Thus far, I have examined only long, high-intensity wars. The repeated appearance of commitment problem concerns among the central motivations to fight in these cases provides strong evidence in favor of the claim that concerns about an adversary’s inability to commit—whether because of adverse shifts in relative capabilities or because of a belief in the opponent’s dispositional commitment to war—produce unusually destructive wars. As I noted at the outset, however, a research strategy that involves examining only large wars is problematic. Most fundamentally, if similar commitment concerns are also present in more minor wars, then it is hard to argue that they account for the difference between bigger and smaller conflicts. Moreover, an explanation for large wars becomes more compelling to the extent that we also have a coherent explanation for more limited conflicts, such as those provided by the informational and principal-agent mechanisms.

This chapter and the next thus examine a set of wars that were more limited in severity and, typically, duration than the wars discussed thus far. In every case, either the informational or the principal-agent mechanism turns out to have been most significant in bringing about the war, while significant commitment problem concerns were absent. This chapter examines the Persian Gulf War of 1991 and the Anglo-Iranian War of 1856–57, each of which I argue was driven by the informational mechanism.

The Persian Gulf War

The optimism about international peace that followed the end of the Cold War did not last long—the Iraqi invasion of Kuwait in the summer
of 1990 and the subsequent war between Iraq and a coalition led by the United States demonstrated that the use of force would remain a feature of the new international system. Whereas in the Iran-Iraq War, Saddam Hussein launched his invasion in response to a fear of relative decline, in this case closer analysis reveals that his attack was a consequence of overoptimism about the likelihood that outside powers would intervene to reverse his fait accompli and the degree to which Iraq could successfully defend its conquest in the event of war. It is thus not surprising that the war followed the logic of the informational mechanism quite closely. Starting from quite disparate expectations about the likelihood and consequences of war, the participants advanced incompatible demands that precluded a peaceful settlement. Indeed, this process occurred twice, first with the failure of negotiations between Iraq and Kuwait prior to the Iraqi invasion, and then with the inability to reach a negotiated settlement to the resulting crisis between the Iraq and the coalition that mobilized against it. Ultimately, it took war to bring these expectations into alignment: once the fighting started, and especially once the shift to ground combat demonstrated that the Iraqi army was not capable of resisting the opposing coalition effectively, Saddam Hussein rapidly scaled down his demands, ultimately accepting a settlement far worse than the one on offer prior to the outbreak of violence.

A Brief Review of Events

Iraq emerged in 1988 from its eight-year war with Iran militarily and economically exhausted. In the course of that conflict, the country had accrued significant debts, including loans from many of its neighbors that it subsequently resisted repaying, on the rationale that its armed opposition to Iran had served the general interest of the Arab Gulf states. The refusal of these governments to forgive the loans after the war thus was a source of significant discord, as was overproduction of oil by some of the smaller OPEC countries—notably including Kuwait—which drove down the price of Iraq’s primary export and thus hindered the country’s efforts to recover financially and rebuild after the war.1 At the same time, the Iraqis had long claimed, on the basis of the internal organization of the Ottoman Empire, that Kuwait was legitimately part of the province of Basra and hence should be governed by Iraq.2 When the Kuwaitis refused to address Iraqi grievances, war thus was an attractive option, and the Iraqi army duly invaded on August 2, 1990, completing the conquest of Kuwait within a day.

The first Bush administration in the United States reacted strongly to the Iraqi invasion, helping to secure a series of UN resolutions
denouncing Iraq’s actions and organizing a coalition of thirty-four countries who sent soldiers, first to defend Saudi Arabia against potential further Iraqi expansion and then to compel an Iraqi withdrawal from Kuwait. When the imposition of economic sanctions failed to induce compliance, the coalition moved first to aerial bombing and then eventually to a ground war. Impartial observers had predicted high potential death tolls on both sides, but American technological superiority, a superior coalition strategy, and the reluctance of most Iraqi soldiers to actually fight resulted in a far more one-sided conflict than many had expected. Rather than launching an attack directly into Kuwait, the coalition instead flanked the Iraqi army by invading from Saudi Arabia into southern Iraq, from where it was in a position to cut off communication between Baghdad and Iraqi forces in Kuwait. The success of this attack forced the Iraqis into a rapid withdrawal at high cost, with the coalition demolishing units retreating along the highway from Kuwait City to Basra. The course of fighting forced Saddam into a rapid political retreat and even opened up the possibility of a march on Baghdad, which the Bush administration however opted not to pursue. The two sides thus reached a political settlement in which Iraq not only recognized Kuwait as sovereign and independent but made a slew of additional concessions, including the imposition of a UN monitor force, the destruction of its weapons of mass destruction program, and the payment of reparations to Kuwait.

A number of questions follow from this brief history. Why did Iraq take the gamble of invading Kuwait? Why did Saddam refuse to back down even after a broad coalition had formed to compel withdrawal? And, most important from the perspective of this project, why, given the sharply contrasting positions that each side established prior to war, did fighting end so quickly? To preview findings, the war conformed well to the predictions of the informational mechanism. Although the Americans arguably underestimated the ease with which they would win, they still expected to do far better than the Iraqis believed they would. Meanwhile, Saddam Hussein had serious (and not unfounded) doubts about the willingness of the United States and others in the West to actually fight, the strength of American resolve once war began, and the ability of the coalition to expel Iraqi forces from Kuwait at anything approaching acceptable costs. Given this divergence in expectations, pre-war negotiation failed to identify a mutually acceptable political settlement. Once fighting began, both sides learned about the other’s capabilities and resolve, and the new information forced Saddam to acquiesce rapidly to demands that he had ruled out as unacceptable prior to the war.

[161]
The Sources of Optimism

The central disagreement for the Persian Gulf War was between the United States and Iraq, but it obviously followed a prior disagreement between Iraq and Kuwait. A few words on Kuwaiti policy will thus be useful, as there is reason to believe that the entire conflict might not have happened had the Kuwaiti leadership understood that Saddam was willing to invade. The two countries had a number of disagreements—including the aforementioned loan disputes, Iraqi sovereignty claims, the possibility of Iraqi usage of the Kuwaiti islands of Warba and Bubiya, Iraqi grievances about Kuwaiti overproduction of oil, and disagreements related to the Rumaila oil field—on which the Kuwaitis consistently refused to make any significant concessions. This unyielding negotiating position was in large part a function of confidence, grounded in reports from the United States and from Arab intermediaries—most notably Egypt’s Hosni Mubarak—that Saddam was bluffing and would not in fact invade. From the opposite perspective, however, this situation meant that an Iraqi government that was willing to resort to force, and that enjoyed overwhelming military superiority, was confronted with an opponent whose demands were unreasonable given the likely outcome of a war. The ensuing fighting amply demonstrated that Saddam was willing and able to enforce his demands, but his capability and resolve could only be credibly demonstrated to the Kuwaitis by the actual use of force.

Of course, invading Kuwait was attractive only so long as the Iraqis had reason to think that they could get away with doing so. The available evidence, however, indicates that Saddam did in fact believe that the consequences of an invasion for Iraq would be mild. This expectation was grounded in three separate beliefs: that the United States lacked the resolve to fight over Kuwait; that even if the Americans were willing to fight the Saudis would not be willing to supply the bases that would be needed for an effective response; and that in the unlikely event of a war the Iraqi army would be able to impose significant casualties on its opponents, which in turn would lead the Americans (and hence the broader coalition) to back down. It is thus worth spending some time discussing the origins of these beliefs.

Expectations of American irresolute—whether with respect to the willingness to oppose the Iraqi conquest or the willingness to stick it out in a war once fighting began—stemmed in large part from a belief that the American people, and thus the government that represented them, were unwilling to suffer significant casualties, especially in pursuit of political aims in peripheral parts of the world. Thus, for example, in his first and only meeting with US ambassador April Glaspie, on July 25, Saddam
observed that “yours is a society which cannot accept 10,000 dead in one battle.” This belief was not new. A decade earlier, in the face of fears of Iran engendered by that country’s revolution in 1979, Saddam concluded that relying on the United States for protection was a losing proposition, as the Americans had found that providing such protection was too costly and risky a venture to undertake. Likewise, at a meeting of the Arab Cooperation Council in February 1990, he had observed that “all strong men have their Achilles’ heel. . . . We saw that the U.S. as a super-power departed Lebanon immediately when some Marines were killed.”

Moreover, Saddam increasingly came to think that his opponents cared more about his citizens than he did, and that images of Iraqi suffering would undermine public support for the war even in the event that his armies were unable to impose significant casualties on coalition forces. Thus, for example, at one point during the crisis the Iraqi leader predicted that CNN coverage of the first two bombing runs on Baghdad would generate enough public pressure on the administration that Bush would have to call off the air war.

The belief that the Americans lacked the resolve to respond to the invasion with force gained support from the policy of “constructive engagement,” under which the United States tried to remain conciliatory in the face of Iraqi provocations in the expectation that moderation would turn Iraq into a regional partner, if an admittedly sometimes difficult one. In this regard, the criticism that Ambassador Glaspie endured for not providing a stronger deterrent threat—by some accounts giving Saddam a “green light” to invade—is overblown, as she was simply carrying out the administration’s policy of trying to avoid directly antagonizing Saddam. For advocates of constructive engagement, the policy held out the prospect of turning the regime with the largest army in the Middle East into a responsible member of the international community. This policy had the downside, however, that the United States made no clear deterrent threats, thus giving the Iraqis reason to believe that, although not welcomed, their invasion would not be forcefully opposed.

Even in the event of a ground war, Saddam had reason to think that Iraq might emerge, if far from unscathed, politically victorious. The Iraqi army had gained an incredible amount of experience in the eight-year war with Iran, and the elite Republican Guard units were expected to pose a significant challenge to even well-trained Western units, while the huge size and impressive armament of the Iraqi general army—the fourth-largest in the world, larger than that of the United States—meant that even absent elite training it would pose a formidable obstacle.

Meanwhile, by the time the air war broke out, the Iraqis had had almost six months to reinforce their defensive line in Kuwait with minefields,
oil-filled fire trenches, sand embankments, and other obstacles. In this context, Saddam could quite reasonably believe that any attempt to push the Iraqis out of Kuwait would prove sufficiently costly for the coalition that public opinion would force a premature end to the campaign.

A less-recognized basis for Iraqi overoptimism was the belief that the Saudis would not invite American soldiers to operate out of their territory. Thus, for example, Saddam reportedly ordered the occupation of all of Kuwait, rather than just disputed border areas, in the expectation that doing so would preempt a Kuwaiti invitation to the United States, which in turn would leave the Americans with no place to establish bases. It was in keeping with this expectation that the Iraqis indicated a desire for an Arab solution to the problem, which in practice presumably would have involved significant concessions to Iraq, likely financed by the Saudis and others in the Gulf. In practice, this expectation was undone by standard security dilemma fears: the Iraqis have consistently professed not to have had any plans to attack Saudi Arabia, but even if their intentions were limited the Saudis could not know that for certain, while the massive Iraqi army in Kuwait obviously was capable of continuing its offensive southward. It was this fear that induced King Fahd to accept the significant domestic political costs associated with inviting the United States and other Western powers into Saudi Arabia.

The Evolution of Beliefs

Political disagreements and a difficult financial situation provided Iraq with reason to consider an invasion of Kuwait, and the weakness of the Kuwaiti defenses, the refusal of Kuwaiti leaders to make significant concessions, and a belief in American irresolution and inability to act indicated that the consequences of an invasion would likely be minimal. In this sense, Saddam Hussein had good reason to be quite confident when he ordered the invasion of Kuwait. Subsequent events, of course, would belie that initial optimism. Some bases for optimism had to be abandoned quite quickly: the Saudis took only five days to decide to invite in American forces. Others, however, persisted.

In particular, the belief that the United States was fundamentally irresolute continued to provide a reason for optimism, even as world opinion turned quickly against Iraq. The confidence in irresolution can be seen in Saddam’s prewar diplomatic strategy, which consisted of repeating with great frequency the claim that American resolve would not withstand the first few encounters. These beliefs can only have been strengthened by the internal divisions in the United States, where only a slender majority of the public supported the threat to use force and many in Congress were arguing vociferously for greater time to allow
sanctions to take effect and calling on Congress, in the words of Senator Edward Kennedy, to “stop this senseless march to war.” When former secretary of state Cyrus Vance testified that in the event of war the United States would find itself “virtually alone in a bitter and bloody war that will not be won quickly or without heavy casualties,” he was capturing—and broadcasting for anyone to hear—the sentiments of a large portion of the American public that believed that war would be a disaster. President Bush was willing to go to war even without the support of Congress, but that kind of willingness was impossible to demonstrate credibly before the fact, while the Iraqis could listen to Democratic senators arguing that even a few thousand casualties in the liberation of Kuwait would be too high.

The American leadership was fully aware that Iraq might not take American threats seriously, given the incentives to bluff. Indeed, at the final meeting between leaders on the two sides, Secretary of State James Baker started his comments by observing to Iraqi Foreign Minister Tariq Aziz, “You think we are bluffing.” This task was complicated, however, by the need to maintain the domestic and international coalition for war, which involved demonstrating that all possible efforts had been made to reach a diplomatic settlement. Thus, regional experts reported that Bush’s proposals for last-minute talks between Baker and Aziz—undertaken to demonstrate that all efforts to reach a peaceful solution had been taken—would be interpreted in Baghdad as evidence of irresolution, no matter how frequently Bush insisted that the American position was nonnegotiable.

At the same time, Bush administration officials were growing more confident, with Bush in particular coming to believe that the United States could “knock Saddam Hussein out early.” A critical development was the discovery that, contrary to initial beliefs, the ground on the Saudi-Iraqi border west of Kuwait was firm enough to support tanks and large transport vehicles, thus permitting a ground war plan based on maneuver and envelopment rather than one based on a costly frontal assault on prepared defenses in Kuwait. As a result, whereas in October military and political leaders were reporting that “if war comes, its human, economic, and political costs are likely to be high” and that “there is little prospect of winning a neat ‘victory’ in such a conflict,” by January Bush was operating on the basis of military estimates that put the likely toll below two thousand casualties. Meanwhile, Bush was increasingly convinced that the Iraqi government’s hold on power was tenuous, at one point in his diary comparing Saddam Hussein to Nicolae Ceaușescu, the Romanian communist dictator whose decades-long rule disintegrated in 1989 in the face of a general uprising. Overall, then, American decision makers believed that war might provide significant
benefits beyond the liberation of Kuwait. In this context, as the crisis developed American demands for unconditional withdrawal rose to include other demands such as reparation payments to Kuwait and the weakening or elimination of Iraqi WMD programs.  

Beliefs and Bargaining Positions Once Fighting Began

Ultimately, large shifts in bargaining positions only occurred once the war began, and especially as the Iraqis came to realize that their expectations, and hence their strategies, were flawed. Before the war started, Saddam repeatedly averred that Iraq would withdraw from Kuwait only as part of a general Middle East peace agreement, and that absent that effectively impossible condition Kuwait was now an inviolable part of Iraq. He held to that position even as the deadline for withdrawal approached, apparently believing that American irresolution would manifest itself either in a decision not to attack or, after initial clashes, in a reluctance to pay the costs necessary to evict the Iraqi army from Kuwait. Once the bombing began, he continued with the same basic strategy of trying to split the coalition by linking its activities to Israel—now by launching missile attacks on Israel designed to provoke an Israeli response—while promising to impose unacceptable casualties on coalition forces should an invasion take place.

Diplomacy did not stop once the bombing began, although it took on a new character. Particularly illuminating were repeated Soviet peace efforts in the period immediately prior to and during the war, which induced the two sides to reveal their bargaining positions. These peace efforts foundered initially on Saddam Hussein’s refusal to withdraw from Kuwait on anything approaching the terms contained in prior Security Council resolutions. Even when after almost a month of bombing the Iraqis indicated a willingness to accept Security Council Resolution 660, which called for unconditional Iraqi withdrawal, their acceptance was subject to extensive qualifications that guaranteed American rejection. This position changed with the approach of the ground war, however. In the days prior to the February 24 start to the ground assault, the Iraqi government made an increasing range of concessions, ultimately agreeing to a Soviet cease-fire proposal in which the Iraqis would withdraw unconditionally, removing all forces from the country within three weeks. By this point, however, the success of the air war and the failure of Iraqi attempts to split the coalition increased confidence in the Bush administration, which now concluded that the Iraqi army was caught in a trap. In response, Bush raised American war aims to include not only the liberation and restoration of Kuwait but the substantial weakening of the Iraqi army. In line with this goal, he stated that peace was on offer
only if the Iraqis would agree to complete withdrawal within one week, which would have required the abandonment of almost all of the army’s heavy equipment. As a result, substantial Iraqi concessions—to terms that would have guaranteed peace prior to the start of fighting—were no longer sufficient to purchase a cease-fire. At the same time, however, the vast disparities between the two sides’ demands prior to fighting had narrowed substantially.

Peace ultimately came only after the four days of ground war brought about the weakening of the Iraqi army that the Americans sought. The flanking of the Iraqi army in Kuwait precipitated that force’s headlong withdrawal into Iraq, during which the coalition air forces were able to annihilate any retreating heavy forces, although key Republican Guard units, which never had been deployed to the front lines, were able to escape back toward Baghdad. At this point, Bush declared victory, announcing a unilateral cease-fire that the Iraqis were only too happy to accept. In practice, the specific terms of that cease-fire were worked out in a meeting between opposing generals at Safwan on March 3, after which the Iraqis begrudgingly acquiesced to additional Security Council Resolutions that codified the political terms of Iraq’s defeat, including the establishment of no-fly zones over southern and northern Iraq in which the Iraqi air force was not permitted to operate and the creation of an intrusive monitoring regime to identify and eliminate Iraq’s chemical, biological, and nuclear weapons programs.33

A separate but important question is why, given the success of the ground war, the Bush administration ruled out an escalation of the war to ensure regime change in Baghdad, a possibility that was discussed both at the time and subsequently. The decision to stop the war with Saddam still in power arose primarily from two concerns: that the coalition would collapse if the United States were to attempt to move beyond the mandate provided by the Security Council, and that an attempt to overthrow Saddam would involve the American forces in an indecisive hunt through hostile regions of Iraq that would expose them to urban guerrilla warfare that would inflict the high level of casualties that a more restrained military strategy had managed to avoid.34 Thus, having attained their primary political aim and having furthermore substantially weakened Iraq, the Bush administration saw no further goals that could be achieved at acceptable cost.

After six months of inflexibility, then, both sides’ bargaining positions shifted quite dramatically once the bombing started, and especially once the fighting intensified with the move toward the ground war. These developments are entirely consistent with the predictions of the informational mechanism: while credible signaling was difficult prior to the outbreak of costly fighting, the course of events once fighting began
demonstrated that Iraqi expectations were profoundly mistaken and thus justified a dramatic shift in political demands. This process occurred quickly: even before the ground war rendered a continued Iraqi occupation of Kuwait impossible, Saddam had indicated his willingness to concede the central issue at stake in the war. Consistent with theoretical expectations, in a war in which divergent expectations led both sides to anticipate a relatively quick victory, only a relatively brief period of fighting was necessary for the two to identify a settlement that they preferred to continued war. Moreover, demands shifted more rapidly as the war intensified, with the shift to the ground war rapidly bringing about near-total Iraqi capitulation despite the survival of the elite Republican Guard units. Overall, then, we have a case in which divergent expectations about the likely course and outcome of conflict led to a war, but once that war began both expectations and bargaining positions shifted rapidly in a manner that allowed for quick settlement.

The Secondary Significance of Alternate Mechanisms

In contrast to the strong evidence of the relevance of the informational mechanism, the commitment and principal-agent mechanisms were of at best secondary significance. One possible commitment problem argument would be that the incorporation of Kuwait would increase Saddam’s capabilities to such an extent that he would be able to engage in further aggression at a later date. While Bush administration figures certainly worried that Iraq might engage in further aggression, in particular with a further strike against Saudi Arabia, they did not express concern that control over Kuwaiti oil would facilitate such aggression. Indeed, the most worrying scenario was an attack on Saudi Arabia immediately after the initial invasion (before coalition forces had appeared in the region), before Iraq would have had an opportunity to turn Kuwaiti oil into a strategic resource. Instead, intervention was undertaken primarily because of the economic cost of allowing the consolidation of oil production under hostile governments and because of a sincere belief that permitting aggression to succeed would undermine international stability.

A separate argument about shifting power concerns the American insistence that Iraq abandon its heavy armor in Kuwait as a condition for peace. This requirement clearly was designed to weaken Iraq, and Saddam’s reluctance to agree even after he had conceded the central political issue—Kuwait’s independence—indicated that it was precisely this anticipated shift that prevented a settlement prior to the ground war. In chapter 1, however, I noted a distinction between beliefs in long-term decline that are a feature of preventive wars and these sorts of intrawar military shifts, as with the anticipated weakening of
the German Wehrmacht after the encircling of the Sixth Army at Stalingrad. The anticipated destruction of an entrapped force certainly may hinder settlement, but it will not produce the dynamics that make preventive wars difficult to resolve over an extended period of time, largely because the anticipated shift is likely to happen quickly. Thus in this case it took only four additional days of fighting for the Iraqi army to suffer the losses that it would have incurred through acceptance of American terms, after which this problem ceased to exist as an obstacle to settlement. It is for this reason that I argue that these sorts of intrawar shifts generated by anticipated military defeat do not invoke the broader dynamics of preventive wars.

A principal-agent interpretation is even weaker. As an absolute autocrat, Saddam worried about the possibility that he might be violently overthrown, either by his own army or by a Shi’a uprising, likely with Iranian backing. Indeed, he took extensive precautions against such a possibility throughout his reign—the steps that he took to forestall overthrow by the military, including executing prominent generals and limiting the military’s training in urban warfare, rendered his regime almost totally coup-proof, while a repressive internal system prevented more general uprisings. He thus had no reason to use a war to cement his hold on power; indeed, defeat in war was one of the few developments that might significantly have threatened his hold on power, as the Shi’a and Kurdish risings after the defeat in Kuwait demonstrated.

Likewise, it is hard to sustain a diversionary argument about the coalition response. While disagreement about whether the war was worth fighting existed, few denied that Iraqi control over Kuwait was contrary to the interests of the United States or the international community more generally, and critics of the war argued primarily that its costs would likely be unacceptably high, something that the Bush administration, with better information, was confident would not be the case. While Bush’s popularity did rise significantly as a consequence of Desert Storm, there is no evidence that he chose to respond to Iraq because he believed that it would improve his hold on power, nor has there been any evidence of substantial misrepresentation of the situation in the Middle East or the likely outcome of a war.

Summary

The Persian Gulf War thus follows the classic pattern of a conflict driven by private information and divergent expectations. Saddam Hussein underestimated the both the resolve and the ability of the international community, and in particular the United States, to respond effectively to the Iraqi invasion. Conciliatory diplomacy prior to the
invasion and an apparent pattern of casualty aversion gave Saddam the impression that the United States would not be willing to use force to reverse the conquest of Kuwait, and even if they were he doubted that the Saudis would permit Western forces to operate from their country. While some of these expectations were disabused over the course of the crisis, until the ground war Saddam remained convinced that the Iraqi army would impose high enough casualties on coalition forces to force a political retreat. It thus took a war to convince Saddam that his expectations were incorrect. Once the war began, however, the errors in his expectations rapidly became apparent, at which point he revised both his expectations and his demands downward, facilitating a quick political settlement.

More than a century before the Persian Gulf War, soldiers from the West found themselves in the area of the Persian Gulf fighting to preserve the autonomy of a small but important region. As with the Iraqi invasion of Kuwait, the crisis created by the Iranian (Persian) takeover of the Afghan city of Herat culminated in a short war that restored the initial status quo. Given the minimal consequences of the war—the deaths of a few thousand soldiers, but no political change—the war has left little imprint on the historical record. While information on this conflict is limited, a clear picture nonetheless emerges of a war driven by overoptimism, in which the information gained through fighting facilitated a quick negotiated settlement, even in the face of great obstacles to negotiation.

An Overview of a Little-Known Clash

The central stake in the war was the city of Herat, which, given the 1824 division of Afghanistan by three rival princes into the principalities of Herat, Kabul, and Kandahar, had been effectively independent for several decades. This situation presented an opportunity for Iran to re-establish control over Herat, which had belonged to Persia under the previous Safavid dynasty, and two separate shahs attempted to do so in 1838 and again in 1852. These incursions, however, inserted the country more centrally into the Great Game, the Anglo-Russian rivalry in Central Asia. In that contest, the British attempted to ensure that local leaders in Afghanistan, through which an invasion of India would have to come, remained friendly; they were in particular opposed to an increase
in Iranian control because of a belief that the Russians, who posed a significant threat to Iran, exerted excessive influence in Tehran. Thus, in both 1838 and 1852 the British responded to Iranian moves with a vigorous diplomatic effort that ultimately convinced the shah to withdraw his forces and agree to treaties repudiating any claim of sovereignty over Herat.

The issue returned to the fore in 1856, however, when an internal Afghan dispute led the prince of Herat to request Iranian aid. With this request as a pretext, the Iranians then moved in and, despite facing significant local opposition, established control over the city in October 1856. With negotiations between British and Iranian representatives in Constantinople dragging on, the British government in London authorized the governor-general in Bombay to declare war, which he duly did on November 1.

Subsequent military developments can be summarized quickly. Given the disasters that had accompanied previous incursions into Afghanistan—including an extended expedition from which only one British soldier escaped death or capture—the British were reluctant to intervene directly in Afghanistan and instead opted to land troops along the Persian Gulf coast. After doing so in December 1856, they seized the important trading center of Bushehr and then moved inland, where they defeated a significant Iranian force outside the town of Khushab in the largest battle of the war. Declining to pursue the defeated army inland, the British then moved up the Euphrates River, taking the city of Mohammerah (now Khorramshahr). On April 4, a few days after the fall of Mohammerah, plans for a significant incursion into the Iranian interior were interrupted by the arrival of news of a peace agreement signed in Paris a month earlier.39

Thus, less than four months passed from the British landing to the news of the peace agreement (and less than three from the start of fighting to the political settlement); while the British had had the best of it, at the point that the war ended, little fighting had occurred.40 Nothing that the British had done directly compelled the Iranian army to withdraw from Herat. Why then did the shah agree to give up something that, as the repeated attempts over several decades to subvert Herat’s independence clearly demonstrated, was quite important to him? The evidence in this case clearly indicates that the shah and his advisers underestimated British resolve—there existed a number of plausible reasons for them to do so—and thus adopted a policy that ran a risk of war with Britain were the British actually resolved, as they turned out to be. Consistent with the predictions of the informational mechanism, settlement came quickly, with the war ending basically as soon as the
two sides could meet to negotiate and convey the peace agreement to the armies.

Private Information and Divergent Expectations

In contrast to the situation in many wars, the Iranians were not particularly optimistic about their military prospects in the event of a direct clash with the British.\(^4^1\) They had reason, however, to doubt that such a clash would actually occur. While the British had established a clear interest in the fate of Herat (and the region more generally) in the past, several developments had provided reason to doubt the continued strength of that commitment.

The most important development was the annexation during the 1840s of the Sikh regions and the Punjab, which together provided a more natural and easily defensible northwest border for India.\(^4^2\) Given this development, a reasonable observer might wonder whether the British would still fight for Herat, as the loss of Afghanistan to Russian control (no certain prospect) was now less threatening. Indeed, at the outset of a renewed Iranian attempt to annex Herat in 1851, Prime Minister Palmerston wrote to the ambassador in Tehran to say,

> You will, therefore, still endeavour to dissuade the Persian Government from advancing on Herat, and you may truly say that such a move would not be viewed with indifference by HM’s Govt.; but you will be careful not to make any specific threat which HM’s Govt. might not be disposed afterwards to carry into execution.\(^4^3\)

While this communication obviously was not conveyed to the Iranians, it was a reasonable response to a changed strategic situation, which the Iranians undoubtedly had noticed. This particular incident ended with the shah pulling out of Herat, which the Iranians had successfully taken, in response to a forceful British remonstration. However, the British never actually mobilized troops, meaning that they did not actually demonstrate any willingness to spend blood and treasure to keep Herat independent. The shah might therefore still think that they would acquiesce in a seizure, especially if it was timed to coincide with British troubles elsewhere.

At the same time that the new frontier decreased Britain’s incentive to fight for Herat, events in the disastrous First Afghan War—in which a British attempt to install a more amenable potentate in Kabul miscarried badly, with a significant British force utterly destroyed during an attempted retreat back to India—provided reason to expect the British to be reluctant to intervene directly in Afghan affairs.\(^4^4\) Indeed, British
Short Wars of Optimism

officials clearly were influenced by this precedent, with the governor-general in India (Lord Canning) sending a message back to London during the Herat crisis specifically to advise against any strategy that involved incursions into Afghanistan. The British certainly still had tools at their disposal, as the invasion of the Gulf coast demonstrated, but those tools provided a less direct way of securing the independence of Herat.

Moreover, if the more secure border and the difficulty of intervening directly in Afghanistan provided a reason to think that when distracted the British might let the Iranians get away with seizing Herat, then the Crimean War appeared to provide the necessary distraction. While the British and their allies ultimately won the war, they did so only at high cost and in a less than completely decisive manner. Indeed, by some accounts far more was known in Asia about the Russian victory over the Turks at Kars than was known of the Russian defeat at Sevastopol; the Iranians apparently started their expedition with the belief that the British would have no spare troops to use against them. Under these circumstances, Iranian decision makers could reasonably believe that the British might decide against diverting strength from a major conflict to prevent the seizure of a remote city whose strategic value had declined significantly. The end of the war with Russia removed this obstacle to British action, but unfortunately by the time that the Iranians realized their error and sought to correct it, events had proceeded too far, and they were unable to carry the negotiations through before the British, thinking that the Iranians would not back down without a fight, had committed to war.

Thus, in October 1857, after the capture of Herat, the shah’s primary adviser assured him that Britain would never resort to “coercive measures” and “would not move a soldier or a ship in connection with Herat.” It is thus unsurprising that the shah “appeared incredulous” when told that the British might occupy Bushehr, and that both he and his primary adviser expressed surprise on receiving the British declaration of war. The arrival of war, however, and especially the demonstration—through the occupation of Bushehr and the battle at Kushab—that the British were sufficiently resolved to fight forced a major revision in Iranian expectations. Initial clashes demonstrated that Iran lacked the manpower, money, or morale necessary to repel Britain; of the defeat at Kushab Amanat observes that “nothing could have more forcefully enhanced the image of British invincibility in the eyes of the demoralized Persian government.” The revision in expectations is perhaps clearest in the shah’s note to his representative in Istanbul: “Of course! Of course! Swiftly settle the matter in whatever manner you deem advisable. Do not allow it to come to severe hostility.”
From this point, all that stood in the way of peace was the time that it took for communication to occur. Indeed, the governor-general in India, whom London had empowered to act, declared war ignorant of negotiations in Constantinople in which differences between the parties had narrowed. After a brief delay when the news of the declaration arrived, negotiations moved to Paris, where an agreement was signed on March 4. The final agreement was quite close to the prewar British position, most obviously in Iran’s agreement to withdraw from Herat, although they abandoned some tangential demands. The extent of updating is clear, however, in the cheerfulness with which the shah greeted the news of the final terms, which although forcing him to back down on the central issues in the war at least permitted him to retain his throne. The Iranians duly ratified the treaty and withdrew their army from Herat in short order.

To summarize, the Iranians saw a favorable opportunity to try to accomplish a fait accompli because they believed that when pushed the British might not prove willing to intervene. In this case, British resolve was private information, and because it was private (and not readily conveyed in a credible manner) it was possible, and indeed reasonable, for the Iranians to underestimate it. By contrast, there was relatively little divergence in expectations about the likely victor once war began. Thus, on the central issue leading to war, updating of beliefs occurred quite quickly, and the Iranians rapidly scaled back their ambitions, withdrawing from Herat (despite the fact that they had already captured the city and that the British lacked the means to force them out directly) based on the expectation that further fighting would only increase costs without improving the odds of political gains. These results thus are quite consistent with theoretical expectations: private information in this case could be effectively revealed only once the British demonstrated their willingness to fight, at which point the Iranians revised their expectations, lowered their political demands, and got out of the war, all in a relatively short period of time.

The Limits of Alternate Mechanisms

Neither of the other two primary mechanisms provides as convincing of an explanation for this case. A commitment problem explanation runs into the fundamental problem that no source that I have found records comments from anyone involved to the effect that war now would prevent a worse conflict at a later date. The British did wish to maintain the Afghan buffer against the possibility of Russian encroachment, but no one argued that the loss of Herat would put the Russians in a position to invade India or to demand concessions with respect to the subcontinent.
A more convincing interpretation of British policy is instead that they were interested in maintaining a clear reputation for protecting strategic interests in Central Asia—a view, given the centrality of private information to reputation, that is more consistent with an informational interpretation. Moreover, if one did believe that the war was motivated by a commitment problem, it would be hard to understand the course of events. Once the war began, the British did not directly relieve Herat, and the damage that they did to the Persians along the Gulf coast was hardly crippling. In other words, nothing in the military events on the battlefield would have resolved a significant commitment problem had it been contributing to the war.

At the same time, the case does demonstrate that mere suspicion of an opponent’s intentions or reliability is hardly grounds for war without end. The Iranians had repeatedly violated agreements to respect the independence of Herat (most notably in 1838 and 1852) when they believed that they might be able to get away with doing so. Indeed, contemporary British sources consistently depict the Iranians as duplicitous and untrustworthy. In response to this duplicity, however, the British did not launch a crusade to remake the Iranian polity; rather they acknowledged that they would need to make clear their willingness to intervene to the Iranians to deter further interventions. From this perspective, it is clear that simple mistrust associated with international anarchy is not sufficient to produce a major war.

Principal-agent dynamics likewise provide an unconvincing explanation for the war, although developments within the conflict provide some support for ancillary principal-agent hypotheses. In both Britain—at this point a partial democracy—and completely autocratic Iran, leaders had little to gain personally from fighting. In Iran, the Qajar Dynasty had developed an institutionally autocratic political system that gained legitimacy and internal cohesion from a system of effective negotiation and consensus. Thus, despite (or because of) the existence of significant external threats, the shah did not have to worry greatly about his hold on power, at least so long as foreign powers like Britain did not attempt to unseat him. In Britain, Prime Minister Palmerston had gained broad esteem through his handling of the initially badly mismanaged Crimean War, and thus stood to gain little from further foreign adventures.

That said, there is some indication that in fighting the war Palmerston—who could be described as an early advocate of liberal internationalism—was pursuing a policy that a substantial fraction of the British public did not endorse. News of the war reportedly was greeted in England “with a mixture of disgust and derision, and with anti-war demonstrations in Bradford and Newcastle”; the Morning Star referred to the war as a “Don Quixote campaign.” Moreover, the government was
also under fire for its conduct in the Second Opium War (aka, the Arrow War), which began concomitantly. Given this situation, Palmerston used, and most likely abused, some of the prerogatives of the prime minister. Parliament went out of session in July 1856, while the dispute in Persia was only just getting started, and only returned in February, by which time the war was well underway and there was little the opposition could do. Indeed, the government adopted the time-honored strategy of arguing that debate on an ongoing war would be contrary of the interests of the nation, arguing that it would be pointless to hold an open debate on the war without the relevant papers being available, but then refusing to release those papers because of the ongoing negotiations with Iran. As a result, the opposition was limited in its criticisms to what could be gleaned from letters that soldiers sent home, which arrived long after the events they described had occurred. Nonetheless, the significant criticism that the government received was an indication of the limits to Palmerston’s ability to engage in international adventures, especially when one considers that his government collapsed roughly contemporaneously over its handling of the Arrow War with China.

Summary

As with the Persian Gulf War, the course of the Anglo-Iranian War is best explained by the informational mechanism. Although it was clear that the British would win in a military conflict, the Iranian government had reason to think that they might not be sufficiently resolved to fight one, given the stronger defense of India provided by the acquisition of the Sikh regions and the Punjab, British disinclination to involve itself directly in Afghanistan after the First Afghan War, and the apparent distraction provided by the Crimean War. In practice, however, this calculation proved to be incorrect: the Crimean distraction was gone by the time that the crisis over Herat had reached its peak, while the British found a military strategy that did not force them to fight in Afghanistan. Every indication is that the Iranians were prepared to back down as soon as it became apparent that the British were willing to fight—had faster communication between the adversaries been possible, it is unlikely that the war would have lasted more than a month, had it occurred at all. By contrast, the commitment problem mechanism appears to have been completely absent, while the principal-agent mechanism played at best a subsidiary role.

The Persian Gulf and Anglo-Iranian wars are not typical of the wars that scholars and policymakers tend to focus on, but they are typical of the wars that actually occur: both were short and not particularly deadly, at
least by the standards of interstate wars, and both ended well before the
loser’s final military defeat, despite the potential for further escalation
and a longer and deadlier war. In the Persian Gulf War, the coalition
elected not to take advantage of initial victories to march on Baghdad; in
the Anglo-Iranian War the politicians brought an end to the war just be-
fore the British force moved away from the coastline to launch a major
incursion into the interior. In both cases, the best explanation for both the
onset of the war and its short duration is provided by the informational
mechanism. Leaders who initially miscalculated about their opponents’
capabilities and resolve revised their expectations in response to the
fighting and reduced their political demands accordingly, facilitating
settlement. These findings thus accord with the prediction that informa-
tional wars, although prevalent, tend to be limited.

Equally important, the absence of significant commitment concerns in
either case is consistent with the prediction that commitment problem
wars will typically be more destructive. In neither case did leaders refer
to concerns about shifting power in the way we observed in previous
chapters, nor were there good reasons for them to worry about such con-
cerns. This observation allays any potential concerns that commitment
problem concerns are simply omnipresent in wars, in which case they
would have provided a less convincing explanation for war destructive-
ness. Instead, it appears that these concerns are present only in the most
severe of wars.