CHAPTER 1

We Have No Quarrel with the People

The United States has no quarrel with the Iraqi people.
—*U.S. President George W. Bush,*
September 12, 2002

DESPITE THIS AND OTHER declarations of friendship, President Bush ordered U.S. troops to invade Iraq on March 20, 2003. Dropping bombs is an unusual way to express felicitations. Yet, as Bush stated, U.S. anger was directed toward the political leadership of the Iraqi government, not the people themselves. British Prime Minister Tony Blair in a November 2, 2002, speech was even more explicit: "[W]e have absolutely no quarrel with the Iraqi people. We want you to be our friends and partners in welcoming Iraq back into the international community." The targets of foreign policies are often political leaders rather than the nations they represent. This book explores the implications of, what we shall call *leader specific punishments.*

We assigned the initial quotation to U.S. President George W. Bush. However, by simply substituting nationalities we might equally well have assigned the quotation to many recent presidents, be it Ronald Reagan discussing Libya, George H. W. Bush discussing Iraq, or Bill Clinton discussing Yugoslavia. Robert Fisk (2002), in an article for the *Independent* newspaper, describes the statement "we have no quarrel with the people of . . ." as "the mantra that means this time it's serious."[1]

One might argue that Bush's targeting of leaders is nothing more than rhetoric that makes his action more palatable to domestic and international audiences alike. After all, whether U.S. policies were targeted against Saddam Hussein or Iraq more generally, it was still the Iraqi people who suffered the loss of loved ones, their homes, and their livelihoods. Yet, we shall argue that leader specific punishments have a profound and, at times, surprising impact on the dynamics of interstate relations.

---

[1] More flippantly, in their list of categorized quotations, the Web site www.righteouswarrior temple.org describes this phrase as "a well known presidential code-phrase, used many times in the past, which roughly translates as 'We're about to bomb your monkey asses into the Stone Age.'"

From the perspective of the nation issuing the threat, leader specific punishments have two important (and related) properties. First, by targeting a specific leader, rather than the nation as a whole, leader specific punishments identify an end to sour relations and an opportunity to rejuvenate good relations. Commitments, to impose sanctions for example, are not open ended. They last only as long as the recidivist leader remains in power. Targeting leaders provides a mechanism to restore good relations. In comments to the BBC World Service, July 20, 1999, on the eve of the Kosovo conflict, NATO's Supreme Allied Commander in Europe, General Wesley Clark, stated, "[I]t is a real political problem for the people of Yugoslavia because I think world leaders have made very clear that they don't see Yugoslavia really being readmitted into the European Community of nations or receiving the kinds of reconstruction that it really needs while he's [Milosevic] still in place as the President." General Clark was correct in his assessment. Following the deposition of the Serbian president Milosevic, economic assistance flooded into Yugoslavia.[2] When punishments are leader specific, leader turnover ends the punishments.

Acrimonious relations often end with leader change. This provides nations with an opportunity to start afresh. After years of failed attempts to negotiate a settlement with the Palestinians, Prime Minister Ariel Sharon explained in a speech before the Israeli Knesset on April 8, 2002, that the Palestinian people were not the problem. Rather, he argued, the obstacle to peace was the Palestinian leadership, which consistently showed itself unwilling or unable to maintain agreements. "We have no quarrel with the Palestinian people and we want to see the Palestinians, like us, live in peace, security and dignity. . . . But peace can only be attained if, once we evacuate the territories, we find a responsible Palestinian leadership, willing to accept the primary responsibility of every regime—to prevent the use of its territory for the purpose of killing and murdering its neighbors. Peace negotiations can commence and move forward only after terrorism has ceased." New Palestinian leadership is needed if the Israelis and Palestinians are to move beyond past recriminations and start constructive negotiations.

Second, leader specific punishments create internal political cleavages within the targeted state. Threats against a nation often create internal cohesion, a phenomenon often referred to as the "in-group, out-group" effect (Coser 1956). In contrast, leader specific punishments partly mitigate the risk of interstate relationships descending into a feud, by creating internal divisions. Since leader turnover normalizes relations, the citizens in the targeted state can end the punishment by deposing their

---

[2] See for example, "Aid Talks after Milosevic Drama," CNN.com, June 29, 2001.

leader.[3] Whether or not leader specific punishments lead to the removal of the targeted leader depends, in part, on how difficult it is to overthrow the leader. We study how the costs of leader removal affect the effectiveness of leader specific punishments. The *New York Times* argued in its discussion "Were Sanctions Right?" (February 28, 2003) that "[b]y making life uncomfortable for the Iraqi people, [sanctions] would eventually encourage them to remove President Saddam Hussein from power." Leader specific punishments encourage citizens to depose their leader, as it triggers a restoration of cooperative relations. British Prime Minister Neville Chamberlain, in his September 3, 1939, speech before Parliament declaring war on Germany at the start of World War II expresses amicable relations with the German people and the view that regime change within Germany would remove all need to resort to war: "We have no quarrel with the German people, except that they allow themselves to be governed by a Nazi Government. As long as that Government exists and pursues the methods it has so persistently followed during the last two years, there will be no peace in Europe."[4]

LEADER SPECIFIC PUNISHMENTS AND INTERSTATE RELATIONS

Although in motivating the topic above, we discussed largely conflictual events, such as war, the impact of leader specific punishment is just as relevant in explaining the everyday economic, financial, and diplomatic interactions between nations. Indeed, for most of the book, we focus on cooperative interactions betwveen states. The theory we develop examines the interplay between individual leaders, political institutions, and interstate relations. We articulate some of the main insights and derive a simple exposition of the theory in the context of the prisoners' dilemma.

Consider a simple example of nations wanting to establish norms of cooperation and trust between themselves in order to provide mutual benefits for both sides. Although both nations are better off if they cooperate,

---

[3] In a related argument with respect to interethnic cooperation, Fearon and Laitin (1996) describe how a combination of between-group and within-group punishments best maintains intergroup cooperation. While the majority group might easily threaten to punish the minority group (an intergroup punishment), the same threat has much less bite for the minority group. Instead, Fearon and Laitin argue that the minority group should internally punish those members of its group that cheat members of the majority in order to maintain good relations with the majority.

[4] Russett (1993) quotes U.S. President Woodrow Wilson, who also expressed that the United States had no quarrel with the German people on April 2, 1917 during the First World War.

each side could make itself even better off if it allowed the other nation to make the greater contribution to the common good. The incentive for each nation to renege on its contribution makes cooperation difficult. The standard Liberal approach, a literature we shall discuss in detail later in this chapter, explains the evolution of cooperation via the use of reciprocal punishment strategies (see for example Keohane 1984 and 1986). For instance, if nation A threatens to withdraw all future cooperation if nation B cheats, then provided that nation B values long-term cooperation more than the myopic gains from cheating, such a threat is sufficient to sustain cooperation. Following Liberal arguments we shall develop these ideas within the context of an infinitely repeated prisoners' dilemma game.

Cooperation evolves when nations choose reciprocal punishment strategies. Yet while treating nations as unitary actors is a convenient device, it is political leaders and not some personified nation that choose foreign policies. Suppose, therefore, that instead of directing reciprocal punishment strategies against a foreign nation, leaders implement strategies against the opposing leader that cheated them. That is to say that once the leader in nation B cheats, the leader of nation A refuses to cooperate with this leader ever again. However, since the punishment is directed against a specific leader, once that leader leaves office, nation A will restore cooperation with the new leadership in nation B (who after all has never cheated nation A). The replacement of a leader who previously cheated rejuvenates interstate relations.

Leader specific punishments enable the citizens of a nation to avoid punishment by simply replacing their leader if she cheats. Whether the citizens choose to do so, however, depends upon domestic political institutions and in particular how these institutions shape the cost of replacing a leader. If the value of restored cooperation exceeds the cost of leader replacement, then the citizens depose their leader if she cheats. Under institutional arrangements that make it difficult to replace leaders, however, the benefits of restored cooperation are too small to justify the high cost of leader replacement.

The Prisoners' Dilemma

We now formalize these arguments using the standard metaphor for international relations, the prisoners' dilemma. This game, shown in figure 1.1, captures the inherent difficulties of international cooperation. In each period, nations A and B choose between cooperate (C) and defect (D) and the payoffs are such that $T > R > P > S$. Nations have a dominant strategy to play defect, since whether nation B plays C or D, nation A's payoff is improved by playing D. This results in the noncooperative outcome of

| | | Nation B | |
		Cooperate, C	Defect, D
Nation A	Cooperate, C	R, R	S, T
	Defect, D	T, S	P, P

Figure 1.1. The prisoners' dilemma. $T > R > P > S$ and $R \geq (T + S)/2$.

(D,D). Yet both nations improve their payoff if they mutually cooperate. Unfortunately, once one nation cooperates, the other has the incentive to exploit their cooperation by defecting to obtain the temptation payoff, T. It is this mix of mutual gains from cooperation and incentives to cheat that has made the prisoners' dilemma such a powerful metaphor for international interactions.

In the single-shot game the prospects for cooperation are dismal. Yet, through the use of reciprocal punishments, in which nations condition their willingness to cooperate on past behavior, cooperation is possible provided nations are sufficiently patient. Patience is measured using the discount factor δ ($1 > \delta > 0$), which states the proportionate value of having to wait until the next period to receive a payoff. When δ is high, nations are patient and discount future payoffs relatively little. In contrast, δ is low for impatient nations who strongly discount the value of future payoffs.

We start our exposition of how mutual cooperation can be maintained between unitary actor nations through reciprocal punishments by considering the Grim Trigger (GT) strategy. Afterward we will adapt this strategy to explain the logic of leader specific punishments. In the GT strategy each nation starts cooperating and continues to do so in every future period unless either nation ever defects. Once either nation plays D, nations refuse to cooperate in all future periods. The GT has several advantages for illustrating how cooperation can be fostered through reciprocal punishment strategies. First, it provides the simplest illustratation of how the threat to withdraw future cooperation induces cooperative behavior. Second, this strategy is a limiting case. Since the threat to withdraw cooperation permanently is the harshest threat that a nation can make, if this threat is insufficient to support cooperation, then cooperation is impossible. Third, it is straightforward to mathematically show how the strategy shapes incentives to cooperate and to derive the limits of cooperation, as we shall now show.

If both nations play the Grim Trigger strategy then they cooperate in every period and so receive the payoff R in every period of the game. The net present value of this stream of payoffs is

$$R + \delta R + \delta^2 R + \cdots = \sum_{t=0}^{\infty} \delta^t R.$$

An extremely convenient mathematical result is that the value of this infinite sum of payoff equals $R/(1 - \delta)$. In the immediate period nation A could improve its payoff by defecting (D). However, if nation B is playing GT, then this ends all future cooperation. Therefore the net present value of playing defect is

$$T + \delta P + \delta^2 P + \cdots = T + \sum_{t=1}^{\infty} \delta^t R = T + \frac{\delta P}{1 - \delta}.$$

Nations can only commit to cooperate if the value of future cooperation relative to immediate rewards is sufficiently high. Consistent with standard approaches we can express this by finding the minimum discount factor such that maintaining cooperation is each nation's preferred option, that is,

$$\frac{R}{1 - \delta} \geq T + \frac{\delta P}{1 - \delta} .$$

If this condition holds, the GT strategy is a subgame perfect Nash equilibrium. This result implies that cooperation is possible if nations are sufficiently patient,

$$\delta \geq \frac{T - R}{T - P} .$$

Modeling Leader Specific Punishments

Although it is convenient to personify nations, it is national leaders, and not nations, who set foreign policy. We consider a simple principal-agent structure within each nation. Nation A is composed of leader α and citizens (a). Leader α sets policy that, in the context of the prisoners' dilemma exposition of international cooperation, means choosing between C and D. Nation B is led by leader β, who chooses whether to cooperate or defect on behalf of nation B. Having observed the outcome of the prisoners' dilemma interaction, the citizens can replace their leaders at cost K_A and K_B, respectively. The game is shown in figure 1.2.

In addition to receiving the payoffs associated with the outcome of the prisoners' dilemma (that is T, R, P, or S), leaders receive a payoff of Ψ for each period they are a leader and citizens pay the costs K_A or K_B if they decide to replace their leader. To reflect our belief that leaders are primarily office seeking, we assume the reward for office, Ψ, is large relative to payoffs from the prisoners' dilemma. After deposition, leaders become ordinary

Leader β

		C	D
Leader α	C	R, R	S, T
	D	T, S	P, P

Figure 1.2. The prisoners' dilemma game between representative leaders. Step (1) Leaders α and β choose Cooperate (C) or Defect (D) in the prisoners' dilemma game. Step (2) The citizens in nations A and B decide whether to replace their leaders at costs K_A and K_B, respectively.

citizens and receive the payoffs from PD only.[5] For technical convenience we assume there is an infinite pool of alternative leaders.

Domestic political institutions shape the ease of leader replacement. In democratic systems, deposing a political leader is relatively costless; citizens need only vote for the challenger rather than the incumbent. In autocratic regimes, deposing leaders is much more costly. In chapter 3 we use Bueno de Mesquita and his colleagues' (2003) selectorate model of domestic competition to examine how domestic political institutions shape the policy incentives of leaders and how this in turn affects the ease of political survival. For the time being, we distinguish between political regimes only in terms of the cost of replacing a leader, K_A and K_B, and, for ease of language, refer to low replacement cost regimes as democracies.

The Leader Specific Grim Trigger strategy (LSGT) utilizes the reciprocal punishment strategies of GT, but conditions punishments at the level of leaders. If leader α plays the LSGT strategy, then initially she plays "cooperates" in the PD. Indeed she will continue to play cooperate provided that neither she nor the current leader of state B (β) has ever cheated. If, however, leader β ever cheats, then leader α will never cooperate with her again and plays D in all subsequent periods. Leader α conditions her punishment strategy against the specific leader that cheated her and not the nation she represents. If incumbent leader β is replaced, then leader α returns to cooperating with β's successor.

The key conceptual distinction between LSGT and the unitary actor GT strategy described above is that the LSGT conditions punishment—that is, the refusal to cooperate in the future—against the specific leader who cheated and not against the nation she represented. Below, we specify the LSGT for leader α. Leader β's strategy is analogous. In this description we

[5] Goemans (2000a, b) argues that many deposed leaders are killed or punished following deposition. He further argues the probability of punishment differs by regime type with democratic leaders least likely to be punished and autocratic leaders most likely to be punished. The prospects of postdeposition punishment further enhance leaders' officeholding motivations.

use the term leader i "cheats" to mean that leader i plays D while leader j plays C.

The Leader Specific Grim Trigger (for leader α)

1. If β, the current leader in state B, has ever cheated, then α plays D.
2. If leader α has ever cheated, then α plays D.
3. Under all other contingencies, α plays C.

Part (1) of this definition indicates that α uses the reciprocal punishment strategy of refusing to cooperate if the current leader β has cheated. Whether prior leaders in nation B have ever cheated is immaterial with respect to α's punishment decision. It is important to note that under the LSGT leader β need not have actually cheated against the current leader in nation A, but just have cheated some leader of nation A. For instance, although Cuba's Fidel Castro "cheated" during the Eisenhower administration by nationalizing U.S. interests in Cuba, all subsequent U.S. administrations recognize Castro's regime as having previously cheated.

We now turn to examining the impact of leader specific punishment strategies on the relations between nations and how the introduction of a leader specific component to the strategy affects the possibility of interstate cooperation. The analysis separates into two distinct cases depending upon the cost of leader replacement. When the cost of leader replacement is high, in particular K_A, $K_B \geq (R - P)\delta/(1 - \delta)$, then the citizens never replace their leader whatever the state of relations between the nations. Under these conditions, behavior is equivalent to the unitary actor GT case. Therefore, as in the GT case, cooperation is possible only if nations are sufficiently patient that the value of maintaining cooperation outweighs the short term gains from defection:

$$\delta \geq \frac{T - R}{T - P} \ .$$

If leaders play the LSGT strategy and the cost of leader replacement is low, specifically K_A, $K_B \leq (R - P)\delta/(1 - \delta)$, then the citizens replace any leader who cheats. Remember that under the LSGT, leader α only refuses to cooperate with the specific leader who cheated her nation; she will cooperate with this leader's successors. If leader β cheats, then the citizens in nation B can end the punishment phase (that is, noncooperation) by replacing leader β. This desire to replace cheaters in order to restore cooperation helps prevent cheating in the first place, since leaders do not want to be removed from office. We formally state the conditions under which LSGT is an SPE (subgamee perfect equilibrium) with low leader replacement cost. We explain the logic of the argument via the process of proving the following claim.

Proposition 1.1: *In the infinitely repeated PD between representative leaders, if $K_A \leq (R - P) \delta/(1 - \delta))$ and $K_B \leq (R - P) \delta/(1 - \delta))$ and $\delta \geq (T - R)/(T - R + \Psi)$, then leaders α and β playing the LSGT and citizens deposing any leader who cheats or has cheated in the past is an SPE.*

To prove that the above constitutes an SPE requires showing that under every possible contingency, for every player, playing these strategies is a best response given the strategies of all other players and play in future periods. In particular, if this strategy profile is an SPE, then there cannot be any profitable single period deviation from this equilibrium path for any player.

First we consider the contingency that leader α has either cheated in the current period or cheated in some previous period. Given this instance of cheating, leader β refuses all opportunity to cooperate in the future as long as leader α remains in power. Leader β does not hold a grudge against nation A per se, however, and will resume cooperation with α's successor. The leader specific component of LSGT offers the citizens in A an opportunity to restore cooperation if they remove α.

If the citizens of nation A replace their leader, then they must pay cost K_A to do so. In the PD between representative leaders, the only dimension on which the citizens evaluate their leader is the outcome of the PD game. If their leader's integrity is intact, meaning their leader has never cheated in the past, then future international cooperation occurs whether they replace their leader or not. Replacing their leader under this circumstance only imposes additional costs with no benefits.

Now consider the case where leader α has tarnished her integrity by cheating. If the citizens in nation A retain her, then leader β will refuse to cooperate in the next period, and the outcome of the PD game will be (D,D), giving the citizens a reward of P. If the citizens retained leader α indefinitely, then cooperation ceases indefinitely, the net present value of which is $\delta P + \delta^2 P + \ldots = \delta P/(1 - \delta)$. If instead the citizens replace their leader, at a cost of K_A, then in the next period leader β cooperates, and under LSGT the cooperation continues in every future period. The net present value of deposing leader α is $-K_A + \delta R + \delta^2 R + \ldots = K_A + \delta R/(1 - \delta)$. Provided that $K_A \leq (R - P) \delta/(1 - \delta)$, then the citizens in A prefer to depose α immediately. The leader specific component of β's punishment strategy means that on average the value of the challenger is $(R - P) \delta/(1 - \delta)$ greater than the value of retaining an incumbent who has cheated.

The condition $K_A \leq (R - P) \delta/(1 - \delta)$ was derived by considering whether the citizens remove α or retain her indefinitely. Technically, the proof that proposition 1.1 is an SPE requires consideration of single period defections from the equilibrium path. In this context a one

period deviation from the path means retaining leader α after she has deviated for one period before removing her. The payoff from doing so is $\delta P - \delta K_A + \delta^2 R/(1 - \delta)$. Comparing this with the payoff from removing α immediately, $-K_A + \delta R/(1 - \delta)$ also yields the same condition $K_A \leq (R - P) \, \delta/(1 - \delta)$. When the cost of removing leaders is low, then the citizens replace leaders who cheat. With the effects of cheating on domestic political survival established, how do leaders play PD?

Suppose neither leader has previously cheated. Under this contingency LSGT dictates that leader β plays C. If leader α plays C, then her payoff is $R + \Psi$ in the current period, plus $R + \Psi$ in every future period. The net present value of cooperation is thus $(R + \Psi)/(1 - \delta)$. However, leader α can improve her payoff in the immediate period by playing D. This yields the temptation payoff and officeholding rewards in the current period, but α is removed by the citizens and so does not receive the officeholding benefits in future periods. However, since α's replacement will restore cooperation, α receives the reward payoff R in all future periods. The net present value of cheating is therefore $T + \Psi + \delta R/(1 - \delta)$. Provided that $\delta \geq (T - R)/(T - R + \Psi)$, the value of cooperation exceeds that of cheating and so α cooperates in every period.[6]

When the cost of leader replacement is low ($K_A, K_B \leq (R - P) \, \delta/(1 + \delta)$), citizens replace leaders who cheat or who have cheated in the past. Given this replacement strategy, leaders who cheat lose office. We believe leaders are primarily driven by officeholding motives; that is, Ψ is large relative to $T - R$. Therefore, the condition $\delta \geq (T - R)/(T - R + \Psi)$ ensures that except under all but the smallest discount factors, full cooperation can be sustained.

Figure 1.3 graphs the minimum discount factor required to support full cooperation in the PD between representative leaders for high and low costs of leader replacement. When the cost of leader replacement is high, leaders who cheat retain office but forgo future cooperation. When the cost of leader replacement is low, leaders that cheat are removed from office. If, as we believe, leaders are primarily motivated by officeholding motives, then the punishment threatened for noncooperation in the latter case is much larger than in the former case. The greater the threatened punishment becomes, the easier it is for leaders to commit to cooperate.

Figure 1.3 is plotted assuming $T = 4$, $R = 3$, $P = 2$, and $S = 1$. It shows, once leaders value officeholding at least as much as the difference between the temptation and reward payoffs ($\Psi > T - R = 1$), then cooperation can

---

[6] We also need to check the optimality of LSGT under the remaining contingencies. If β has cheated in a previous period, then under LSGT, β will play D. Leader α's best response is also to play D. If leader a has previously cheated, then leader β will play D. Leader α's best response is to also play D.
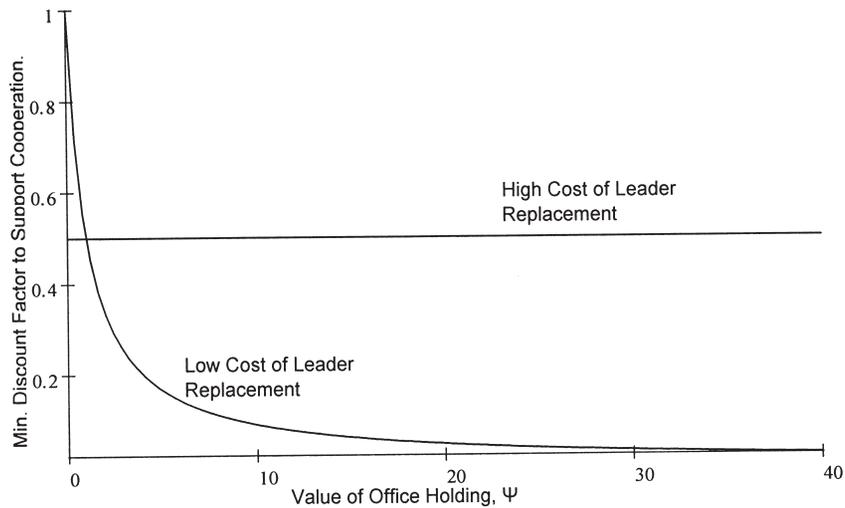
Figure 1.3. The minimum discount factor required to support full cooperation in the PD between representative leaders for high and low costs of leader replacement.

be maintained in the PD game under a wider range of conditions when leader removal is easier. If, for example, leaders care about office holding ten times more than the value of cooperation, then leaders require a discount factor of only $1/11$ to maintain cooperation. In contrast, when leader removal is hard, the maintenance of cooperation requires a discount factor greater than $\frac{1}{2}$. It is interesting to note that the leader, as agent of the citizens, can commit to cooperate under conditions that the principals themselves could not commit to cooperate under. If the citizens themselves choose the nation's foreign policy directly, then they would only cooperate if $\delta \geq 1/2$. Furthermore, when the cost of leader replacement is sufficiently low ($K_A < T - R$), the citizens want their leader to cheat. By doing so, the citizens gain the temptation payoff and can then replace their leader at cost K_A and so avoid the punishment phase. Of course, the leader does not want to cheat despite the public popularity of such a policy. Although by cheating the leader is carrying out the citizens' wishes, such a popular course of action will still lead to her dismissal.

By simultaneously examining interactions at the level of interstate relations, domestic political institutions, and individual leaders, leader specific punishments theory provides predictions about (1) how domestic political institutions shape the level of international cooperation, (2) how foreign policy outcomes affect the survival of leaders, and (3) the dynamics of interstate relations and how they depend upon domestic institutions. Since the theory combines different levels of analysis, it is appropriate to pause and examine the units of the international system.

PROPER NOUNS IN INTERNATIONAL RELATIONS

Who are the actors in international politics and what are the organizing principles in the study of international relations? The most common answer to these questions would be nations. Certainly the concept of nations provides a useful organizing principle. However, the extent to which nations are the "true" actors of international politics is questionable. While it is convenient to say the United States decided to invade Iraq on March 20, 2003, or to say France opposed the United States' actions in the United Nations, neither of these statements is strictly true. The United States never decided to invade Iraq. The decision was made by U.S. President George W. Bush's administration, having gained congressional approval for the use of force on October 11, 2002, with a 296–133 vote in the House of Representatives and a 77–23 vote in the Senate.[7] On March 10, 2003, Jacque Chirac, president of France, announced that he would direct the French ambassador to the United Nations, Jean-Marc de La Sabliere, to veto U.S. calls for the UN to authorize the use of force to remove and destroy Saddam Hussein's alleged weapons of mass destruction.[8]

One might argue that these distinctions are pedantic. After all, leaders enact policies that best fulfill national goals! Of course social choice theorists show us that there is no such thing as a "national will." National will is a construct of the institutional rules used to aggregate preferences. Had butterfly ballots in Palm Beach County Florida not confused so many voters, Al Gore, the Democratic candidate in the 2000 U.S. presidential election, would in all likelihood have been elected instead of Bush and perhaps the 2003 war with Iraq might never have taken place. Despite these social choice niceties, it is often useful to simplify a problem by classifying some outcomes as preferred to others by a massive majority in a nation. We are as guilty as anyone else of using this assumption. Throughout this book we assume that, all else equal, a nation unambiguously prefers to win a war rather than lose. However, whether the United States decides to launch a war against Iraq depends upon whose preferences get represented.

The 2003 U.S.-Iraq war is poorly conceived as a war between the United States and Iraq. At least according the Bush administration's foreign policy statements, the *causi belli* was Iraqi President Saddam Hussein. President Bush was explicit: "The U.S. is the friend of the Iraqi people." If Bush's foreign policy statements are to be believed, the Iraqi people were

---

[7] "Congress Backs Bush War Powers," BBC News, October 11, 2002. http://news.bbc.co.uk/1/hi/world/americas/2318785.stm.

[8] "Timeline: Steps to War,"BBC News, March 20, 2003, http://news.bbc.co.uk/1/hi/world/middle_east/2773213.stm.

not the target of the bombing. The target was Saddam Hussein and his re-gime. The targets of U.S. foreign policy are often not nations, but specific leaders, administrations, or political regimes in foreign nations.

In historical terms, treating nations as the actors of international rela-tions makes even less sense. When the English King John (1199–1216) rowed with the French, classifying the conflict as a war between England and France would badly misconstrue the dispute. The feudal system in place at that time was a series of hierarchical structures, in which people lower down the structure paid homage to those above. In England, King John ruled through the support of the barons, each of whom was obli-gated to do him homage and provide military resources at times of war. In return these barons held lands. As authors of histories are always quick to remind us, our modern concept of nations presumes distinctions that people of the medieval times would not recognize (Warren 1997). The politics of the time were between kings, barons, knights, and ecclesiasti-cal actors and not between states. For example, the foundations for what we regard as wars between England and France were over feudal rights. Although John was king of England, he was also duke of Normandy and via his mother, Eleanor of Acquintane, he also held great swaths of land in southwestern France.

The feudal system was arcane. In one regard John as king of England was an equal of Philip Augustus, king of France. Yet on another level he was subservient to the French king, as he held his French lands as a vassal of the French king. To provide further complications, the pope would like to have claimed supremacy of over all secular authorities. Rather than war between England and France, John and Phillip fought for control of lands that are in modern-day France, such as Normandy, which John held as part of his feudal rights that were independent of his role as king of England.

At the same time, of course, John faced rebellion in England from his barons. These barons had no interest in John's confrontations with the French king over rights to lands in France in which they had no stake. Their feudal obligations made them duty bound to provide the king re-sources for his wars in France, but it was his fight, not theirs. It is perhaps small wonder then that the barons rebelled, forcing John to make conces-sions in his famed signing of the Magna Carta in 1215. While the Magna Carta is often thought of as having enormous historical significance in establishing the rights of commoners and nobility vis-à-vis the king, as a contemporary document its importance was minimal. Pope Innocent III quickly annulled the document; he liked John's policies of Crusading and war with France.

While from a twenty-first-century perspective we refer to the wars of King John's reign as Anglo-French wars, few if any people at the time would have identified themselves as English or French. King John himself

did not even speak English; Edward I (1272–1307) was the first (post-1066) English king to speak English (Gunaratne 2003). During the medieval period international politics rarely had much to do with nations.

According to many international relations textbooks (for example, Russett, Starr, and Kinsella 2005) the nation-state came into existence in the Treaty of Westphalia of 1648, a series of treaties that concluded the Thirty Years War. The Treaty of Westphalia contains provisions that allowed princes to freely choose the religion of lands they controlled (although the people themselves still had little choice in their religion). The Thirty Years War is often seen as a battle over religious rights between Catholic and Protestant rulers. While contemporary international relations scholars focus predominantly on these aspects of the Treaty of Westphalia, a greater percentage of the treaty dealt with which individuals get what in terms of lands, rents, and other booty.[9]

In the post-Westphalian era many wars can indeed be portrayed as conflicts fought between nations for national interests. One does not have to dig deeply to see that many of these wars are not based around national competition, however, but rather are driven by domestic political interests. For instance, the Prussians started the Wars of the French Revolution by invading France to restore the monarchy of Louis XVI.

Although our modern conception of the international system is organized around nations, the policies nations pursue depend upon who is national leader and which domestic interests this leader represents. Who gets to be leader of a nation and which interests the leader represents depend upon the nation's domestic political institutions. Foreign policies are drawn up with goals and targets. While in some case these targets might be a nation (the national unit as a whole), in other cases foreign policies are either implicitly or explicitly targeted against a specific feature of a nation's polity, such as the leader. This book treats international politics not as competition between amorphous national groups, but as foreign policies composed by political elites in one nation with regard to elites in another nation. Domestic political institutions and the preferences of leaders interact to shape the types of policies that elites choose. Domestic institutions in a foreign nation determine how its political elites respond to these foreign policies.

This book builds a theory of international politics based on the actions of individual leaders constrained by domestic political institutions. The

---

[9] Bueno de Mesquita et al. (2003, pp. 432–34) categorized the clauses of a number of treaties according to whether they concerned public policy, private benefits (such as the allocation of lands and rents to individuals) or implementation and procedural issues. Of the treaties 128 clauses, they code only 36 (28 percent) as being involved with issue of public concern. In contrast, 55 (43 percent) clauses concern the allocation of private benefits. The remaining 37 clauses mainly concern implementation and procedural issues.

interactions between nations depend upon the goals of these individual leaders and political institutions. The theoretical basis of our arguments is individual choice and the aggregation of preferences. However, we do not generate our results by considering wide varieties of preferences and arguing in favor of different sets of preferences to explain different events. Instead we consider a simple set of goals for each actor. For instance, we assume political leaders primarily want to retain their jobs and that all the citizens of a nation have a common objective function with respect to international outcomes.[10] From this sparse framework we examine how the strategies of leaders interact to produce international outcomes and how domestic institutions modify these interactions.

INTERNATIONAL COOPERATION

Cooperation and coordination represent fundamental problems in international relations. Although later, in chapter 7, we consider to more conflictual relationships, the bulk of our study investigates cooperation. In a domestic setting if two groups wish to work together for some common goal, they can sign a contract. If one party subsequently shirks its obligations, the other group can sue it in a court of law and receive compensation. The threat of being sued is sufficient to ensure that both parties contribute to the joint goal in accordance with their agreement. Unfortunately, the anarchy of international relations makes international cooperation much more difficult. Without courts and police to enforce contracts, nations have little incentive to honor their obligations.

The problems involved with international cooperation are well known to political scientists. Following Keohane (1984, p. 12), we take international cooperation to mean "mutual adjustment." Keohane (1984, chap. 4) is careful to distinguish between harmony—a situation where nations' interests are already sufficiently aligned that by default they want to take actions that are mutually beneficial—and cooperation, where there is some discord between the objectives of each nation such that adjusting their policies for mutual gains requires changing policies from those the nations would adopt absent attempts to mutually improve welfare.

Whether the topic under consideration is trade and tariff arrangements, coordination of monetary policy, arms control agreements, sharing of

---

[10] There is a substantial literature that argues individual characteristics of leaders, such as their gender, age, marital status, and even birth order, affect national policy. Others argue a leader's psychological and behavioral makeup plays an important part (Goldstein 2001; Horowitz, McDermott, and Stam 2005; Hermann, Tetlock, and Diascro 2001; Post 2003; Rosen 2005). Although we do not want to dismiss these factors, here we show that the effects of leader turnover can be explained even when all leaders are assumed to be homogenous.

common pool resources, environmental agreements (or other externality problems), contributions to a common defense commitment (Olson and Zeckhauser 1966), or shared technology and research, the difficulty of cooperation can be explained in terms of a collective action problem. Of course not all the interactions between states can be thought of as opportunities to cooperate: however, such cases present a wide range of interesting problems from which to start (Simmons and Martin 1998).[11] Later we discuss more conflictual and less cooperative interactions.

Although they can be divided into a number of different categories, such as the provision of public goods, externalities, or common pool resources problems (Olson 1965), collective action problems share some basic features. If all parties agree to undertake some socially preferred set of policies, then each party is better off than if all parties acted myopically. Depending upon the topic under consideration, there are numerous examples. For instance, in the context of international trade between two nations, free trade is mutually beneficial relative to each nation being autarkic. In another example, international liquidity aids international trade, investment, and capitalism. If all nations contribute to a fund to ensure liquidity for distressed banks and other financial institutions overseas, all nations benefit from avoiding international financial shocks. International liquidity is a public good that benefits all members of the international community. No member of the international community can be excluded from the benefits of a robust international economy; neither does any nation's enjoyment of a robust international economy diminish another state's enjoyment. Once the public good of international liquidity is provided, all parties benefit, whether or not they provided any of the resources required to produce the good. Given they benefit whether or not they contributed to the public good, each party wants to minimize its contribution. The net result is an underprovision of the public good. Of course in the domestic setting, legal and contractual arrangements can be used to overcome the collective action problem. For instance, governments collect compulsory taxes from their citizens to provide for such public goods as national defense and public health. Theories of public goods provision have been well developed, and we need not go over the details here (Olson 1965). As with other collective action problems, the key features are that even though all parties could make themselves better off by coordinating their actions, each party could make itself even better off by free riding on the efforts of others.

As we saw earlier, the prisoners' dilemma is a convenient model that encapsulates the inherent problems of international collective action (Axelrod 1984; Axelrod and Keohane 1986; Bendor 1987; Downs and

---

[11] Morrow (1994) provides a useful framework from which to consider coordination problems.

Roche 1990; Gourevitch 1996; Milner 1992; and Pahre 1994). In the classic explication of the game, two criminals have been caught. The district attorney separately offers each a deal if he agrees to testify against his partner in crime. The payoffs T, R, P, and S reflect the length of sentences the criminals can expect to receive depending upon who rats on whom or who remains silent.

To see why the prisoners' dilemma serves so well as a model for international cooperation, we need only change the actions of each party. Returning to the example of international financial liquidity for instance, cooperation (C) means contributing resources to the provision of the public good, while defecting (D) means shirking. Both parties are better off if each contributes its share ($R > P$), but each party is even better off if it keeps its resources while benefiting from the partial provision of the public good by the other nation ($T > R > P$).

Whether we think of the prisoners' dilemma as a game between prisoners negotiating with the DA, nations contributing to a public good, or any other collective action problem, each party wants to defect. Cooperation is difficult in PD because whatever the strategy of the other side, each party has a dominant strategy to defect. That is to say, if player A chooses C, then player B obtains the maximal payoff T by defecting. If player A defects, then player B chooses between cooperating, which produces S, the worst payoff, or defecting and obtaining the payoff P. In either case, player B is better off playing D to C. In the context of PD, the prospects for international cooperation appear bleak. Yet, international cooperation frequently occurs: therefore, at least one of the assumptions in the prisoners' dilemma model of international cooperation must be wrong.

Hegemonic Theory

Hegemonic stability theorists argue that hegemons—that is, nations that predominate over all other states—promote international cooperation (De Cecco 1975; Feis 1930; Ford 1962; Kindleberger 1981; Lindert 1969; Wallerstein 1980). In the context of economic issues, such as trade and financial liquidity, they argue that a hegemon controls such a large proportion of the world economy, and its share of the gains from the provision of public goods is so large that this outweighs the cost of providing the public good. In the context of the prisoners' dilemma, this would be to say that the hegemon's preferences are $T > R > S > P$. Given that the hegemon would prefer to unilaterally provide the public good rather than see it not provided, the hegemon's optimal strategy is to unilaterally provide the public good by playing C. The smaller state has no incentive to provide the public good, since the hegemon is already doing all the work.

One convenient way to think of the hegemonic argument is a division of a pie. By providing a public good, nations can increase the size of the overall pie. However, when a large number of nations each has a small share of the pie, no nation individually gains from investing in the public good, since it receives only a small share of the increased size of the pie. In contrast, a hegemon receives such a large share of the pie that it pays for the hegemon to increase the size of the pie since such a large proportion of the increased pie goes to the hegemon.

Hegemonic arguments have been used to explain the expansion of trading and international banking under the hegemony of the British prior to 1914, Pax Britannica. Despite explaining cooperation as the presence of a hegemon in these cases, Carr (1962) argues that the collapse of the world economy during the 1930s was largely due to the United States' unwillingness to provide public goods despite its hegemonic position. Although hegemonic arguments explain cooperation in the presence of a hegemon, they fail to explain cooperation in the absence of a hegemon. Indeed, one might argue that increasingly high levels of cooperation achieved through such organizations as the World Trade Organization in the latter part of the twentieth century, coupled with declining U.S. hegemony, falsifies hegemonic arguments. Liberal theory offers an explanation for cooperation in the absence of a hegemon.

Liberal Theory

In his classic book, *After Hegemony*, Robert Keohane (1984) explains that even in the absence of a hegemon, international cooperation can arise if nations use reciprocal strategies. That is to say, nations condition their current and future play on the outcome of past interactions. As Liberal theorists argue, provided that nations are sufficiently patient, such strategies make cooperation possible, and collective action problems can be solved (Axelrod and Keohane 1986; Axelrod 1984, 1986; Baldwin 1993; Busch and Reinhardt 1993; Goldstein 1991; Gowa 1986; Keohane and Nye1977; Krasner 1983; Milgrom, North, and Weingast 1990; Milner 1992; Oye 1986; Ruggie 1993).

Earlier we formalized these arguments within the context of the prisoners' dilemma using the Grim Trigger. Of course, GT is just one strategy through which cooperation can be obtained. When nations are sufficiently patient, there are infinitely many SPEs that exhibit many patterns of behavior—a result known as the Folk theorem (Fudenberg and Maskin 1986). To prove the existence of these other patterns of equilibrium behavior, Folk theorem type results find a punishment schedule that nations want to implement if a nation deviates from a prescribed pattern of

play. The GT strategy is a limiting case because it utilizes the harshest punishment schedule—the permanent removal of cooperation. In short, if GT does not work, there is no other way of securing cooperation in equilibrium.

Although GT defines the theoretical limits of cooperation in infinitely repeated PD, from a practical perspective, cooperation is much harder to achieve. In reality, once cooperation fails, nations want to try to restore it. Unfortunately, the desire to try to renegotiate cooperation once the punishment phase starts undermines the threat of the punishment in the first place. Formally, equilibria that are immune from attempts to renegotiate an end to the punishment phase are referred to as *renegotiation proof* (Farrell and Maskin 1989). GT is not renegotiation proof. Given the prospects of indefinite punishment, both parties prefer to negotiate a return to cooperation. Unfortunately, this undermines the strength of the punishment threat in the first place.

The GT strategy assumes a noiseless world of perfect information and no errors. Unfortunately, the real world is a noisy place where nations are liable to misinterpret each other's actions. Further, in practical terms nations do not choose between C or D. In reality, nations' choices are much more complex. For example, nations might choose how high to set tariffs or whether to impose non-tariff barriers. These actions can be thought of as shades of gray rather than the black-and-white choices of C and D. In our formal analysis of international cooperation in chapter 2, we allow for the integration of both continuous choice action spaces and noise.

Despite these limitations, the Liberal interpretation of international cooperation as an infinitely repeated PD with cooperation made possible via reciprocal strategies has been a powerful idea in political science. The importance given to this result is well founded. These ideas explain how cooperation is possible when collective action problems suggest the prospects for cooperation are poor. Unfortunately, whatever the value of the infinitely repeated PD analysis, it inherently remains a possibility result. Above we showed that GT is an SPE. However, both nations always playing D in every period is also an SPE. The analysis tells us that the former cooperative equilibrium is possible when the temptations to cheat are not too great, when the rewards from cooperation are large, and when nations are patient: specifically, $\delta \le (T-R)/(T-P)$. Beyond these limits, however, the result tells us nothing about when cooperation is most likely. The SPE analysis provides us no comparative static results with regard as to whether cooperation is more likely between nations A and B or between nations X and Υ.

Empirical analyses suggest some pairs of states are more likely to cooperate with each other than are other pairings. Scholars such as Russett and Oneal (2001), Leeds (1999), and many others show that democratic

dyads cooperate at far higher levels than do other dyadic pairings of states.[12] The basic Liberal analysis cannot explain these differences without resorting to arguments that PD payoffs for democratic states differ from the PD payoffs for other states, or that democracies are inherently more patient.

We believe the failure of Liberal theory to predict which nations are most likely to cooperate is a major limitation of the approach. Liberal scholars have also done little to show the dynamics of reciprocal strategies in the pattern of cooperation (Goldstein 1991). The GT strategy predicts that once defection occurs, cooperation permanently ends. As the brief anecdotes at the start of this chapter indicated, the end of cooperation is rarely permanent. After several years of harsh economic sanctions and international isolation, Yugoslavia (Serbia) is again an active member of the international community that receives economic assistance and investment from Western states. Obviously, GT cannot account for the restoration of interstate relations. While other strategies, such as Tit-for-Tat (Axelrod 1984), allow for the restoration of cooperation on the equilibrium path, Liberal theorists have done little to provide empirical evidence that reciprocal strategies explain the restoration and termination of international cooperation.

Liberal theory treats nations as unitary actor states. As we argued above, nations are not the only proper nouns of international relations. Although Liberal theory has made an appropriately huge impact on the study of international relations, it cannot explain how institutional differences between states shape the level of cooperation. Neither can it explain the dynamics of cooperation. We believe a theory of leader specific punishments addresses these deficiencies.

Leader Specific Punishments and International Cooperation

Nations are not unitary actors interested in maximizing social welfare. Leader specific punishment (LSP) theory dispenses with the unitary actor assumption and replaces it with a principal-agent framework in which representative leaders are the agents and the citizens are the principals. Although leaders are assumed to care somewhat about international outcomes, they

---

[12] The literature emphasizing the ability of democracies to cooperate more than autocracies is wide and varied: Bliss and Russett 1998; Busch and Reinhardt 1993; Gaubatz 1996; Gowa 1994; Leeds 1999; Mansfield, Milner, and Rosendorff 2000; Mansfield and Pevehouse 2000; Mansfield and Pollins 2001; Martin 1993; McGillivray 1997, 1998; Milner 1997; Milner and Rosendorff 1997; Morrow, Siverson, and Taberes 1998; Oneal and Russett 1997, 1999a, b, 2000, 2001; Oneal, Russett, and Berbaum 2003; Polachek 1997; Pollins 1989; Remmer 1998; Reuveny and Kang 1996, 1998; Reuveny 2000, 2001; Russett and Oneal 1999, 2001; Verdier 1998.

primarily want to keep their jobs. Domestic political institutions affect the ease with which citizens can replace their leader. When the cost of leader replacement is high, a leader's political survival is relatively detached from her ability to produce successful foreign policy outcomes. In contrast, when a leader is easily replaced, her political fate depends upon being able to deliver good international outcomes.

LSP examines how targeting punishments against individual leaders, rather than the nation they represent, affects the level of cooperation between states, the survival of political leaders, and the dynamics of interstate cooperation.

LEVEL OF COOPERATION

Democratic dyads—that is, pairs of democratic states—cooperate at higher levels than do other dyadic pairings of states. This empirical result has been established in numerous settings. We dwell on these extant results for a moment. Although leader specific punishment theory predicts these results, our empirical tests do not focus on the level of cooperation between states. As we are about to summarize, the impact of domestic political institutions on the level of cooperation has been well established in the empirical literature. To repeat similar tests would be largely redundant, as it would provide little new information. Instead, our empirical tests will focus on the novel and relatively underinvestigated results regarding the dynamics of cooperation and leader change.

That pairs of democracies behave differently from other dyadic pairings of nations has been a common theme of the international relations literature over the past decade. The impetus for this research stems from the democratic peace, an observation that democracies do not fight each other (Maoz and Abdolali 1989; Ray 1995; Bremmer 1992). Although numerous cases, from ancient Greece to the United States imperialist wars against Native American tribes, have been proposed as potential examples and counterexamples (see Russett 1993 and Weart 1998 for discussion of many of these cases), the result appears to have been generally accepted by much of the discipline. Indeed, Jack Levy (1988) has gone so far as to call it a *law*. The principal democratic peace result is that democracies do not go to war with each other. However, democracies do become involved in wars with nondemocratic states (Maoz and Abdolali 1989). Democracies also become involved in violent conflict with other democracies: it is just that these conflicts do not escalate to war (Oneal and Russett 1997; Senese 1997). The literature appears conflicted as to whether democracies are more or less aggressive, in terms of overall war participation, than other regime types (Benoit 1996; Ray 1995). The conflict behavior of democratic states has also been shown to differ greatly from that of other

states in a variety of ways. Democracies generally win the wars they fight
(Lake 1992; Reiter and Stam 1998a, 2002). Further, they typically win
quickly and with relatively few casualties (Reiter and Stam 1998a, b; Si-
verson 1995; Bueno de Mesquita et al. 2004). Democracies also tend to
fight for policy change or regime change, while nondemocratic states are
more likely to fight for land (Bueno de Mesquita et al. 2003). Democra-
cies are also more likely to use conflict management techniques (Brecher
and Wilkenfeld 1997; Dixon 1994; Mousseau 1998; Raymond 1994).
Democracies often initiate conflict against nondemocracies (Reiter and
Stam 1998b). Transitional status and size also appear to affect the con-
flict involvement of democracies (Mansfield and Snyder 1995; Ward and
Gleditsch 1998; Morgan and Campbell 1991).

Many of the theoretical efforts to explain these regularities have fo-
cused on either normative arguments or institutional constraints (Maoz
and Russett 1993). Unfortunately, few of these theoretical arguments
have satisfied the criteria of explaining all the known empirical regularities
and predicting novel hypotheses (Rosato 2003); although we believe that
Bueno de Mesquita and his colleagures' selectorate politics explanation
of the democratic peace has made substantial progress in this direction
(Bueno de Mesquita et al. 1999, 2004).

Motivated in part by a desire to explain the democratic peace result,
scholars have sought to find other regularities associated with democracy
outside of conflict behavior. For example, numerous studies have found
that regime type influences trade (Bliss and Russett 1998; Gowa 1994;
Mansfield and Pevehouse 2000; Mansfield and Pollins 2001; Milner and
Rosendorff 1997; Morrow, Siverson, and Taberes 1998; Oneal 2003;
Oneal and Russett 1997, 1999a, 1999b, 2000, and 2001; Polachek 1997;
Pollins 1989; Reuveny 2000 and 2001; Reuveny and Kang 1996 and
1998; Verdier 1998). Even controlling for their typically large economies
and regional concentration, democratic dyads appear to trade with each
other to a greater extent than do other pairs of nations. Scholars such as
Russett and Oneal (2001) argue this affinity between democratic states
extends beyond simple trade and affects their propensity to invest in each
other, join international organizations together, and generally cooperate
at a high level. (For evidence on the greater propensity of democratic
states to join international organizations see Jacobson, Reisinger, and
Mathers 1986; Shanks, Jacobson, and Kaplan 1996; Russett and Oneal
2001; Russett, Oneal, and Davis 1998; Mansfield, Milner, and Rosend-
orff 2002; Mansfield and Pevehouse 2006).

Leeds (1999) finds further evidence of greater cooperation between
pairs of democratic states using COPDAB data (Azar 1982). These data
are compiled through the reporting of news events. She finds that pairs of

democratic states have systematically more cooperative relations than do other pairs of states.

The empirical evidence portrays a clear picture. Relations between states with democratically accountable leaders are more cooperative than relations between other pairs of states. Although we shall show further evidence of this in our subsequent analyses, we do not focus on this result. It is already well established, and repeated analysis does not help us understand why. Many of the works cited above propose theoretical explanation for this result. For example, Leeds (1999) argues democracies cooperate because their leaders face audience costs from breaking their commitments. Russett and Oneal (2001) compare a wide range of structural and normative approaches to explain democratic behavior. Unfortunately, based only on evidence relating to the level of cooperation, it is impossible to distinguish between rival theoretical explanations, all of which predict elevated levels of cooperation between democracies. Only by extending the analyses to consider dimensions on which the theories have differing predictions can we separate them. In this book most of our empirical tests have this goal in mind. We are less interested in describing behavior that is broadly predicted by many theories than we are in testing the hypotheses generated by LSP theory that distinguish it from other approaches.

IMPACT OF FOREIGN POLICY ON LEADER SURVIVAL

Democratic dyads cooperate more than other dyads. According to the theory of leader specific punishments, democratic leaders cooperate because a failure to do so costs them their jobs. The theory predicts a relationship between policy choice and leader removal. Democratic leaders who cheat on their agreements or otherwise violate norms of international behavior and so incur the ire of other states are removed from power. In contrast, the high cost of leader removal in authoritarian states means autocrats can incur the wrath of the international community and trading partners with impunity, at least with respect to domestic political removal. Unfortunately, directly testing this hypothesis is extremely difficult as leaders do not make policy choices that jeopardize their own political survival. There is a selection effect. If cheating on an international agreement would cost a leader her tenure in office, she does not cheat. Therefore, instances where we observe a leader being removed for cheating are extremely rare.

The term *audience costs* is commonly used to describe any costs leaders face as a result of their foreign policy decisions (Fearon 1994). Fearon argued that leaders involved in crises face domestic political repercussions from escalating crises and then subsequently backing down. He argued that

democratic leaders, being more accountable, face higher audience costs than autocratic leaders. These higher costs enable democrats to more effectively commit themselves and help them prevail in crises and maintain cooperative agreements (Bueno de Mesquita and Lalman 1992; Eyerman and Hart 1996; Guisinger and Smith 2002; Leeds 1999; Mansfield, Milner, and Rosendorff 2002; Martin 1993; Partell and Palmer 1999; Schultz 1998, 1999, 2001, 2002; Smith 1998).

The basis articulation of audience cost theories simply asserts the existence of audience costs without deriving their origin within the political system. Unfortunately, this creates something of a time inconsistency problem in the credibility of the audience costs. Audience costs allow a leader to tie her own hands, thus enabling her to commit to a course of action that she would not otherwise take (Fearon 1997). Audience costs turn bluffs into credible commitments. However, should the commitment fail to get the opposing leader to concede, the citizens do not want the leader to carry out her stated policies. Yet it is the threat that the citizens will punish their leader that causes her to stay the course and enact the policies that she and the citizens do not want. For the citizens, enforcing audience costs is against their interests once a leader's bluff has failed. Leader specific punishment theory resolves this inconsistency because it simultaneously derives the origins of audience costs and their effect on bargaining and other relations between states.

Audience costs affect interstate relations by making it possible for leaders to commit to carry out threats or commit to cooperate (depending upon the context). Unfortunately, it is difficult to directly test the audience cost mechanism. The basic argument is that leaders face costs for taking particular actions. Unfortunately, we cannot effectively measure whether leaders are indeed punished for these actions. If audience cost theories are correct, democratic leaders who escalate crises and back down or who break agreements are likely to be punished domestically. However, the larger the audience cost is likely to be, the smaller the chance becomes that we actually observe the audience cost being imposed. This creates sample selection problems in that we can only assess audience costs when they are modest. Schultz (2001) demonstrates why this makes the direct observation of audience costs impossible (Gelpi and Grieco 2000) and why it creates biases in many other empirical tests.

One immediate criticism of leader specific punishment is the lack of direct evidence for it. Leaders who are easily deposed and who violate international agreements should be removed. Analyses of public opinion, such as Hermann and colleagues (2001), suggest it is indeed costly for leaders to violate agreements. Unfortunately, direct evidence of democratic leaders being removed for cheating should be (and is) rare. Such leaders are unlikely to cheat if it costs them their jobs. Throughout this book we

offer examples of nondemocratic leaders cheating and the restorative effects of their subsequent replacement. We can offer far fewer examples of democrats cheating. This is precisely what the theory predicts. We cannot assess the impact of cheating on the domestic political tenure of democrats because democrats typically don't cheat.

DYNAMICS OF COOPERATION

The most novel and interesting hypotheses derived from leader specific theory concern the dynamics of leader change and interstate cooperation. To our knowledge, outside of our work, these dynamics have not been systematically explored before. Leader specific punishments endogenously provide opportunities to restore cooperative relations. It is individual leaders, rather than the nations they represent, who choose to cheat. It is, therefore, perhaps natural to expect that punishments are targeted against leaders.

If the leader of nation A adopts leader specific punishments against nation B and the leader of nation B cheats, then nation A withdraws cooperation, or otherwise imposes sanctions, until the leader in nation B changes. Leadership change refreshes sour relations as the following example illustrates. During the 1991 Gulf War, Jordan's King Hussein sided with Iraq. Although Jordan did not become involved militarily, it kept its border with Iraq open, making the enforcement of multilateral sanctions much more difficult. Jordan is relatively devoid of natural resources and has traditionally received substantial financial support from other, wealthier, Arab states. Most Arab states joined the U.S.-led coalition to remove the Iraqi forces that had occupied Kuwait. In retaliation for Jordan's support of Iraq, most Arab states cut off their traditional economic support for Jordan. The February 7, 1999, death of King Hussein provided the impetus to renew relations. Once Hussein's son ascended the throne, Arab states renewed their economic assistance despite few signals of policy change (*New York Times*, February 19, 1999. p. A3).

Leadership change brings about shifts in policy and reshapes external relations. These dynamics are not constant across all political systems, however. Domestic political institutions play an important role in determining the extent to which the citizens hold political leaders accountable for international outcomes. As we already argued, the greater ease of leader replacement in democratic nations encourages the citizens of these nations to replace their leader if she is caught cheating. The desire to avoid such a removal from office enables democratic leaders to commit to not cheat. This allows for greater levels of cooperation between democratic states than is possible between other pairs of nations. It also means that instances of sour relations between democratic nations are unlikely.

Further, leadership turnover in democratic states has little impact on the
restoration of sour relations because it is unlikely that the relations were
sour as a result of the democrat's actions in the first place.

DOMESTIC POLITICAL INSTITUTIONS

Thus far we have derived the effects of leader specific punishments on
interstate relations in terms of the ease of domestic leader replacement.
For convenience of language, we have substituted the term *democracy* for
systems with *low cost of leader removal*. However, equating these terms is
not strictly accurate. To our knowledge, political institutions are never
classified as democracies on the basis of the ease of leader removal. To
operationalize leader specific punishment theory we require a metric for
the cost of leader removal. Bueno de Mesquita and his colleagues' (2003;
hereafter BdM2S2) theory of selectorate politics, classifies institutions ac-
cording to the size of the winning coalition (W)—the number of loyal
supporters whom the leader needs to retain power—and the selectorate
(S), the size of the group from which these supporters are drawn.

BdM2S2 argue that small winning coalitions systems, especially in the
presence of a large selectorate, induce a strong loyalty norm toward the
incumbent, which makes it relatively easier for such leaders to survive
relative to leaders in large coalition systems. A leader's policies provide
rewards for individual supporters (private goods) as well as public goods
that benefit all members of society. The number of supporter whose loy-
alty a leader must maintain to survive in office shapes the balance of her
policies between private rewards for her supporters and the provision of
public goods. When a leader requires the support of only a small coalition
to survive in office, she can effectively enrich this small group by provid-
ing them with private benefits and the particularistic policies they desire.
However, as coalition size increases it becomes increasingly expensive for
leaders to buy support with private goods, and leaders must rely increas-
ingly on public goods to reward supports.

In large coalition systems most of a leader's resources and energy goes
toward the provision of public goods. Although large coalition leaders
provide their supporters with some of their particularistic wants, the focus
of government policy is on good public policy. In a large coalition system
members of an incumbent's coalition jeopardize relatively little if they
defect to a political challenger. Although once in power the challenger is
likely to reorganize his coalition of supporters and potentially replace the
defector, the supporter has little to fear from being excluded from the
coalition. In a large coalition system most of the rewards are provided
in the form of public goods. All members of society benefit from these
goods whether they are coalition members or not. Indeed, since private

goods make up only a small proportion of the rewards in a large coalition system, those outside the coalition are only slightly worse off than those inside the coalition. Potential defectors have relatively little to fear from being excluded from the coalition and thus have little loyalty to the incumbent. If the challenger can better provide public goods (such as international cooperation), coalition members readily defect.

Small coalition systems engender a strong norm of loyalty toward the incumbent. The leaders of such systems predominantly rely upon private goods to reward their supporters. The welfare difference between those inside and those outside the winning coalition is therefore large. Supporters of the incumbent are reluctant to risk losing the highly valuable private goods the incumbent supplies. Although the challenger might offer a supporter huge rewards to defect, supporters are aware that once ensconced in power, the new leader is liable to reorganize a coalition. When the coalition size is small, so the leader needs only a limited number of supporters, and the selectorate is large, so the leader can choose the supporters from a large pool, then supporters recognize that there is a substantial chance of their being excluded from the coalition (and therefore from access to the valuable private goods) if they defect. When coalition size is small, such that access to private goods is very valuable and the risk of exclusion from the challenger's coalition is high, the incumbent's supporters are extremely loyal.

Selectorate politics, in the process of deriving the types of policies pursued under different political institutions, generates a metric for the ease of leader replacement. The selectorate model also suggests that leadership change in small coalitions produces much large variability in policies than occurs when leaders change in large coalitions. In large coalition systems leader survival is predicated on the effective provision of public goods. Leader change does not change this policy goal. The incoming leader, like his predecessor, enacts those policies that best further the interests of the nation at large. In contrast, small coalition leaders survive by pandering to interests of their small number of supporters. Providing rich rewards for these supporters is more important to a leader's survival than effective governance. However, since leader change often leads to a change in coalition membership, wild shifts in policy can occur as the incoming leader drops the particularistic interests of his predecessor's coalition and panders instead to the wants of his supporters.

The essential public goods focus of large coalition systems remains unchanged by leader change. Leader change therefore has relatively little impact on the relations between states. However, leader change in small coalitions creates great policy variability. This volatility potentially disrupts relations between states. For instance, trade can be severely disrupted as the government switches from favoring one sector of the economy to another.

The selectorate model of politics provides a metric for the cost of leader replacement that is an essential component of LSP theory. In chapter 2 we develop our leader specific punishment theory through a careful examination of the underlying assumptions of the theory and two formal models. These models relax the simplistic assumptions of the prisoners' dilemma game between representative leaders developed in this chapter. As noted above, the prisoners' dilemma makes the unrealistic assumptions of perfect observation of a binary action choice. The first model introduces randomness into the payoffs of the prisoners' dilemma and assumes that leaders face mortality risks. The second model is a continuous choice prisoners' dilemma model with noise. This is to say, a leader chooses an action on a continuum and the other leader cannot perfectly observe the leader's action. Instead each leader receives a noisy signal from the other leader's actions. The models formally derive the properties and dynamics of leader specific punishments to which we have informally alluded in this chapter. These models are developed in terms of the cost of leader replacement.

In chapter 3 we examine the selectorate model of politics and use it to link the cost of leader replacement, a vital component of LSP theory, with political institutions. The selectorate theory also generates an additional series of hypotheses concerning how political institutions affect the variability of policy change associated with leader change and its implications for international cooperation.

Broadly the theory predicts the following relationships between political institutions, leadership change, and interstate relations:

1. Nations with large winning coalitions maintain higher levels of cooperation than other nations.
2. Nations with small winning coalitions experience greater volatility in their external relations with other states than do nations with large coalition governments.
3. Leader change in a small coalition nation is more likely to alter interstate relations than is leader change in a large coalition system.
4. Leaders from small coalition systems are more likely to cheat on international agreements or otherwise incur the ire of the international community than are leaders from large coalition systems.
5. Leader turnover helps reinvigorate tarnished relations between states, although due to the selection effect that large coalition leaders are unlikely to take actions that lead to the breakdown of cooperation, this effect is only generally observed for small coalition leaders.

Chapters 4, 5, and 6 turn to testing the implication of the theory. Chapter 4 describes a series of human subject experiments designed to

closely mirror the prisoners' dilemma game between representative leaders discussed in this chapter. The evidence from the experiments supports the intuitive plausibility of leader specific punishments. In particular we show that a change in leader reduces the dependence between the past history of play and a leader's choice, which aids in the restoration of cooperation. At least in the experimental setting, it appears that leader turnover helps rejuvenate previously sour relations.

Chapter 5 examines how institutions and leader change affect trade relations between nations. Our analyses of dyadic trade flows support theoretical predictions. Pairs of nations with large coalition systems experience greater trade than do other pairs of nations. The effects of leader change on trade flows depend strongly on political institutions. In large coalition systems, leader change has no appreciable affect on trade flows. However, as implied by the high policy volatility in small coalition systems, leader change in such systems substantially reduces trade. We code instances of sour relations between states by identifying collapses in the value of trade between nations. These collapses occur more often between trading partners that include small coalition systems rather than large coalition systems. Consistent with LSP, we find that during periods of sour relations, leadership turnover in small coalition states provides a major boost to trade.

Chapter 6 examines sovereign debt. The terms of sovereign debt depend upon the lenders' beliefs that they will be repaid. We examine the impact of leader specific punishment in the context of sovereign debt borrowing by developing a simple formal model of borrowing and repayment. Consistent with the themes developed throughout this book, LSP allows large coalition leaders to credibly commit to repay loans. We test the dynamic predictions of the theory using sovereign debt bond indices. These indices reflect changes in the willingness of investors to hold regularly traded U.S. dollar denominated sovereign debt bonds. In particular, we examine how institutions moderate how these indices respond to leader change. In large coalition systems leader change has no appreciable effect on the value of sovereign debt bonds. In contrast, in small coalition systems leadership change generally lowers the price of the index, reflecting a decreased willingness of investors to hold these bonds. After examining instances of sovereign default, we examine the effect of leadership change on the bond price. Consistent with LSP ideas, leadership change following default helps increase the value of sovereign debt bonds.

In chapter 7, we examine leader specific punishments in conflictual situations of crisis bargaining and economic sanctions. Through a discussion of a number of historical events we examine the impact of LSP within crises. We use McGillivray and Stam's (2004) analysis of leadership change and the termination of economic sanctions to motivate a discussion of LSP in economic warfare. Consistent with the anecdotal evidence offered

in this chapter, leader turnover has a major impact in the termination of sanctions.

In chapter 8 we conclude by considering the broader policy implications of LSP theory. In particular, we address how leader specific punishments could improve the efficacy of a nation's foreign policy in crisis bargaining and lead to a deepening of cooperation and compliance within international agreements.