

Beyond Anarchy

David A. Lake

The Importance of Security Institutions

Security institutions are central to patterns of conflict and cooperation within the international system.¹ Although bipolarity would undoubtedly have brought the United States and the Soviet Union into competition anyway, the depth and length of the Cold War were exacerbated by the security institutions created after 1945. The Soviet Union's informal empire in Eastern Europe threatened the West because it secured Moscow's control over the region; a looser alliance in which each state could choose its own foreign policy and form of internal rule would have been far more fractious and diffuse than the solid bloc that emerged. The East-West conflict evaporated in turn when this informal empire collapsed in 1989. Likewise, the ability of the United States to fight the 1991 Persian Gulf War depended crucially on the protectorate it established over the Gulf states, wherein they gave up their independent foreign policies and followed Washington's lead in the crisis, and on the multilateral coalition it created to bind its own hands. Neither the Cold War nor U.S. hegemony in the Persian Gulf would have taken the form it did without the particular security institutions that lay at their core and eventually came to be among their most prominent characteristics.

The search for how and to what extent international institutions "matter" has largely played out in the realm of international political economy.² To the extent that scholars look for institutional effects, it is mostly at the level of universal or at least broad-based multilateral institutions in the areas of trade,

David A. Lake is Professor and Chair of Political Science at the University of California, San Diego, and co-editor of *International Organization*.

For comments on previous versions of this article, I would like to thank Peter Gourevitch, Robert Keohane, Lisa Martin, Celeste Wallander, and Barbara Walter.

1. For contrary views, see Robert Jervis, "Security Regimes," *International Organization*, Vol. 36, No. 2 (Spring 1982), pp. 357–378; and John J. Mearsheimer, "The False Promise of International Institutions," *International Security*, Vol. 19, No. 3 (Winter 1994/95), pp. 5–49.

2. See Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, N.J.: Princeton University Press, 1984); Lisa L. Martin, *Coercive Cooperation: Explaining Multilateral Economic Sanctions* (Princeton, N.J.: Princeton University Press, 1992); and Ronald B. Mitchell, "Regime Design Matters: Intentional Oil Pollution and Treaty Compliance," *International Organization*, Vol. 48, No. 3 (Summer 1994), pp. 425–458. For reviews of this literature, see Robert O. Keohane and Lisa L. Martin, "The Promise of Institutional Theory," *International Security*, Vol. 20, No. 1 (Summer 1995), pp. 39–51; and Lisa L. Martin and Beth A. Simmons, "Theories and Empirical Studies of International Institutions," *International Organization*, Vol. 52, No. 4 (Autumn 1998), pp. 729–757.

finance, standards, and so on. With several noteworthy exceptions,³ analysts presume that in the “high politics” realm of security affairs, states will eschew institutions and depend on their own unilateral capabilities.

This presumption rests on a severe truncation of the variation in international institutions. Even anarchic institutions—those premised upon the full sovereignty of all members—can solve problems of coordination among states, as in the case of multilateral sanctions, and influence international politics. But where national interests that would be significantly compromised by opportunism are at stake, states can be expected to either avoid cooperation or form more hierarchic institutions, such as informal empires and protectorates, that exert greater constraints on their members. Through the common practice of restricting analysis to only anarchic institutions, scholars bias their selection of either their independent or dependent variables. I hope to show here that we can better grasp the impact of institutions on world affairs by broadening the range of variation.

Both neorealists and neoliberal institutionalists are blinkered by their common assertion that all important relationships within the international system are anarchic. International relations is a variegated system composed of a range of institutions, some anarchic, some hierarchic. Identifying the elements of hierarchy at the core of both the Cold War and Gulf War reveals how taking anarchy as the defining characteristic of international politics—although useful in some circumstances as a simplifying device—actually distorts our vision.

Current studies seek to measure the effects of institutions by examining the behaviors of member states. Because institutions are typically created by these same states, however, it is difficult to identify the independent effects of interests and institutions. Here, I attempt to gauge the effects of institutions by assessing the reactions of nonmembers to institutional creation. Given underlying political cleavages, such as those in the Cold War, if third parties react strongly to institutional innovations, we can infer that the institution itself was relatively important. If these other parties react weakly or not at all, we can conclude that the institution exerted little effect. I also employ counterfactual reasoning to augment and support these inferences.

3. Celeste A. Wallander, *Mortal Friends, Best Enemies: German-Russian Cooperation after the Cold War* (Ithaca, N.Y.: Cornell University Press, 1999); Helga Haftendorn, Robert O. Keohane, and Celeste A. Wallander, eds., *Imperfect Unions: Security Institutions over Time and Space* (New York: Oxford University Press, 1999); and John S. Duffield, “International Regimes and Alliance Behavior: Explaining NATO Conventional Force Levels,” *International Organization*, Vol. 46, No. 4 (Fall 1992), pp. 819–855.

This exercise yields two conclusions that together offer small comfort to either neorealists or neoliberal institutionalists. First, institutions do matter, and they are likely to matter in security affairs precisely because important national interests are at stake. The skepticism of neorealists toward institutions is unwarranted. Second, anarchic institutions may resolve coordination problems and enable cooperation. But because they do not greatly constrain their members, such institutions are less effective when participants have incentives to defect. Anarchic institutions may produce large aggregate benefits by facilitating trade liberalization or exchange rate coordination, for instance, but they do not influence the behavior of states to the same extent as hierarchic institutions.

The remainder of this article seeks to explicate the above claims. The first section identifies a more complete range of security institutions in world politics and examines problems of inference in existing studies. The second section provides theoretical tools for explaining institutional variation. The third probes the role of institutions in the Cold War, and the fourth does the same for the Persian Gulf War. The final section discusses the proper role of security institutions in the twenty-first century.

Security Hierarchies

Robert Keohane, a leading neoliberal institutionalist, defines institutions as “persistent and connected sets of rules (formal or informal) that prescribe behavioral roles, constrain activity, and shape expectations.”⁴ This definition is widely shared in the field of international relations, even by prominent critics of neoliberal institutionalism.⁵ Although Keohane and other neoliberal institutionalists typically focus on “cooperation under anarchy,” nothing in this definition limits institutions to voluntary agreements between sovereign states. Indeed this definition is similar to those of Douglass North, who is primarily interested in the role of authoritative states in setting property rights, and Jack Knight, who is broadly concerned with the role of institutions in social conflict.⁶ Institutions may be hierarchic as well as anarchic, forged through

4. Robert O. Keohane, *International Institutions and State Power: Essays in International Relations Theory* (Boulder, Colo.: Westview, 1989), p. 163.

5. Mearsheimer, “The False Promise of International Institutions,” p. 8, offers a similar and largely consistent definition.

6. Douglass C. North, *Structure and Change in Economic History* (New York: W.W. Norton, 1981), pp. 201–202; and Jack Knight, *Institutions and Social Conflict* (New York: Cambridge University Press, 1992), pp. 2–3.

coercion or voluntarily negotiated, and rest on differences in power or equality. What is important is that the relevant actors share an understanding of the rules that shape and constrain their interactions.⁷

Although neorealists and neoliberal institutionalists do not assume that all relationships and in turn all institutions within the international system are anarchic, they exclude nonanarchic actors and relationships from their purview through a series of progressively restrictive assertions. Recognizing that other types of actors may exist, Kenneth Waltz, the dean of neorealism, reasonably posits that “states are not and never have been the only international actors. But then structures are defined not by all of the actors that flourish within them but by the *major* ones.” Waltz then asserts that “so long as the *major* states are the major actors, the structure of international politics is defined in terms of them.”⁸ Keohane similarly writes that “institutionalist theory assumes that states are the *principal* actors in world politics.”⁹ Notice the narrowing in both schools. Although open to variations in actors and relationships, they posit that anarchy is the foundation of international structure because it characterizes relations between the *major* actors.

Anarchy is indeed a useful assumption for some purposes, especially for explaining relations among great powers. Yet at the dyadic level, at least, anarchy is only one possible set of relationships. Institutions in security affairs are considerably broader than the alliances, concerts, and collective security organizations that are normally taken to define the set.¹⁰ A more complete accounting would include such increasingly hierarchic security institutions as:

- spheres of influence, in which dominant states prevent subordinates from entering into security relationships with third parties (the United States and

7. To illustrate, slavery and imperialism are both institutions with long historical practice. Neither rests on anarchy, voluntarism, or equality—and neither produces equitable outcomes. Both, however, are premised on clear (if usually informal) rules understood by all parties that prescribe roles, constrain behavior, and shape expectations.

8. Kenneth N. Waltz, *Theory of International Politics* (Reading, Mass.: Addison-Wesley, 1979), pp. 93–94 (emphasis added).

9. Robert O. Keohane, “Institutional Theory and the Realist Challenge after the Cold War,” in David A. Baldwin, ed., *Neorealism and Neoliberalism: The Contemporary Debate* (New York: Columbia University Press, 1993), p. 271 (emphasis added). For contrast, see Robert O. Keohane and Joseph S. Nye, *Power and Interdependence: World Politics in Transition* (Boston: Little, Brown, 1977).

10. This is not to say that scholars have not varied institutions at all, but these are typically minor variations in rules over decisionmaking within what remain broadly anarchic institutions. On dimensions of variation in institutions, see Miles Kahler, *International Institutions and the Political Economy of Integration* (Washington, D.C.: Brookings, 1995), pp. 2–6. See also Barbara Koremenos, Charles Lipson, and Duncan Snidal, “The Rational Design of International Institutions,” *International Organization*, Vol. 55, No. 4 (forthcoming 2001).

Latin America under the Monroe Doctrine; the Soviet Union and Finland after 1945);

- protectorates, where subordinate states yield control over their foreign and defense policies to dominant powers (Great Britain and Afghanistan until World War II; the United States and the Federated States of Micronesia today);
- informal empires, through which dominant states control substantial areas of policy in subordinate polities—even some typically regarded as purely domestic—but subordinates continue to interact with third parties on the basis of sovereignty (the United States and Central America under the Roosevelt Corollary to the Monroe Doctrine; the Soviet Union and Eastern Europe under the Brezhnev Doctrine); and
- empires, where dominant states formally control subordinates that retain no independent international “personality,” and therefore no right to enter international agreements in their own name.¹¹

Perhaps because of their focus on systemic anarchy, international relations scholars typically ignore such hierarchies. Although it is true that these are institutions that normally govern relations between major and minor states, excluding such institutions unnecessarily ignores important phenomena—including nineteenth-century imperialism, arguably one of the fundamental causes of World War I. As I hope to show below, ignoring such international hierarchies distorts our understanding of both the Cold War and Persian Gulf War.

WHY HIERARCHY MATTERS

By focusing exclusively on anarchic institutions, analysts not only ignore important phenomena but also introduce a strong selection bias that significantly weakens their ability to draw causal inferences.¹² Limiting the range of variation in institutions as a dependent variable underestimates the effect of independent or causal variables, suggesting that those factors that produce international institutions are more significant than past studies have revealed.¹³

11. David A. Lake, *Entangling Relations: American Foreign Policy in Its Century* (Princeton, N.J.: Princeton University Press, 1999), chap. 2. An analogous continuum of economic institutions would range from preferential trading agreements to customs unions with common external tariffs to economic unions.

12. Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, N.J.: Princeton University Press, 1994), pp. 129–135, 137–138.

13. See Lake, *Entangling Relations*, chap. 3 and pp. 270–271. Most important, this selection bias has led analysts to largely ignore what I call “governance costs” (see below).

Transactions costs, informational asymmetries, the need for credibility, and other variables that neoliberal institutionalists have used to explain institutions may be more important than even they recognize.¹⁴

Limiting variation in institutions as an independent variable—the problem addressed in this article—makes estimates of their effects more uncertain. Indeed selecting only one type of institution for analysis renders any such estimate highly problematic. Implicit in most studies is a comparison of anarchic institutions, which may be common internationally, and hierarchic institutions, which are often presumed to characterize domestic political systems.¹⁵ The real question is not whether international institutions matter in an absolute sense, but how much and in what ways they matter relative to courts, legislatures, and other hierarchic institutions that are presumed to characterize domestic politics. If one were to plot points for institutions of increasing hierarchy on the X axis, thereby capturing this comparison, and effectiveness on the Y axis, the research design most commonly employed in international relations would produce a stack of points near the “anarchic” origin, with some vertical dispersion reflecting natural variation in effectiveness. With no further observations at other values of the X axis, almost any regression line drawn through this stack of points would be just as valid as any other, thus preventing us from determining whether international institutions matter a little, a lot, or not at all. Given this research design, it is not surprising that reasonable scholars have been able to read the record of institutional effectiveness very differently.¹⁶

IDENTIFYING INSTITUTIONAL EFFECTS

In addition to the selection biases just noted, the study of institutions has been bedeviled by what Robert Keohane and Lisa Martin term an “endogeneity trap.”¹⁷ That is, if institutions are constructed to serve their members, identifying the independent effects of institutions is extremely hard. If institutions are endogenous, demonstrating that they matter beyond the interests that gave rise to them in the first place is well-nigh impossible—and explains why the

14. On transactions costs and information, see Keohane, *After Hegemony*, pp. 85–98. On credibility, see Martin, *Coercive Cooperation*, pp. 11–12.

15. See Waltz, *Theory of International Politics*, pp. 81–93. For an alternative, see Helen Milner, “The Assumption of Anarchy in International Relations Theory: A Critique,” *Review of International Studies*, Vol. 17, No. 1 (January 1991), pp. 67–85.

16. Although the problem of institutional effectiveness has not been formulated in precisely this manner, recognizing it helps explain the focus of many institutionalists on the European Union. As the most hierarchic of the voluntarily negotiated institutions, it provides crucial analytic leverage on the question of institutional effectiveness.

17. Robert O. Keohane and Lisa L. Martin, “Institutional Theory as a Research Program,” in Colin Elman and Miriam Fendius Elman, eds., *Progress in International Relations Theory* (forthcoming).

initial study of international regimes was drawn to periods “after hegemony” where scholars could reasonably argue that interests had changed but institutions had not.¹⁸

In this article I employ two methods to estimate the effects of institutions. Neither provides a precise metric against which we can measure the effects of institutions, but both aid in revealing whether institutions matter a little, a lot, or not at all. First, I examine not the behavior of the members of an institution but the reactions of third parties to the creation of the institution. The greater the change in the behavior of nonmembers, we can infer, the greater the effect of the institution. This method is particularly appropriate for the kinds of conflictual relations examined here.¹⁹ As members and nonmembers alike are assumed to act in their self-interest, it is still difficult to separate the effect of the institution per se from the larger set of actions and reactions, but careful attention to sequence and the perceptions of the actors themselves can shed some light on causal import.

Second, I employ counterfactual thought experiments, asking the reader to imagine a different institution and then to project the logical outcome of this imagined world.²⁰ Because the Cold War and Persian Gulf War were unique events, we have only one observation for each. We cannot rewind the tape of history, modify the institutions, and replay these events. To establish the importance of institutions, then, I posit the next most likely alternative institution and try to assess what would have changed as a consequence. Counterfactual reasoning is entirely consistent with the type of equilibrium analysis employed throughout this article. Any rationalist account of decisionmaking presupposes that actors compare across alternatives and consider what is likely to happen “off the equilibrium path.” Counterfactual analysis is merely a structured way of inquiring into the consequences of “near misses.”²¹ It remains an imperfect method, of course, both because we lack completely specified theo-

18. On static institutions but shifting interests, see Stephen D. Krasner, “Regimes and the Limits of Realism: Regimes as Autonomous Variables,” in Krasner, ed., *International Regimes* (Ithaca, N.Y.: Cornell University Press, 1982), especially pp. 359–361.

19. But it is not limited to relations between antagonists. The creation of the European Free Trade Association in response to the formation of the European Economic Community is a similar indicator of the latter’s importance.

20. On counterfactual analysis, see James D. Fearon, “Counterfactuals and Hypothesis Testing in Political Science,” *World Politics*, Vol. 43, No. 2 (January 1991), pp. 169–195; and Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, N.J.: Princeton University Press, 1996).

21. See Bruce Bueno de Mesquita, “Counterfactuals and International Affairs: Some Insights from Game Theory,” in Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, pp. 211–229; and Barry R. Weingast, “Off-the-Path Behavior: A Game-Theoretic Approach to Counterfactuals and Its Implications for Political and Historical Analysis,” in *ibid.*, pp. 230–243.

ries and because human interaction contains a random element. Nonetheless, counterfactual analysis gives us some analytic and empirical purchase on institutional effects.

Restricting analysis to the next most likely alternative is vital to maintaining the plausibility of any counterfactual analysis. By altering the independent variable as little as possible, one minimizes the changes in the antecedent conditions that are required. Within each case, as a result, I confine my discussion to what the theory suggests is the next most reasonable institutional choice. The plausibility of the counterfactual is then confirmed by an analysis of contemporary policy debates. In each case, the alternative suggested by the theory was actually advocated by a significant bloc of policymakers.

Counterfactual analysis cannot prove the causal role of institutions. Yet the plausible rereadings of history below make a strong case for the role of institutions. Counterfactual analysis also shows how a theory of institutional variation is essential for identifying institutional effects, and explains why, in an article on the effects of institutions, I devote so much space to a theory of institutional choice.

Explaining Institutional Choice

Institutions are both a product and a cause. Actors create institutions to serve their interests, and these institutions influence subsequent behavior. Institutions either enable actors to achieve outcomes that might otherwise be impossible or constrain actors from undertaking behaviors that would otherwise be chosen. They are intended to channel behavior in predictable ways, and this intent forms part of the explanation for the particular institution. It is this dual role that makes the separate effects of interests and institutions hard to distinguish.

Nonetheless, interests as commonly described in international relations seldom determine a particular institutional choice. To increase security or counter threats from revisionist states, countries may augment their own capabilities by cooperating with others. Yet the resources and efforts of the partners can be more or less effectively aggregated through an alliance, a protectorate, or an empire. Whether and how states can design an agreement to cooperate is determined not only by their interests but also by features of their relationship with one another. Are the interests of the parties aligned so that the first need not fear the defection of the second? Would the defection of the second impose high costs on the first? What are the costs of establishing a hierarchy through which the first would control more or fewer of the actions of the second, and

thus possibly limit the ability of the latter to shirk its obligations? Moreover, are the synergies from combining resources and efforts greater than the expected costs of defection and the likely costs of constructing some institution? Because the answers to these questions are not determined by the sorts of primordial security interests such as territorial integrity normally attributed to states, institutions themselves may have an independent effect on world politics.

Elsewhere I have explained the choice of security relationships in terms of three variables: joint production economies, which determine the gains from security cooperation and derive from scale economies, positive externalities (or beneficial spillovers from the actions of others), and the division of labor; the expected costs of opportunism, or the costs incurred when a partner defaults on an agreement to cooperate; and governance costs, the resources and efforts devoted to creating and maintaining the relationship. I argue that:

- the larger the joint production economies, the more willing states are to expend resources to obtain more effective cooperation;
- the expected costs of opportunism decline with hierarchy because subordinates are less able to act against the wishes of dominant powers; and
- governance costs rise with hierarchy. Dominant states must induce subordinates to give up their valued freedom and compensate them accordingly, engage in costly actions that signal their benign intent and restrict their own behavior so as to limit their ability to exploit their increasingly vulnerable partners, or coerce subordinates into relationships against their will.²²

It follows that anarchic institutions are most likely to be created under two circumstances. First, when joint production economies are small, a state will enter an institution only if it expects that its partner will live up to their agreement—because a high risk of opportunism is likely to negate the limited benefits of cooperation—and the agreement does not greatly constrain its own behavior, implying that the governance costs of cooperation are also low. In other words, when joint economies are small, we should expect the creation of anarchic institutions only when it is in the interests of the parties to abide by agreements and the accords themselves do not impose significant costs on their members. Anarchic institutions are chosen, in short, when the need to alter the behavior of the contracting parties is smallest. That such institutions often fail to modify significantly the behavior of states should not be surprising.

22. Lake, *Entangling Relations*, chap. 3.

Second, we can also expect anarchic institutions to arise when joint economies are large but governance costs rise rapidly with hierarchy. In these cases, the potential gains make cooperation attractive, but hierarchy is not a feasible solution for potential opportunism. States will then choose to cooperate as long as the expected costs of opportunism and governance costs do not exceed the benefits created by the joint economies. Despite a high probability or cost of opportunistic behavior by partners, states may still enjoy net gains from cooperation—even though the overall benefits will appear small as the large gains from joint economies are dissipated by frequent opportunism. Under these circumstances, we will typically observe lots of defections, states largely ignoring or accepting these defections, and—surprisingly—continuing cooperation.

Critics of the importance of international institutions correctly identify the limits of such anarchic relationships, but fail to recognize that the observed weakness of extant institutions is the product of rational selection: States get the institutions they need or can afford. In a way not appreciated by either institutionalists or their critics, anarchic institutions are most likely to be constructed by states precisely when cooperation produces little value added, states possess few incentives to defect, or states are willing to tolerate substantial opportunism.

Hierarchic institutions will tend to arise when more is at stake. Hierarchies are most likely when there are significant benefits from cooperation, such as technological scale economies that allow force to be projected at low cost and large positive externalities that make it as efficient to defend a larger area as a smaller one; high costs if a partner acts opportunistically, which in turn are a function of the specific assets in the relationship such as ports or forward defense sites in unique locations; and low costs for creating and maintaining a relationship that do not increase rapidly with hierarchy, such as those that arose from the power disparities between the European imperialists and the peoples they encountered during the age of expansion. When the potential benefits of cooperation are large and the costs of opportunism by partners are great, governance costs are the determining factor.

If the costs of creating a hierarchy are high, states will eschew cooperation—which is the traditional realist prediction. If the costs of building a hierarchy are not too high, states will invest in institutions to control the behavior of their partners. The greater the potential benefits of cooperation or the greater the risks of defection by partners, the more likely states will be to forgo cooperation or create a hierarchic institution. When cooperation is both attractive and risky, hierarchy is the likely result.

Security Institutions and the Cold War

Despite similar structural positions as the poles in a bipolar system, the Soviet Union and the United States adopted very different security institutions after World War II. Where the former imposed a relatively hierarchic informal empire, the latter chose a relatively anarchic institution (the North Atlantic Treaty Organization). This difference in security institutions mattered. Although both types of institutions influenced the course of international politics—and therefore mattered in an absolute sense—the Soviet Union’s informal empire prompted a greater reaction from the United States than NATO did from the Soviet Union.

INFORMAL EMPIRE IN THE EAST

The Soviet Union built an informal empire in Eastern Europe for at least three interconnected reasons.²³ First, the Soviet Union could gain greatly from cooperation with Eastern Europe. The significant changes in military technologies that reached fruition during or soon after the war—most notably mobile armored warfare and atomic weapons—spurred a shift in both superpowers to a forward-based defense posture. The Soviet Union wanted to defend itself against a resurgent Germany or a potentially hostile capitalist coalition as far west as possible. It also wanted to project force into the heart of Europe and, through its expanded ability to intimidate others, to influence events on a world scale.²⁴

Second, the costs to the Soviet Union of opportunistic behavior by its partners were very high. Forward defense depends on the existence of a complete perimeter; any gaps threaten to vitiate the whole by allowing opponents to surge through and threaten areas to the rear. Each state in Eastern Europe was therefore crucial to the success of this strategy, and thus the territory of each became a highly specific asset for the Soviet Union. Because one partner could not simply be replaced by another, each acquired power over Moscow. By

23. In addition to the contractual problems emphasized here, the Soviet Union also sought rents from its relationship with Eastern Europe. On autocracy, state rent-seeking, and imperialism, see David A. Lake, “The Rise, Fall, and Future of the Russian Empire: A Theoretical Interpretation,” in Karen Dawisha and Bruce Parrott, eds., *The End of Empire? The Transformation of the USSR in Comparative Perspective* (Armonk, N.Y.: M.E. Sharpe, 1997), pp. 30–62.

24. See David M. Glantz, *The Military Strategy of the Soviet Union: A History* (London: Frank Cass, 1992), especially pp. 176–180; and R. Craig Nation, *Black Earth, Red Star: A History of Soviet Security Policy, 1917–1991* (Ithaca, N.Y.: Cornell University Press, 1992). Edward N. Luttwak, *The Grand Strategy of the Soviet Union* (New York: St. Martin’s, 1983), p. 19, suggests that a forward-based strategy was not in fact central to postwar planning and politics.

threatening to defect, each could extract bargaining advantages from the Soviet Union—indeed by threatening to defect, any one partner could potentially appropriate all the benefits from cooperation received by the superpower. To avoid this outcome, the forward defense strategy required that the Soviet Union effectively control the behavior of its partners and thereby limit their demands. Some significant degree of hierarchy was essential.²⁵

Third, governance costs were low, largely because the costs of creating the informal empire had already been absorbed in the war against Germany. Hierarchy is usually costly. Dominant states can offer concessions to induce subordinates to give up their valued freedom. The Soviet Union rejected this course, at first, electing to extract resources from Eastern Europe rather than share its benefits from cooperation. Indeed, by one estimate, Moscow withdrew nearly \$1 billion per year from the region until 1956.²⁶ By the late 1950s, however, the flow of resources reversed, and by the 1980s the Soviet Union was subsidizing Eastern Europe to a total of about \$17 billion per year.²⁷ Dominant states can also build loyalty and compliance by credibly committing not to exploit their subordinates. There is little evidence that Moscow considered this route. Dominant states can also coerce subordinates into yielding control over their foreign and defense policies and, with sufficient force, even over their sovereignty. Imperialism by one sovereign state against another is rare, largely because the costs of conquering an organized, modern state are typically high; in most cases, the increment of control so gained is worth less than the incremental cost—that is, the marginal costs of hierarchy typically exceed the marginal benefits at something less than empire.²⁸ The Soviet Union, however, was in the nearly unique position of having established control over the states of its informal empire in the process of defeating Nazi Germany. Rather than being

25. The United States faced the identical problem in Asia, where the forward defense required bases in a consistent line along the Pacific perimeter. It responded with a series of hierarchic institutions, including a formal empire in Japan under the Occupation relaxing into a weak protectorate by the mid-1950s; a protectorate in the Philippines; and a formal empire in Micronesia. See Lake, *Entangling Relations*, pp. 142–148, 169, 180–185.

26. Zbigniew K. Brzezinski, *The Soviet Bloc: Unity and Conflict*, rev. and enl. (Cambridge, Mass.: Harvard University Press, 1967), pp. 285–286.

27. Randall Stone, *Satellites and Commissars: Strategy and Conflict in the Politics of Soviet-Bloc Trade* (Princeton, N.J.: Princeton University Press, 1995), p. 45. See also Valerie Bunce, “The Empire Strikes Back: The Transformation of the Eastern Bloc from a Soviet Asset to a Soviet Liability,” *International Organization*, Vol. 39, No. 1 (Winter 1985), pp. 1–46.

28. This is not to argue that conquest never pays in absolute terms; see Peter Liberman, *Does Conquest Pay? The Exploitation of Occupied Industrial Societies* (Princeton, N.J.: Princeton University Press, 1996). Rather it is to suggest that the marginal benefits of greater control that come with conquest typically exceed the marginal cost of acquiring that control. The relevant calculation is the costs and benefits of formal empire relative to informal empire.

deterred from establishing hierarchy by the high costs of defeating subordinates, the Soviet Union now had to consider only the costs of maintaining its informal empire—which, given the resource flows out of Eastern Europe in the immediate postwar years, may have been quite small (but which grew progressively over time).

Compliant regimes subservient to the Soviet Union further reduced the costs of maintaining the informal empire. By imposing political regimes that lacked significant domestic support and that therefore depended on the Soviet Union for their continued rule, Moscow created loyal agents in the heart of each state who would faithfully respond to guidance from the Kremlin. The lack of domestic legitimacy was not an impediment to the consolidation of Soviet authority, as often averred, but rather a central element in maintaining Russian control. Moscow's costs of domination and its need to engage in day-to-day monitoring of East European affairs were thereby reduced. Direct, imperial rule would have been far more costly—but it was also unnecessary.

The establishment of an informal empire in Eastern Europe did not proceed from any grand vision. Josef Stalin was an opportunist first, an imperialist second. He exploited Western trust and disinterest to manipulate events to the Soviet Union's advantage. Nonetheless, the informal empire was rooted in the large benefits of cooperation, the high risk of opportunism, and the low costs of governance—ideal conditions for creating a relatively hierarchic set of security institutions.

THE EFFECTS OF THE INFORMAL EMPIRE

The literature on who or what caused the Cold War is enormous, and the question will obviously not be settled here.²⁹ My concern is with the role of the informal empire in general and relative to other possible security institutions in stimulating the conflict. In the literature, four positions can be identified. Traditionalists place primary responsibility for the Cold War on the imperialist nature and desires of the Soviet Union. Revisionists argue for a greater responsibility for the United States, especially through its efforts to construct a liberal, capitalist international order that would rebuild Europe and reintegrate Germany but exclude the Soviet Union. Postrevisionists acknowledge a sort of moral neutrality in the conduct of the two superpowers, emphasizing the competition inherent in bipolarity and the series of seemingly innocent steps taken

29. For a review prior to the end of the Cold War, see John Lewis Gaddis, "The Emerging Post-Revisionist Synthesis on the Origins of the Cold War," *Diplomatic History*, Vol. 7, No. 3 (Summer 1983), pp. 171–190.

by each that were nonetheless misperceived or, following a security dilemma logic, were misconstrued by the other as more hostile and threatening than intended. Neotraditionalists, benefiting from recently opened Soviet archives and hindsight, have returned to a more traditionalist position.³⁰ They self-consciously describe the Soviet sphere in Eastern Europe as an empire, place responsibility for the conflict on Stalin's lust for total security or territorial expansion, and conclude that Stalin himself caused the Cold War.

Both postrevisionists and neotraditionalists place the creation of the informal empire in Eastern Europe at the center of their explanations for the Cold War. Where postrevisionists emphasize the misunderstandings that followed from the consolidation of Soviet control over this region, neotraditionalists assign more malign intentions to the Soviet Union. Nonetheless, for both, the creation of the informal empire was a crucial break in East-West relations.

The effects of the informal empire can be gauged in part by the sequence of Soviet actions and Western reactions in the immediate postwar years. As early as the Yalta conference in 1945, the United States and Great Britain complained about the communist regimes being installed in Bulgaria and Romania; within a year, communists controlled the governments in those two states and East Germany. By late 1947, communists were also in power in Poland and Hungary. The coup in Czechoslovakia on February 25, 1948, where the Red Army backed the seizure of power by the local communist party, was a decisive turning point for the West.³¹ Also important was the Soviet Union's demand, made two days after the Czech coup, that Finland sign a friendship treaty. These events signaled that the Soviet Union intended to exert firm control over the states of Eastern Europe and that it might have ambitions outside the region that it had inherited through its military successes and sacrifices in World War II. By early 1948, well before the outbreak of the Korean War, the West had concluded from Soviet behavior in Eastern Europe that some sort of postwar conflict was inevitable.

The United States had two principal reactions to these Soviet moves. First, it greatly expanded its level of defense spending. A strong desire for a balanced budget and fears of creating a garrison state had led President Harry Truman to propose deep cuts in defense spending immediately after World War II.

30. See John Lewis Gaddis, *We Now Know: Rethinking Cold War History* (New York: Oxford University Press, 1997); and Vojtech Mastny, *The Cold War and Soviet Insecurity: The Stalin Years* (New York: Oxford University Press, 1996).

31. See Joseph Held, ed., *The Columbia History of Eastern Europe in the Twentieth Century* (New York: Columbia University Press, 1992).

Even the Truman Doctrine and then the Marshall Plan did not seriously alter budget priorities in Washington. Resisting strong pressure from the armed services to increase spending, Truman submitted a modest defense budget in January 1948 of only \$9.8 billion in new spending authority. Following the Czech coup, he planned a supplemental request of \$1.5 billion, which under pressure from the services grew to \$3.1 billion before it was finally submitted to Congress. The legislature then significantly augmented the president's request, bringing the total in new spending authority to \$13.9 billion—\$4.1 billion more than originally envisioned by Truman, for a total increase of more than 42 percent.³² Proportionately, this represents the largest peacetime increase in military spending in U.S. history.

Second, the United States also created a countervailing alliance, its first in more than 150 years (see below). Prior to the Czech coup, Washington had consistently rebuffed Britain's suggestions for a formal military alliance. As late as January 1948, the United States had rejected outright London's calls for a regional security arrangement. Within a month after the Czech coup, however, Secretary of State George Marshall reversed this position, endorsed Britain's proposal for an alliance, and began talks on a possible security treaty. In June the Senate passed the Vandenberg Resolution, articulating its support for any regional security arrangements the president might choose to negotiate. That same month, the United States informed Britain that it was ready to begin formal negotiations. Diplomats from Belgium, Britain, Canada, France, Luxembourg, and the Netherlands convened in Washington in July. The North Atlantic treaty was signed in April 1949 and ratified by the United States just three months later.³³

32. On the budget struggle following the Czech coup, see Michael J. Hogan, *A Cross of Iron: Harry S. Truman and the Origins of the National Security State, 1945–1954* (New York: Cambridge University Press, 1998), pp. 91–113.

33. The negotiation of the North Atlantic treaty is summarized in Lake, *Entangling Relations*, pp. 133–142. Of course, some argue that it was the Korean War that put the “O” in NATO, making it a stronger and more distinctive security institution. At one level, this is true, but it does not vitiate the argument developed here. The relevant question for my purposes is whether the creation of the informal empire in Eastern Europe, which occurred between 1945 and 1948, had an important effect on the West and especially the United States. On this score, I am confident that the Truman Doctrine, the Marshall Plan, the signing of the North Atlantic treaty, and other policy initiatives prior to the outbreak of the Korean War were consequential departures from American tradition. At a second level, however, I also question the extent to which the creation of an integrated command structure within NATO transformed the organization. Many of the key elements were already in place prior to 1949, including the indefinite stationing of American occupation troops in Germany as hostages to a Soviet invasion and the commitment to come to the aid of alliance partners in case of attack. In my view, the institutional superstructure built after 1951 is the culmination of a process started years earlier.

The significance of the Soviet's informal empire is reflected in the magnitude of the West's response. In reply to events in Eastern Europe, the United States rebuilt a peacetime army. It also broke decisively with its historic policy of unilateralism and created a security institution of its own. The United States did of course have other reasons to alter its strategy and form NATO (explored below). Nonetheless, the consolidation of Soviet control over Eastern Europe had a significant impact on the West and especially the United States.

The effect of the informal empire can also be assessed through counterfactual analysis. Theory indicates that the relevant alternative is a slightly less hierarchic security institution, perhaps a Soviet-led protectorate or a sphere of influence. Supporting this prediction, the former was actually proposed by Czechoslovakia's last noncommunist president, Edvard Beneš, as a means of escaping outright Russian domination, and the latter was imposed on Finland.³⁴ Although we will probably never know as much about the internal deliberations in the Soviet Union as in more democratic societies, there is some evidence that such alternatives were considered. Most provocative are hints that some leaders were sufficiently dissatisfied with Stalin's strategy toward Eastern Europe to consider support for a neutral Germany.³⁵ Despite the strong forces pushing the Soviet Union toward an informal empire, other less hierarchic options appear to have been considered.

A less hierarchic security institution in Eastern Europe would have had four primary effects. First, the Soviet Union would have faced a more fractious set of partners, complicating relations and constraining its ability to unilaterally determine the fate of the Eastern bloc. More independent voices within the bloc would have reduced Moscow's ability to challenge the West, especially by projecting force from its less secure bases in Eastern Europe. Second, the forward defense strategy would have been less effective. Consequential gaps might have arisen that the West could exploit, and Soviet military power would have appeared less foreboding. Third, the increased bargaining leverage obtained by the states of Eastern Europe would have boosted their share of the benefits earned from cooperation and diminished the attraction of the informal empire for the Soviet Union, perhaps facilitating a natural rollback. Fourth, the Soviet Union would have appeared less aggressive to the West, and

34. Gaddis, *We Now Know*, p. 17.

35. Lavrentii Beria, Stalin's chief of the secret police, proposed the creation of a neutral Germany as a major plank in his platform in the leadership struggle that ensued after Stalin's death in 1953. Beria was of course eventually defeated in this struggle, but his belief that such a proposal would help rally his Politburo colleagues to his side suggests at least some latent dissatisfaction with Stalin's approach to Eastern Europe. *Ibid.*, p. 130.

there would have been fewer misunderstandings with the potential to escalate into major war.

If the sequence of events and analysis outlined above is correct, each of these effects would likely have reduced conflict with the West. Cooperation between the United States and Western Europe and the level of Western military spending would have declined, and the advanced preparations for war—including the permanent stationing of U.S. troops in trip-wire positions—would have been scaled back. A smaller Soviet threat would have elicited a more muted Western response, and the cycle of superpower escalation would have slowed. This is not to argue that the Cold War would have been averted if only the Soviet Union had not created an informal empire in Eastern Europe. The two superpowers would inevitably have had tensions and disagreements, and would have clashed over the nature of the international order that each desired. But had the Soviet Union chosen a less imperialist institution to regulate its relations with Eastern Europe, the tensions and hostilities that characterized this era would have been significantly mitigated.

In this case, the security institution created by the Soviet Union clearly appears to have mattered. Indeed the final evidence on how great an impact the informal empire had on postwar politics is provided by the collapse of communism in Eastern Europe. As soon as the Berlin Wall fell, many analysts proclaimed the end of the Cold War. The Soviet Union did not disintegrate for another two years, however, and its aggregate power and nuclear deterrent did not evaporate overnight. The structure of the international system did not suddenly shift. Nonetheless, the West's view of the Soviet Union and the threat posed by Russia changed dramatically—demonstrating forcefully how strongly the informal empire affected international politics.

THE WESTERN ALLIANCE

The relatively anarchic alliance built by the United States after World War II shares some essential similarities and causes with the Soviet Union's informal empire. First, as in the Soviet case, there were large benefits for the United States from cooperation with its European allies. The same technological advances that prompted Moscow to choose a forward-based defense led Washington to adopt a similar strategy. The United States was not only more exposed to the force that could be projected by others, a lesson imprinted on the American mind at Pearl Harbor, but it was also able to extend its own reach abroad more easily.

In addition, because its partners were more trustworthy (see below), the United States was able to build a small but consequential division of labor

within NATO in which it specialized in nuclear and naval forces, while Europe concentrated on land forces and tactical air power.³⁶ Finally, although they free rode on the United States, counting on their superpower ally to protect them despite their own limited defense efforts, the European allies did contribute something to the joint defense and thereby relieved the Americans of an even greater burden.³⁷ NATO also developed a set of institutional practices to enhance the contributions of the allies, such as the annual review that called laggards to task publicly and shamed them into doing more. The question is not whether the United States contributed a disproportionate amount toward the collective defense—which it surely did—but whether the alliance facilitated greater equality of effort than if each acted independently, which is also surely the case.

Second, there was some risk that the Europeans would act opportunistically, but this was substantially lower for the United States than for the Soviet Union. Where the thinness of its East European buffer led Russia to adopt a rigid forward strategy, the United States took advantage of its greater strategic and economic depth. With a series of allies extending from Germany through France to the Iberian Peninsula and British Isles, American defense plans called for a gradual pullback to the European periphery in the face of a Soviet assault and then a massive mobilization of economic and military muscle to repel the Red Army. The most forward-deployed American troops were intended to play a more political than military role. The greater flexibility enjoyed by the United States lessened the costs of possible opportunism by its partners. Specifically, it prevented any one ally from being able to undermine its forward strategy or, by threatening to defect, from being able to appropriate the gains from cooperation earned by the United States. Indeed France's withdrawal from NATO's command structure in 1966 had little noticeable effect on the alliance's defense doctrine.³⁸ This is not to suggest that there were no specific assets at risk in Europe: In the immediate postwar years, for instance, the Azores (governed by Portugal) and Iceland were crucial refueling stops, and

36. Simon Lunn, *Burden-Sharing in NATO*, Chatham House Papers 18 (London: Routledge and Kegan Paul, 1983), pp. 10–11; and Lawrence S. Kaplan, *NATO and the United States: The Enduring Alliance* (Boston: Twayne, 1988), p. 39.

37. Lake, *Entangling Relations*, pp. 156–162. The appropriate counterfactual here is unclear. I take as my baseline the expectation that without NATO, the Europeans would have free rode on each other and the United States to an even greater extent, much as they tried to do during the interwar years.

38. For a comparison of the Western reaction to France's withdrawal from NATO and Russia's reaction to Hungary's attempted withdrawal from the Warsaw Pact (in 1956), see Gaddis, *We Now Know*, p. 219.

some military bases on the continent were crucial. But it is to suggest that the more flexible forward defense created far fewer specific assets in the West than in the Soviet bloc.

The specific assets that were generated within NATO—such as Britain's specialization in mine sweepers—arose from the limited division of labor created under that institution. These assets were the political products of a relatively high degree of trust; because member states believed that NATO would work effectively, they were willing to delegate responsibility to the organization and become dependent, to some extent, on the capabilities provided by other members. Even so, the members consciously restrained the division of labor to limit further the costs of opportunism. The United States invested in nearly all aspects of conventional defense. More telling, fearful of relying exclusively on the American nuclear umbrella, Britain and France developed politically important if militarily negligible nuclear arsenals.

Overall, then, there were fewer specific assets at risk in the West than in the East. Many of the Western assets, moreover, were hedged by cautious countries building a measure of redundancy into their force structures. As a result, hierarchy was less necessary in relations between the United States and its partners than between the Soviet Union and Eastern Europe.

Third, had the United States sought greater hierarchy in Western Europe, the costs would have been prohibitive. Unlike the Soviet Union, which fought World War II in the East essentially on its own, the United States worked within a coalition of states. It did not unilaterally liberate the Western victims of Nazi aggression. The war was a team effort whose success was premised on the restoration of legitimate national governments. After the war, of course, the United States did play a role in shaping European politics, and actively opposed communist parties in France and Italy. Although it could and did withdraw support from regimes it disliked, the United States did not have a free hand to select governments for its partners.³⁹ Thus, for the United States to have imposed an informal empire on the states of Western Europe similar to that imposed by the Soviet Union in the East would have fractured cooperation in the war against Germany, significantly raising the costs of victory there, or required overt coercion against its partners after the war. In short, the governance costs of imposing a hierarchic security institution on Western Europe

39. On the intervention of the United States into the domestic politics of its postwar allies, see Irwin Wall, *The United States and the Making of Postwar France, 1945–1954* (New York: Cambridge University Press, 1991), especially pp. 2–4, 63–95; and James Edward Miller, *The United States and Italy, 1940–1950: The Politics and Diplomacy of Stabilization* (Chapel Hill: University of North Carolina Press, 1986), pp. 213–249.

were not already absorbed as a by-product of defeating the Nazis and, thus, were much higher for the United States than for the Soviet Union.⁴⁰ Overall, hierarchy was both less necessary and more expensive for the United States.⁴¹ An anarchic set of security institutions centered in NATO was the logical result.

EFFECTS OF THE WESTERN ALLIANCE

Revisionist historians place substantial responsibility for the Cold War on the United States, both for its efforts to create a capitalist international order that excluded the Soviet Union and for overreacting to and purposely exaggerating the Soviet threat.⁴² Key events in this view are the announcement of the Truman Doctrine, where the president was encouraged by Senator Arthur Vandenberg (R-Michigan) to “scare the hell out of the American people,”⁴³ and the Marshall Plan, which laid the foundation for rebuilding Europe along American lines.

Although created in a period of escalating mistrust, the Western alliance appears to have elicited only a modest reaction from the Soviet Union. This suggests that its effects on Western cooperation, though not unimportant, were smaller than the effects of the Soviet Union’s informal empire.⁴⁴

Most important, there was little response to the actual signing of the North Atlantic treaty.⁴⁵ Truman and other American officials acknowledged that

40. Based on an assessment of the American occupations of Germany and Japan after the war, it also appears that the governance costs of maintaining a hierarchy over Europe would have been prohibitive. See Lake, *Entangling Relations*, pp. 190–192.

41. Although he exaggerates the degree of hierarchy, it was the positive attraction of American security and economic cooperation, rather than a threat of political domination, that led Geir Lundestad to dub America’s role in Europe after the war an “empire by invitation.” Lundestad, *The American “Empire”* (New York: Oxford University Press, 1990).

42. The classic revisionist history is William Appleman Williams, *The Tragedy of American Diplomacy*, rev. and enl. (New York: Dell, 1972), especially pp. 229–275.

43. Walter LaFeber, *The American Age: United States Foreign Policy at Home and Abroad, 1750 to the Present*, 2d ed. (New York: W.W. Norton, 1996), p. 477.

44. One might argue that Stalin factored the possible creation of a NATO-like institution into his earlier calculus in determining whether to establish an informal empire in Eastern Europe. If so, then the failure to react to the creation of NATO is not evidence that it did not matter. This is unlikely, however, because there was no strong expectation, especially in the first postwar years, that the West would ally. Even if this was considered a possibility, the realization of this event ex post should have led the Soviet Union to update its strategy.

45. Moscow certainly knew about the secret Pentagon talks that eventually led to the formation of NATO through Donald McLean, a Soviet spy in the British foreign service who participated in those discussions, but the outbreak of the Berlin crisis one day after the United States informed Britain that it was ready to begin formal negotiations appears to be more coincidence than response. One week after the formal treaty was unveiled, however, Stalin did send word that he was ready to discuss calling off the Berlin blockade. Don Cook, *Forging the Alliance: NATO, 1945–1950* (New York: William Morrow, 1989), p. 221.

Moscow was likely to see the treaty as threatening.⁴⁶ Yet Vojtech Mastny writes:

Despite his concern about NATO, Stalin did not initially act as if he worried too much about its presumed aggressive intentions. The Soviet military response to its creation was moderate rather than alarmist: a 20 percent increase in defense spending, mainly calculated for public effect [compared to a 42 percent increase in the United States following the Czech coup], a bolstering of the troops stationed in East Germany, the establishment of an office to supervise the modernization of the armed forces of the Eastern European allies. Having previously preferred to keep them weak as potentially disloyal, Moscow now pressed forward a purge of their officer corps, to replace holdovers with more reliable party men.⁴⁷

Moreover, Warren Cohen suggests that in the months preceding the formal signing of the North Atlantic treaty, the Soviet Union was increasingly accepting of the status quo. Far from challenging the developing treaty, Russia appeared on the defensive.⁴⁸

The rearmament of Germany and the integration of its forces into NATO in 1955, of course, led directly to the signing of the Warsaw Pact. Nonetheless this agreement—anarchic in form but reflecting the deeper hierarchy of the informal empire already in place—was largely an empty shell. As Craig Nation observes, “The creation of the [Warsaw Pact] did not alter the balance of forces in the central European theater . . . and the pact added little to the controls already exerted by Moscow since 1948 over the national military contingents of the bloc states.”⁴⁹ Even here, the Soviet response was remarkably mute.

The formation of the Western alliance undoubtedly contributed to the deepening of the Cold War. By aggregating the individual capabilities of the Western states, it posed a more potent threat to the Soviet Union. It confirmed Russia’s expectations of conflict with the West. But as expected by the theory, which predicts more hierarchic institutions when the need to modify the incentives of members is greater, the formation of the Western alliance did not have the decisive effect on Russia that the creation of the Soviet Union’s informal

46. Melvyn P. Leffler, *A Preponderance of Power: National Security, the Truman Administration, and the Cold War* (Stanford, Calif.: Stanford University Press, 1992), p. 308.

47. Mastny, *The Cold War and Soviet Insecurity*, p. 74.

48. Warren Cohen, *America in the Age of Soviet Power, 1945–1991*, Vol. 4, *The Cambridge History of American Foreign Relations* (New York: Cambridge University Press, 1993), p. 49.

49. Nation, *Black Earth, Red Star*, p. 219. Brzezinski, *The Soviet Bloc*, p. 174, finds the pact important for its political symbolism. West Germany’s remilitarization within NATO might also have contributed to a tightening of relationships within the informal empire, reflected in the crackdown on Hungary in 1956.

empire had on the West. In this case, the security institution mattered, but more at the margin.

The most relevant and plausible alternative for the United States was a unilateral strategy designed to deter the Soviet Union through American strength alone. This was indeed the other option discussed in Washington. The unilateralists after World War II were emphatically not isolationists. They recognized the importance of the Soviet threat and America's new global security interests, but they opposed formal alliances and supported instead an extended Monroe Doctrine that would stretch from the Western Hemisphere to Europe and beyond. Senator Robert Taft (R-Ohio), the leader of this group within Congress, argued strongly that the United States should stand aloof from potential allies, fortify itself, and build a nuclear arsenal capable of "pulverizing Russia" if the communist power sought to expand abroad at the expense of American interests.⁵⁰

The primary effect of a unilateral strategy would have been to raise the defense burden on the United States or force it to produce a lower level of overall security (accept more risk). To the extent that there were benefits from cooperating with Western Europe—either through bases from which the United States could project force, less free riding than would otherwise have been the case, or the division of labor—unilateralism would have required greater defense efforts at home to produce the same level of security. The more effective the Western alliance was at producing benefits from cooperation, the greater the additional burden that a unilateral strategy would entail. In actuality, there is always a guns-butter trade-off, and the steeper the defense burden the less private consumption society is willing to sacrifice to gain additional security. Unilateralism would most likely have produced some compromise between additional defense spending and less security for the United States.

Based on this compromise, the secondary effect of unilateralism depends on the extent of the reduction in security selected by the United States and the reaction of the Soviet Union. The smaller the reduction in the level of security pursued—or the greater the unilateral defense burden the United States was willing to tolerate—the smaller the net effect on the Cold War, as presumably the Soviet Union cared more about the overall ability of the United States to fight a potential war than about whether the burden of preparing for this war was shared with the Europeans. To the extent that the United States reduced its

50. On unilateralism, see Lake, *Entangling Relations*, especially p. 132. On Taft, see Timothy Ireland, *Creating the Entangling Alliance: The Origins of the North Atlantic Security Organization* (Westport, Conn.: Greenwood, 1981), especially p. 209.

overall level of preparedness, much depends on the expected reaction of the Soviet Union. If revisionists are correct and the Soviet Union largely responded to initiatives taken by the United States, a lower level of defense preparedness might have quelled the security dilemma, produced fewer misunderstandings, and led to less chilly relations. If the neotraditionalists are correct and the Soviet Union was in fact more imperialist, a lower level of preparedness in the United States might well have led to a more aggressive Soviet response, deepening the Cold War even further.

The evidence provided by the limited response of the Soviet Union to the creation of NATO, however, suggests that unilateralism would not have made a major difference in either American policy or the depth of the Cold War. If the alternative to NATO was unilateralism, and the effect of the latter would have either mitigated the Cold War or opened up significant opportunities for Soviet expansion, we would then expect the reaction of the Soviet Union to the formation of the Western alliance to have been more severe—it would have either changed its own policies to vitiate the need for NATO or reacted more aggressively to intimidate the Europeans into possibly dropping out of the nascent organization. It did neither.

SECURITY INSTITUTIONS AND ORIGINS OF THE COLD WAR

The security institutions adopted by both the United States and the Soviet Union did affect the depth and duration of the Cold War. NATO and the informal empire in Eastern Europe mattered to the behavior of the other. It is also possible to construct plausible counterfactual histories in which different institutions would have led to different outcomes than those actually experienced.

Nonetheless, the informal empire, as expected by the theory, appears to have modified more the incentives of the East Europeans and thus had a larger effect on cooperation than did the more anarchic Western alliance. Accordingly, the United States reacted more dramatically to the creation of the informal empire than the Soviet Union did to the Western alliance.

This finding sheds light on the causes of the Cold War. The neotraditionalists now place primary responsibility for the conflict on Stalin himself. Here, I have suggested that there were understandable strategic reasons for the Soviet Union to have pursued an informal empire in Eastern Europe even though this institution helped to precipitate the Cold War. One need not point to Stalin as an individual to assign blame. At the same time, I agree with the neotraditionalists that the Soviet Union, in the end, bears greater responsibility, although both sides remain guilty of many sins of commission and omission. A small change in the type of security institution chosen by the Soviet

Union would have had a comparatively larger effect on the course of history than a similar change by the United States. Had Moscow chosen a less hierarchic security institution, it might well have altered fifty years of history.

Security Institutions and the Persian Gulf War

Following Iraq's invasion of Kuwait on August 1, 1990, the United States created a de facto protectorate over Saudi Arabia. With the kingdom's full acquiescence, the United States sent an initial deployment of 250,000 troops and thereby assumed control of its partner's foreign policy for the duration of the crisis. With the United States eventually deploying more than 3.5 times the number of troops in the region than Saudi Arabia, decisions on the nature and timing of the war, on the purpose and progress of negotiations, and on acceptable outcomes to the crisis were made not in Riyadh, but in Washington.⁵¹

To demonstrate its commitment not to exploit the control it exercised over Saudi Arabia, and to safeguard the kingdom against any tendency—unwitting or not—toward hegemonic opportunism, the United States constructed a multilateral coalition. In the “inner coalition” were two types of states. Some—such as Britain and France—made substantial contributions to the joint effort, and their withdrawal would have raised the physical costs of the war to the United States. Others, most notably Syria, were “fire alarm” states. Wary of American motives, such partners could be counted on to criticize the United States loudly and publicly should it overstep the consensus on policy toward Iraq that held the coalition together. The “outer coalition” of states in the United Nations Security Council provided a further check on U.S. policy. These were, after all, countries that chose not to join in the war but whose acquiescence was necessary to legitimate American actions to the rest of the world. In short, the multilateral coalition was essential to induce Saudi Arabia to enter voluntarily a security relationship that left it highly vulnerable to the whims of American policy.

This particular security arrangement—a protectorate surrounded by a multilateral coalition—was unique in its particulars but based on persistent and connected sets of rules that nonetheless qualify as an institution. The protectorate between the United States and Saudi Arabia had long been anticipated, reflected in detailed planning and the building of Saudi military bases to American military specifications from the early 1980s. Both Saudi Arabia and

51. On the security relationships between the United States, Saudi Arabia, and other coalition members, as well as the factors leading to these institutions, see Lake, *Entangling Relations*, chap. 6.

the United States knew that any major threat to the kingdom's security would necessitate calling in American troops. Similarly, the primary fighting forces of the inner coalition rested on prior coordination and integration. The ability to marshal such a multinational force was made possible by decades of cooperation within NATO and joint military exercises with Egypt and the Gulf states. Despite their unique nature—replicated in many other cases since—and ad hoc quality—put together for the first time as the crisis developed—the protectorate and the coalition rested on rules designed to facilitate precisely this kind of security cooperation.

This institution was advantageous to the United States for three reasons. First, there were large benefits from cooperation with Saudi Arabia and, secondarily, other coalition members. Access to land bases on the Arabian Peninsula was essential to deter further Iraqi aggression and, if Baghdad refused to withdraw, to fight the kind of war planned by the Pentagon. Without bases, the United States would have been limited to air strikes that might have harassed and hurt the Iraqis but that were unlikely to change significantly their behavior—just the scenario played out after the war in the cat-and-mouse game over the UN weapons inspection program. To crush the Iraqi forces in Kuwait, the United States needed land bases from which to deploy its troops.

The members of the inner coalition also shared significant burdens with the United States. Britain and France in particular provided substantial forces that meshed easily with American troops. Still other members provided financial assistance that reduced the direct costs of war to the United States. Overall, the benefits of cooperating with a broad range of states were highly attractive.

Second, there was a substantial risk that left on its own, Saudi Arabia would act in ways that contravened American interests. In particular, the United States worried that the Saudis would compromise with Iraq as a way out of the short-term crisis, perhaps buying off Baghdad with billions from its oil revenues.⁵² Such a defection would have been extremely costly to the United States. If denied the land bases that it required, its whole approach to the crisis would have been undercut. Having declared that the invasion “will not stand,” President George Bush could not easily reverse himself and rely only on what were expected (and later proved) to be ineffective air strikes. Bases on the Arabian Peninsula were thus highly specific assets. Given America's commitment to restore the sovereignty of Kuwait and its military doctrine of using

52. See George Bush and Brent Scowcroft, *A World Transformed* (New York: Vintage Books, 1998), pp. 321, 328; and Bob Woodward, *The Commanders* (New York: Simon and Schuster, 1991), pp. 251–253.

only overwhelming force, there was no alternative to troops stationed on Saudi territory. Worried about Saudi opportunism, the United States needed to exert some significant control over the kingdom's foreign policy.

Third, although they were substantial, the political costs of creating and maintaining the protectorate and coalition were tolerable—largely because the benefits of cooperation were sufficiently high. The primary costs to the United States were the direct expense of inducing others into the coalition, such as the forgiving of Egypt's \$6 billion debt to the U.S. Treasury, and the indirect costs of constraints on its own policy.⁵³ Of the two, the indirect costs were ultimately more important. These constraints manifested themselves in the limited objectives of the war, which soon led to a collective belief that the military success was a "hollow victory";⁵⁴ in the inability to open serious negotiations with Iraq for fear of splitting the coalition; and in the actual conduct of the war, including the inability to target Saddam Hussein personally, the limited incursions into Iraq, and the need to keep Israel out of the conflict.⁵⁵ Given the sensitivity of the Saudis to even the limited degree of control exerted by Washington, a more hierarchic relationship—perhaps entailing the permanent stationing of a large American force on Saudi soil—would have required even greater constraints on the United States to induce their voluntary acquiescence or perhaps a degree of coercion that would have ultimately proven counterproductive. However costly the constraints on American policy turned out to be, a more hierarchic institution would have been significantly more expensive.

THE EFFECTS OF THE PROTECTORATE AND COALITION

The Cold War cases above used two methods to gauge the effects of security institutions: the reactions of third parties to the formation of the institution and counterfactual analyses of the likely effects of the next-best institution under the same strategic circumstances. In the case of the Persian Gulf War, the first method is possibly less accurate, because the outbreak of violence suggests that the two sides—principally the United States and Iraq—could not accurately read each other's preferences or actions in the crisis.⁵⁶ Indeed Saddam

53. On the direct costs of the coalition, see Lake, *Entangling Relations*, pp. 245–246.

54. Jeffrey Record, *Hollow Victory: A Contrary View of the Gulf War* (Washington, D.C.: Brassey's, 1993). See also U.S. News and World Report, *Triumph without Victory: The Unreported History of the Persian Gulf War* (New York: Times Books, 1992).

55. On these constraints on U.S. policy, see Lake, *Entangling Relations*, pp. 247–251.

56. On bargaining failures and war, see James D. Fearon, "Rational Explanations for War," *International Organization*, Vol. 49, No. 3 (Summer 1995), pp. 379–414.

Hussein appears to have fundamentally misread American resolve and, more important, incorrectly estimated the unity of the coalition. Nonetheless we can still tentatively judge the importance of the security institution created by the United States by examining Saddam's reactions, especially his ineffective but still serious efforts to undermine the coalition. That breaking up the coalition was the centerpiece of his foreign policy both before and during the war indicates just how significant the institution was.

Saddam appears to have hoped that his lightning strike into Kuwait would present the world with a *fait accompli*. Consistent with his political strategy of staking out a vanguard position in the Arab world, he apparently expected the Arab League to split into a moderate bloc that opposed him and a radical bloc that supported him, thus preventing the league from mounting a common response. He also predicted that the Gulf states, too weak to resist his move, would prove unwilling to subjugate themselves to the United States. Although he may have anticipated much bluster and possibly sanctions, he appears to have hoped that his attempt to seize Kuwait and either appoint a puppet regime or incorporate the formerly sovereign state as Iraq's nineteenth province would succeed by its very boldness.

Once the magnitude of his miscalculation became apparent—as the Arab states, with the exception of Jordan, condemned the invasion and Saudi Arabia invited the United States to defend the kingdom—Saddam's principal political strategy switched to breaking up the coalition forming against him. He tried to harness anti-Israeli sentiment in the Arab world by calling for a regionwide peace settlement. He subtly manipulated the release of hostages to exacerbate cleavages in the coalition. He encouraged the Soviet Union, then looking to retain its position as a superpower, and France, determined as ever to distinguish itself from the United States, to negotiate a settlement on favorable terms. Even after the fighting started, he attempted to bring Israel into the war through Scud missile attacks—with the hope that Arabs would find it politically impossible to fight on the same side as the Israelis. By dismantling the outer or inner coalitions, Iraq hoped to strip away the safeguards so created, leave Saudi Arabia vulnerable to American opportunism, and thereby crack the protectorate that was the necessary core of an effective response to Iraqi aggression. The United States in turn worried about the stability of the coalition and worked hard to keep it firm throughout the crisis.

This does not solve the mystery, however, of why Iraq actually went to war. Once the fighting appeared imminent, and the vast array of forces against him became clear, Saddam's decision to press on suggests either something

about his preferences, which a decade later remain unclear, or that some other bargaining impediment also existed. That Saddam may have overestimated his ability to break the coalition does not imply that a more accurate reading would necessarily have avoided the war. Likewise, without a fuller understanding of Saddam's motivations and strategies, we must be cautious in inferring too much from his near single-minded focus on disrupting the coalition. Nonetheless, that the coalition formed such an important object of his diplomacy does suggest that it and the protectorate it enabled were vitally important.

Two counterfactuals are relevant to this case. First, the United States could have adopted a wholly unilateral response in which it acted alone to expel Iraq from Kuwait. The logistical difficulties in such an endeavor, however, would have been enormous. To land almost 500,000 troops and their equipment in heavily defended Kuwait City would have been nearly impossible. The risks entailed would have precluded a land-based assault on Iraqi forces, thereby restricting the United States to air strikes and diplomacy.⁵⁷

Second, the United States could have created a less hierarchic security institution, such as an alliance, in which Saudi Arabia was not just a *de jure* but also a *de facto* equal. The primary implication of this strategy would have been to limit severely the number of American troops in Saudi Arabia. The huge force actually sent by the United States inevitably gave it the ability to dominate policymaking. To limit its influence would have required limiting its presence. This would have had two secondary implications. The coalition forces would have been smaller or composed of more diverse troops from more countries; either would have degraded the fighting power of the coalition and left the outcome of the war less certain. The coalition would also have been less solid politically. Unable to control the policy of its partners, the United States would have been more suspicious of their actions and more cautious in its own strategy. In turn, Saudi Arabia, which was greatly reassured by Bush's initial offer of 250,000 troops, would have been more worried about the depth of the U.S. commitment and thus more willing to consider settling the conflict on terms possibly favorable to Iraq. That the next most likely alternatives would have produced such different outcomes indicates, once again, how important the security institution created by the United States was to the success of cooperation.

57. Lake, *Entangling Relations*, p. 222.

Conclusion

By expanding the range of institutions typically examined in international relations, we can see their effects more clearly. States form institutions because important interests are at stake. Institutions help realize the benefits of cooperation because they constrain, to a greater or lesser extent, the behaviors of their members. In turn, the greater the risks of cooperation, the more likely states are to act unilaterally or build more hierarchic institutions. Although international relations scholars generally ignore hierarchies within the world system, the Soviet Union's informal empire in Eastern Europe during the Cold War and the United States' protectorate over Saudi Arabia in the Persian Gulf War suggest that such institutions are by no means uncommon. The methods employed here do not produce easily replicated measures of the effects of institutions. Nonetheless, examining the reactions of third parties and exploring plausible counterfactuals allow us to estimate the impact of institutions and draw conclusions about their effects. The evidence strongly indicates that security institutions matter—they alter behavior and allow cooperation where it would otherwise almost certainly fail.

The arguments developed above offer mixed support for both neorealists and neoliberal institutionalists who have staked out conflicting positions on the extent to which institutions are important in international relations. As predicted by the theory and observed in the cases, relatively anarchic institutions, such as NATO, exert relatively smaller constraints on members than do hierarchic institutions, such as the Soviet Union's informal empire in Eastern Europe.⁵⁸ Anarchic institutions are unlikely to mitigate wholly the more nefarious consequences of systemic anarchy such as the need for self-help or the security dilemma. At the same time, it would be wrong to conclude that such institutions do not matter in an absolute sense and in some important ways.⁵⁹ The Soviet Union did react to the creation of NATO, and the nascent security

58. That hierarchic institutions appear to exert greater constraints on behavior than do anarchic institutions is not inconsistent with neoliberal institutionalism or neorealism, both of which accept that domestic politics is distinct from international politics precisely because of differences in the type and density of institutions in the two arenas. Where this analysis does differ from these schools is in the possibility and certainly the importance of hierarchies within international relations.

59. Of course, institutions may also solve coordination problems where the interests of the actors are more congruent. In these cases, anarchic institutions are likely to be sufficient to foster cooperation. Because hierarchic institutions are hard to create in the international system, and thus states fail to cooperate in the face of potential gains, anarchic institutions may facilitate more overall cooperation in international affairs than do hierarchic institutions.

institution clearly worsened superpower relations. By implication, then, the institution most likely did serve to facilitate cooperation among its member states and led to a stronger defense than otherwise. Even anarchic institutions can be important in international relations.

Conversely, neorealists argue that international institutions exert no impact independent of the interests of their members. The reactions of third parties and plausible counterfactuals suggest that this argument is incorrect. Institutions do matter, and they are likely to matter in security affairs precisely because important national goals are at risk. This flips the conventional wisdom on its head. Where some neorealists admit that institutions may matter for economic relations between states, they presume that security affairs are different and more immune from institutional effects. The evidence here suggests that when we observe security institutions in place, they may well be more important than their economic counterparts.

Neorealists might still respond that the effects attributed to institutions in this article derive mainly from power and dominance. This is also incorrect. First, the difference in security institutions created by the United States and the Soviet Union after 1945 cannot be explained by international structure, which would predict uniformity in the policies of the two superpowers. Nor can the difference be explained by the threat each posed to the other, for again this would predict a symmetrical response. Bipolarity cannot account for the choice of institutions, and therefore cannot explain their differing effects. Second, under bipolarity states are expected to balance internally, to rely on their own capabilities. Although the contributions of associated states may be helpful, according to Waltz, "they are not indispensable" and realignments of alliances are "fairly insignificant."⁶⁰ If so, then the United States should not have reacted strongly to the Soviet Union's creation of an informal empire in Eastern Europe, which was clearly not the case. The institution itself appears to have substantially augmented the internal capabilities possessed by the Soviet Union. Finally, the multilateral coalition created by the United States during the Persian Gulf War cannot be explained simply in terms of power. According to neorealists, states should be unwilling to bind their own hands and, even if they do so, the institutional constraints should not be credible and should therefore have no effect. But it was because the United States was so powerful relative to the Gulf states that it needed to construct a security institution to constrain its power. That the institution enabled cooperation that would not otherwise have occurred indicates that these constraints were relatively effective.

60. Waltz, *Theory of International Politics*, pp. 168–169.

tive. Power is important in permitting dominant states to define relationships with subordinates that meet their needs, but power explains neither the form of security institutions nor their effects.

More generally, the analysis shows that the international system is not wholly anarchic. If we look below relations between the major powers, the institutions that govern interactions between states are far more variegated than either neorealists or neoliberal institutionalists suggest. Institutionally, world politics is a far more complex realm than we often recognize.

INSTITUTIONS IN THE TWENTY-FIRST CENTURY

The United States today enjoys a position of unprecedented power and influence in the world. For good or ill, intended or not, America's actions have strong effects on others. To date, the United States has used security institutions effectively to shield others from the potential for opportunism inherent in its current hegemony. The multilateral coalition created in the Persian Gulf War has been repeated in Somalia, Haiti, Bosnia, and Kosovo—in short, every time the United States has used significant military force since the end of the Cold War. And it has been employed for the same ends: to control opportunism by its partners and in turn to bind itself to working within the international consensus on the appropriate ends and means of foreign policy. In bipolarity, the competition with the Soviet Union constrained American capriciousness. Today the United States binds itself through institutions that limit its ability to exploit others.

Domestically, a bitter struggle is emerging over the direction of American policy. Unilateralists are skeptical about international institutions. Often reflecting a neorealist view of international relations, they worry that the United States might be placing too much faith in its ability to control the actions of others through international organizations and coalitions. At the same time, and apparently unaware of the contradiction, they bristle at the constraints being imposed on the United States by these same institutions. They call for maintaining America's freedom of action abroad and oppose organizations, such as the UN, that threaten to limit the country's sovereignty. In an era of unrivaled power, constraints on the use of that power are hard to abide.

Multilateralists have greater confidence in the efficacy of international organizations and coalitions. They believe that such institutions are useful instruments in shaping the behavior of others. Multilateralists also accept the need for the United States to work within the international consensus and are comfortable using international institutions as checks on American opportunism. For them, the benefits of cooperation for the United States and the

ability to constrain the behavior of others are typically seen as greater than the costs.

Unilateralists fail to appreciate that constraints on its policy are the price the United States must pay for the right to lead. Even in a world of unparalleled American power, the price of controlling others is the need to constrain oneself as well. If the United States were to throw off the fetters imposed by the international institutions it has, so far, willingly created and forcefully pursue its self-defined goals, it might well drive others into opposition. For the United States to insist that international institutions serve only its interests undercuts their legitimacy and effectiveness in regulating international affairs and undermines U.S. policy in the long run.