclear weapons that threatened Washington's European allies. When President Kennedy was briefed on the Kaysen group's report, he was apparently not shocked to learn that U.S. success in a nuclear war hinged on focusing U.S. strikes against the Soviet *long-range* nuclear forces (meaning, delaying attacks on the Soviet bombers that targeted European cities). Instead, Kennedy asked for a briefing the following day from General Thomas Power, the commander of the Strategic Air Command, to address the mechanics of the plan. Kennedy wrote to Power: "How would you plan an attack that would use a minimum-sized force against Soviet long range striking power only, and would attempt to achieve tactical surprise? How long

94

would it take to develop such a plan?"

U.S. strategic posture in the 1950s is often harshly criticized. Some analysts claim that it left the United States unable to deal with more limited acts of Soviet aggression because Washington would be unwilling to trigger a mutually catastrophic nuclear war to counter such aggression. Other critics claim that the rapid growth of the U.S. arsenal reflected an illogical desire for overkill because there were not nearly enough Russian cities to require such a vast U.S. nuclear force.

But those criticisms miss the mark. In the 1950s, the Truman and Eisenhower administrations were not planning to respond to an act of Soviet limited aggression—for example, against West Berlin—by triggering a *mutual* catastrophe. Rather, they planned to use U.S. nuclear forces to conduct a massive disarming strike to eliminate Soviet nuclear forces during any significant conventional conflict in Europe. In other words, the large U.S. arsenal would have been "overkill" had their mission been to destroy Soviet cities, but they were deployed with an altogether different objective: to rapidly destroy the Soviet nuclear arsenal at the outset of any major war. As long as the United States felt it could destroy the Soviet arsenal (be it an existential or minimum deterrent one) with a nuclear-disarming strike, it relied on plans to do so.

1962 AND BEYOND: THE ERA OF ASSURED RETALIATION

If advocates of assured retaliation are correct that fielding a survivable retaliatory force is the key threshold for creating stalemate, then the 1960s should have seen significant changes in U.S. defense plans. The Soviet Union's deployment of a secure retaliatory force should have created major problems for the United States. After all, how could the United States defend Europe if war erupted, given that a major nuclear strike against the Soviet Union no longer made sense in light of that country's ability to retaliate? The United States also faced a deterrence problem. If responding to a conventional Soviet attack with a nuclear strike no longer made sense, how could Washington use its powerful nuclear arsenal to deter that attack in the first place? And how could U.S. nuclear capabilities be used to assure NATO allies?

If the assured retaliation school is correct, then U.S. leaders would have needed to grapple with these dilemmas in the early 1960s. The United States should therefore have grown more cautious in its dealings with the Soviet Union, and its strategy for deterring and waging war should have changed as well.

The Soviet deployment of an assured retaliation capability did not cause a notable decline in U.S. foreign policy assertiveness, mainly because U.S. foreign policy already aimed at containing the Soviet Union without provoking war. The crisis that erupted just as the Soviets were deploying their first survivable capabilities—the Cuban Missile Crisis—is indicative of both of these threads in the U.S. approach. On one hand, the United States was willing to confront the Soviet Union to contain the expansion of Soviet power. On the other hand, the United States crafted its policies at each stage of the confrontation to reduce the probability of war. At the start of the crisis, the United States chose to blockade Cuba rather than launch air strikes. At the end of the crisis, the United States chose to trade away its missiles in Turkey rather than invade Cuba.

Throughout the Cold War, the United States sought to resist Soviet political and military gains, but do so in a manner designed to avoid direct military conflict. The development of an assured retaliation capability did not compel the United

States to abandon the former, but neither was it necessary for the latter. The more relevant changes in U.S. strategic posture occurred in the realm of war plans.

U.S. leaders reacted to the changed strategic landscape in the 1960s as advocates of assured retaliation would anticipate. The United States realized that U.S. strategy could no longer rest on threats to start a general nuclear war, which would have been suicidal to execute and hence lacked credibility. U.S. leaders, therefore, sought other options for war and deterrence, such as stronger conventional defenses, limited nuclear options, and plans for gradual nuclear escalation. Post-Cold War historical accounts have shown that the actual changes in U.S. war plans occurred slowly, and each concept for implementing limited

95

nuclear options raised new problems for the United States and its allies. Nevertheless, the United States did adjust its war plans in response to the emergence of a Soviet assured retaliatory capability, even though U.S. planners never fully “solved” the challenges posed by the Soviets’ growing arsenal. In short, the growth of the Soviet nuclear arsenal finally forced the United States to abandon its prompt and massive nuclear war-winning plan; in that crucial respect the Soviet force had finally produced real peacetime and wartime security benefits for Moscow.

Previous historical interpretations of the “flexible response” turn in U.S. war plans in the 1960s are consistent with the assured retaliation school of deterrence, but those accounts are not accurate. According to the old interpretation, because the United States could no longer win a nuclear war, its threats to respond to a Soviet invasion of Western Europe with massive retaliation lacked credibility; therefore, the United States and NATO developed a flexible response strategy. Flexible response had two related components. First, NATO would strengthen its conventional forces so they would be powerful enough to thwart minor Soviet incursions. This would mean that only a full-scale Warsaw Pact invasion could overwhelm NATO’s conventional forces. Faced with a full-scale invasion, however, NATO’s threats to escalate to the nuclear level would be more credible—especially if the United States and its allies had options for escalation that might not trigger a full-scale nuclear war. The second component of flexible response was the development of “limited” or “graduated” nuclear options. Rather than force a president to choose between surrender and suicide, flexible response allegedly offered U.S. and NATO leaders a menu of choices for nuclear escalation, to coerce the Soviet Union to halt the war before all sides had plunged into the abyss.

Evidence from the 1950s and 1960s reveals that U.S. leaders from both the Eisenhower and Kennedy administrations diagnosed the problem they faced exactly as the assured retaliation school would expect. Even before the Soviet Union had

96

deployed a survivable arsenal, senior officials in the Eisenhower administration anticipated the coming problem. For example, during a May 1958 NSC meeting, Eisenhower’s national security advisor Robert Cutler declared that the United States needed to reevaluate its overarching national security strategy, given the impending situation in which nuclear forces, “capable of delivering massive nuclear devastation (regardless of which side strikes first) increasingly deters each side from

97

initiating, or taking actions which directly risk general nuclear war.” Cutler argued that the coming era of “nuclear parity and mutual deterrence” would require a “flexible and selective capability ... to deter or suppress limited military aggression.” Secretary of State John Foster Dulles agreed that the United States would need to develop options that would not rely on the threat of “wholesale obliteration,” and suggested that tactical nuclear weapons might contribute to NATO’s defense without

98

triggering Soviet retaliation. Dulles had written an earlier memo describing the looming situation, in which he said it would be “suicide” to engage in a nuclear exchange with the Soviet Union and advocated increased spending on conventional forces

99

and the development of limited nuclear options, which he referred to as “graduated retaliation-coercion.” In a conversation later that year, Dulles told Eisenhower that the existing U.S. strategy was “rapidly outliving its usefulness and we need to

100

apply ourselves urgently to finding an alternative strategic concept.”

But diagnosing the problem and solving it were entirely different matters. Throughout the Kennedy and Johnson administrations, senior U.S. leaders urged the Pentagon to produce limited war options, but were frustrated by the slow progress. In fact, when the Nixon administration took office in 1969, the president and his national security advisor Henry

101

Kissinger were horrified at the lack of flexibility in U.S. war plans, and the enormous scale of even the smallest option. Although recent histories attribute the Pentagon’s resistance to changing the war plans to organizational inertia and interservice rivalries, there is another persuasive explanation. We now know that the Pentagon was working hard to escape stalemate (that is, to render the Soviet arsenal a minimal, rather than assured, deterrent force once again), by tracking Soviet submarines, by exploiting weaknesses in Soviet command and control (C2), and by enhancing U.S. offensive strike capabilities. Pentagon planners understood that executing limited war options might forestall the possibility that subsequent

strikes could disarm the Soviet arsenal—since limited attacks would bring the Soviet arsenal up to full alert, as well as task some delivery systems critical for a disarming strike to those limited operations—and thus they did not want such options to be

102

available. Regardless of the explanation for the Pentagon's reluctance to overhaul U.S. strategic war plans, the changes were indeed slow.

Newer historical accounts provide a necessary corrective to the previous interpretation of strategic plans in the 1960s, but they should not obscure the fact that real changes did occur, and U.S. options did grow. The war plan that was briefed to Nixon and Kissinger in 1969—SIOP-4, which had been the plan since 1966—was less flexible than Nixon or Kissinger wanted, but it provided them with significantly more options than the plans from the era of U.S. superiority. SIOP-4 divided targets in the Soviet Union into three categories: Soviet nuclear forces and command-and-control (Alpha), nonnuclear military forces (Bravo), and urban-industrial targets (Charlie). The war plan offered flexibility in at least four respects: it allowed the president to choose virtually any combination of Alpha, Bravo, and Charlie target sets, and exclude the others; it permitted country-specific “withholds,” meaning that nuclear targets outside the Soviet Union could be spared; it allowed for capital city withholds (Moscow and Beijing), so that Soviet and Chinese leaders would survive and be available for war termination

103

discussions; and it included “at least ninety ‘sub-variations’ of target withholds.” U.S. leaders still wanted more and better options, and even the smallest attack in the war plan (an Alpha-only disarming strike) was a massive operation. The plan did

104

not offer the sort of limited options called for in a flexible response doctrine. But the changes reflected progress in that direction.

Equally important, the United States and its allies possessed a second set of nuclear capabilities—NATO's regional nuclear forces—that appear to have given U.S. and allied leaders much more flexibility. By the mid-1960s, the NATO commander (a U.S. general) had control of more than one hundred nuclear weapons to be delivered by aircraft and ground forces under his command. Furthermore, NATO had by then created a range of nuclear plans for using those weapons that look very much like what “flexible response” advocates would have desired. There were nuclear plans for scenarios involving small-scale Soviet attacks, which called on NATO to use very few nuclear weapons (specifically, five) in demonstration strikes to coerce the

105

Soviets to halt operations. There were also bigger nuclear options for scenarios in which the Warsaw Pact launched a major invasion of Western Europe, involving NATO nuclear strikes against Soviet theater nuclear forces, Soviet tactical nuclear forces, Warsaw Pact conventional forces, and logistical sites in Eastern Europe that would be essential to the Warsaw Pact

106

conventional operations.

To be clear, these NATO nuclear plans had their own limitations, which is why every U.S. administration sought better limited war options. The greatest limitation was that several NATO members—including Britain, France, and West Germany—disagreed among themselves and with the United States about how, when, and whether these limited nuclear

107

options would be executed. But the (understandable) reluctance to execute limited nuclear war plans does not obviate the fact that the plans and capabilities were created, and thus NATO possessed real limited-war options it could have executed in a

108

crisis or war.

To take a step back, the assured retaliation school predicts that the concepts for defending Europe would *not* change in 1950 when the Soviet Union deployed its initial, rudimentary arsenal, nor in 1956 when the Soviets first developed the capability to strike the United States. If the assured retaliation school is correct, then the United States would have retained its war-winning plans throughout the 1950s, undeterred by the Soviet existential and then minimum deterrence postures. Rather, according to the assured retaliation view, U.S. plans for defending Europe—and waging war if it occurred—would be thrown into disarray when the Soviets deployed a truly survivable nuclear capability. That is exactly what happened.

The inability of U.S. leaders to truly “solve” the problem posed by an enemy that possessed an assured retaliation force is not surprising. Getting to stalemate is, indeed, a very useful security strategy. But what this chapter reveals is that the process of getting into stalemate can be long and dangerous. It was long and dangerous for the Soviet Union: their effort to create an assured retaliation force triggered a major arms race in the 1950s and 1960s. Furthermore, establishing stalemate did not end the competition. U.S. war planners worked throughout the Cold War to reestablish superiority (with a level of success that the historical literature has not yet fully absorbed). And even in the condition of stalemate, competition continued, as the United States and its allies developed nuclear capabilities to mitigate the problem that stalemate caused: how to deter or fight an enemy that possessed superior conventional forces and a secure retaliatory arsenal.

Perhaps Plans Are Just Plans?

Skeptics may wonder whether we can draw reliable conclusions about deterrence by looking at changes in war plans. After all, peacetime plans are often discarded the moment war begins. Political leaders rarely focus on the details of military matters unless war seems likely, and once they do, they often demand significant changes to those plans. Furthermore, military planners frequently present multiple plans and options to leaders. In short, according to this criticism, the existence of one war plan in some dusty archive may not reveal anything meaningful about actual leaders' prewar preferences or the option they

109

would select if conflict erupted. The Kennedy administration's efforts to revise U.S. war plans in the midst of the Berlin

110

Crisis of 1961, discussed above, is but one example of plans being revised on the cusp of war.

There are at least three reasons to believe that the changes and continuities in U.S. strategic posture from the 1950s to the 1960s shed light on how the United States would have waged World War III. First, unlike many war plans, the U.S. strategic posture received considerable and sustained attention from senior political officials. In NSC meetings and other settings, the Eisenhower and Kennedy administrations repeatedly debated U.S. plans for defending Europe, including the wisdom of relying on a nuclear-centric strategy. The "principals" themselves, rather than military planners, were ultimately the ones to resolve these debates over U.S. strategic posture.

Second, although leaders facing the outbreak of war are often able to choose from a range of contingency plans, in the 1950s

111

the United States did not have a range of options—it had one plan. The plan steadily evolved during this period, as U.S. Strategic Air Command (SAC) acquired more weapons and bombers to deliver them, and as the number of targets in the Soviet Union ballooned. Indeed, the predelegation of nuclear use authority to military leaders in both the Eisenhower and Kennedy administrations meant that in certain circumstances the SAC commander may have had the ability to use nuclear

112

weapons on his own initiative. Until the mid-1960s, the U.S. president did not have a list of attack options from which he

113

could choose if the Soviets crossed the inner German border; instead, he had only one option.

Third, although there are many examples in history of political leaders modifying peacetime plans as war approaches, one should recognize the limits of those modifications. It is possible to adjust a plan before implementing it—for example, by strengthening the attack in one sector and weakening the effort in another. But adjusting the location of a ground offensive, or adding additional ships to a naval task force, is different from radically changing plans at the last minute about how one is going to fight. The argument that we should not attribute too much significance to what were "just plans" in the 1950s suggests that during a major crisis, U.S. leaders might have fundamentally revised the U.S. military's core concepts for fighting the Soviet Union. Such critics posit that the peacetime plan—a massive strategic nuclear offensive—might have been overturned in favor of a conventional defense of Western Europe with only limited nuclear operations. But it is very difficult to see how U.S. leaders could conjure up such a replacement plan in the midst of a crisis. A change of that magnitude would require entirely different concepts of military operations, as well as a different mix of forces.

One example described above, of the rushed efforts by the White House to create alternative military options during the 1961 Berlin Crisis, illustrates both the validity and limitation of this critique. As the Kennedy administration realized that the crisis over Berlin could lead to war, they reviewed U.S. war plans and were unsatisfied with the only option they had—a single massive nuclear strike on the entire Sino-Soviet bloc. But the White House effort to create an alternative war plan did not come up with a radically different way of fighting the Soviets, such as with a conventional-only defense of Europe. Instead, they created a modified strategic nuclear offensive. The biggest change was a reduction in size of the initial U.S. nuclear strike force in order to help slip bombers through holes in the Soviet radar network, thereby increasing the odds of destroying all long-range Soviet nuclear-delivery systems on the ground. The operational concept (rapid nuclear disarming strike) and goal (victory, by destroying the Soviet arsenal) remained the same.

Leaders can adjust plans during crises, but they are limited in what they can do by the available forces, military doctrine and training, and the need for political and military coordination with allies. The United States faced such constraints in the 1950s, and thus was committed to only one option for defending Europe: launching a strategic nuclear offensive against the Soviet Union. Alternative plans and options were seriously considered only as the Soviet ability to respond with nuclear weapons grew to the point where retaliation was virtually assured.

The development and growth of the Soviet arsenal in the 1950s did not have the impact on U.S. strategic posture that either the existential or minimum deterrence schools would expect. The Soviet Union's initial deployment of nuclear weapons did not make it much safer and more secure. Instead, although the shifting nuclear balance may have led U.S. leaders to back away

from covert actions aimed at subverting the Soviet system, it energized U.S. efforts to build weapons and war plans to rapidly

114

destroy the vulnerable Soviet arsenal if war erupted. American leaders only began to moderate U.S. strategic posture in the early 1960s, when the Soviets began to field a true assured retaliation capability, by seeking ways to avoid triggering a mutual nuclear holocaust at the outset of any war.

Two additional observations about the Cold War case are noteworthy, not least because they reinforce the findings in subsequent chapters. First, the development of a Soviet assured retaliation capability did not end the superpower nuclear arms race or competition more generally. Instead, the onset of nuclear stalemate in the 1960s led U.S. leaders to search for ways to reestablish nuclear superiority. Just as the United States spent the 1950s struggling to delay the onset of nuclear stalemate, U.S. leaders spent the rest of the Cold War trying to escape nuclear stalemate and return to the prior era of strategic superiority. For example, U.S. efforts to track, target, and destroy Soviet nuclear weapons and delivery systems, as well as to monitor and jam Soviet strategic command and control, reflected this mission. Although the details of these efforts are only slowly emerging, and those contemporary nuclear plans remain highly classified, there is substantial evidence that the United States has consistently responded to adversary nuclear developments by enhancing its own counterforce capabilities in case war occurs. This reinforces the conclusion that vulnerable arsenals may bolster deterrence in peacetime but prove dangerous in wartime. And it provides evidence for the argument—which is the focus of the next chapter—that even when a country establishes an assured retaliation capability, its geopolitical challenges continue because its enemies may seek to reverse the condition of stalemate.

The second important observation to flow from the Cold War history is that despite the onset of stalemate sometime in the 1960s, the United States and its NATO allies remained committed to deliberate nuclear escalation as a strategy to fight and deter a Soviet conventional attack. Even when facing the massive, survivable, and redundant nuclear capabilities that the Soviets deployed in the later decades of the Cold War, a pillar of NATO defense plans was to launch tactical and theater nuclear strikes in order to deter or coerce the Soviets to halt any invasion. The challenge of deterring conventional war with nuclear weapons under conditions of nuclear stalemate is the theme of chapter 4.

The analysis in this chapter sheds substantial light on the overarching question of this book: Why have international relations remained so competitive during the nuclear age? One answer—the subject of this chapter—is that at least during the Cold War, deploying a small, vulnerable force was not “enough” to give the Soviet Union the security they needed. A small and vulnerable nuclear force did not give the Soviet Union adequate protection from attack, which explains why the Soviets continued to strive to establish more robust capabilities. Moreover, U.S. behavior—especially the massive arms buildup in the 1950s—makes sense once one understands how highly the United States valued its ability to win a nuclear war if conflict with the Soviets erupted. The result was an intense arms race, concerns over relative gains, and competition for key allies (and the use of their territory for staging nuclear strikes).

This chapter offers one window into power politics in the nuclear age. Although nuclear weapons provide an excellent deterrent against attack, there are good reasons for new nuclear-armed states to be dissatisfied with a small and vulnerable arsenal. Such a force might induce caution during peacetime, but it might also invite attack during a crisis or war. New nuclear powers are therefore wise to build up their nuclear forces and develop truly survivable arsenals. The other side of the coin is that rivals of new nuclear states may prefer to retain war-winning options in case conflict occurs, and hence those rivals have an interest in deploying counterforce capabilities to delay the onset of assured retaliation. The understandable result of this dynamic is arms racing, balance-of-power competition, and continued hostility.

115

As Thomas Schelling once warned, “a fine deterrent can make a superb target.” The Soviet Union’s initial deployments of nuclear weapons in the 1950s likely bolstered deterrence by reinforcing the desire of U.S. leaders to avoid war, but the bigger effect was to vastly increase the number of targets in Soviet territory that would have been hit by U.S. nuclear forces had World War III erupted. It was not until the early 1960s that the Soviet arsenal ceased being primarily a target and started becoming a robust deterrent. The evidence from the Cold War indicates that arsenals that lack robust, capable, and survivable delivery systems do not constitute a solid foundation for nuclear deterrence, much less offer reasons to expect a transformed, more pacific state of international politics.