Ref # 2105

# Allies, Adversaries, and International Trade

*Joanne Gowa*

PRINCETON UNIVERSITY PRESS

PRINCETON, NEW JERSEY

1994

examine this barrier to trade that is its most profound flaw. Because of this and other flaws, hegemonic stability theory should be deposed from its position of prominence. Only a system-level theory that successfully redresses its most serious flaw can administer the coup de grâce.

In the next chapter, I begin to construct a Third-Image theory that takes explicitly into account the impact of the play of great-power politics on free trade.



CHAPTER THREE

# Allies, Adversaries, and Free Trade

THIS CHAPTER CONTAINS the analytic core of the book. In it, I develop at length my argument about the impact of the anarchic structure of international politics on the exchange of goods and services among states. I cast my argument at the macroeconomic rather than microeconomic level. In other words, the argument advanced here is *not* about, for example, attempts to use economic statecraft to embargo exports of particular products or to inhibit the development of technologically advanced industries in other countries. It is instead an argument based on the effect of free trade on the real income and power potential of states.

I argue that strategic interaction among the great powers influences trade among them, because of the security externalities that trade produces. I consider these external effects explicitly. In order to do so, I construct and analyze a very simple game-theoretic model. The analysis makes clear that tariff games between allies differ systematically from those played between actual or potential adversaries.[1] These differences imply that free trade is more likely within than across political-military alliances. The analysis also suggests that alliances are more likely to evolve into free-trade coalitions if they are embedded in bipolar rather than multipolar systems.

The explanatory power of these hypotheses obviously depends directly on the incidence of alliances across nations and time. Less obviously, perhaps, it also depends on the attributes of the process of alliance-formation. In my analysis, I assume that rational-choice theory can explain intra-alliance trade patterns. This assumption makes sense only if closely related decisions—that is, decisions about alliance entry—can also be understood as the outcome of a rational-choice process.

As a result, in the first section of this chapter, I examine the extent

[1] Tariffs are used here as a substitute for all forms of government intervention in trade. Because of constraints imposed by GATT, intervention now usually takes the form of nontariff barriers to trade (NTBs). As long as the home country captures the scarcity rents that accrue from trade barriers, the form in which intervention occurs does not affect the argument presented here.

to which the anarchic structure of the international system induces rational states to ally with each other. Then I develop in detail my argument about the impact of great-power politics on inter-state trade.

## A Rational-Choice Theory of Alliance Formation

I define an alliance in conventional terms, that is, as "a formal agreement between two or more nations to collaborate on national security issues" (Holsti, Hopmann, and Sullivan 1973, 4).[2] At least three different types of alliances exist: (1) a defense pact, which commits its signatories "to intervene militarily on behalf of one another"; (2) a neutrality or nonaggression pact, which obligates the parties to it to remain neutral in the event of war; and (3) an entente, which mandates consultation and/or cooperation if a war occurs (Small and Singer 1969).

The frequency with which great-power alliances have formed over time and across nations suggests that there is nothing puzzling about these alliances. They seem to be an inevitable concomitant of life in an anarchic world. Absent any supranational authority, states exist in a self-help system. In such a system, unless a state can ensure its security unilaterally, it will ally with others in order to do so. The net effect is to maintain the balance of power in the international system as a whole. It is this logic that led Hans J. Morgenthau, for example, to assert that the "most important manifestation of the balance of power" inheres "in the relations between one nation or alliance of nations and another alliance" ([1948] 1973, 181).

Recent literature tends to refine, rather than challenge, the claim about the relationship between system structure and alliance-formation that Morgenthau and others advanced long ago. Stephen M. Walt, for example, argues that it is the balance of threat rather than the balance of power that motivates states to ally with each other (1987).[3] Walt (1987, 1989); Thomas J. Christensen and Jack Snyder (1990); and Conybeare

[2] Bruce Russett defines an alliance in very similar terms: that is, as "*a formal agreement among a limited number of countries concerning the conditions under which they will or will not employ military force*") (1971, 262–63, emphasis in original). Glenn Snyder does so as well. Alliances, he observes, "are formal associations of states for the use (or nonuse) of military force, intended for either the security or the aggrandizement of their members, against specific other nations" (1990, 104). Walt, however, contends that alliances can also be based on tacit rather than on explicit agreements (1987, 12).

[3] For incisive critiques of Walt's argument, see Keohane (1988) and G. Snyder (1991).



(1992), among others, have examined the pre- and post-entry strategy sets available to states, while James D. Morrow (1991) has examined the objectives of alliance entry.[4] Although each of these studies is important in its own right, none of them supplants the claim that alliances are a consequence of a balance-of-power world.

The same statement can be made about attempts to analyze alliance-formation more rigorously than was typical of early literature in the field (e.g., Liska 1962; Holsti, Hopman, and Sullivan 1973). Glenn Snyder, for example, models the process as a PD game.[5] He argues that PD-preference orderings characterize states confronting decisions about whether to seek allies, because "(1) some states may not be satisfied with only moderate security, and they can increase it substantially by allying if others abstain [i.e., $T > R$]; [and] (2) some states, fearing that others will not abstain, will ally in order to avoid isolation or to preclude the partner from allying against them [i.e., $P > S$]" (1984, 462).

Because alliance entry is a dominant strategy, Snyder continues, alliances inevitably populate international systems. This outcome is collectively suboptimal, however; "if all states are about equally strong and are interested only in security, all are fairly well off if all abstain, since each has moderate security against individual others," while alliances "involve various costs, such as reduced freedom of action, commitments to defend the interests of others, and so forth" (Snyder 1984, 462).

Thus, although Snyder formalizes conventional arguments about alliance formation, he does not challenge their underlying logic. As is true of other recent contributors to the literature on alliances, he tends to accept the long-standing argument that alliances are the product of anarchic systems and of concerns about the balance of power that such systems induce.

The process of alliance-formation, however, *is* intrinsically problem-

[4] A variety of candidate strategies exist. Walt, for example, argues that states seeking allies can either balance or bandwagon; that is, they can choose to ally with or against the threatening state (1987; 1989). Conybeare argues that one determinant of pre-alliance entry strategies is an attempt to maximize a "combination of risk and return" (1992). The relevant strategy set available to states that have already entered into an alliance, suggest Christensen and Snyder, includes "chain-ganging" and "buck-passing" (that is, offering an ally an irrevocable rather than an easily reversible commitment) (1990).

[5] Niou, Ordeshook, and Rose (1989) formalize the process of alliance-formation, but they rely on cooperative game theory to do so. As a result, as G. Snyder argues, the empirical relevance of their argument is limited (1991, 132).

atic. Indeed, within the context of balance-of-power theory, alliance-formation is paradoxical. Balance-of-power theory implies, as Arthur A. Stein observes, that states concerned with their own survival will

> act in concert to prevent the emergence of a power that threatens them. This coincidence of interests forms the basis for alliances, which in turn undergird a balance of power. Hence these alliances hardly seem necessary. If, on the other hand, alliances entail commitments that states have no interest in fulfilling, then nations will not keep to their terms, and the accords will have no consequences. Alliances, then, must be either unnecessary or inherently unbelievable bluffs.[6] (1990, 153)

The logic of Stein's argument is compelling. Two interests presumably motivate a state to enter into an alliance.[7] One interest is to increase a state's ability to deter war.[8] Another is to enhance its ability to prevail if a war occurs. If alliance commitments are problematic, however, alliances will neither enhance deterrence nor increase the probability of prevailing in the event of war.

Alliance commitments are almost always problematic. It is very difficult for any state to credibly commit to enter a future war in which its ally is a belligerent. Indeed, if only an imperfect congruence of interests between allies exists, the temptation to renege in the event of war may prove irresistible.[9] If such postcontractual opportunism is plausible and its existence is common knowledge, alliances will not

[6] Stein is not the first or only scholar to have made this point. Morgenthau, for example, noted that when the interests of states are such that they "obviously call for concerted policies and actions . . . [then] an explicit formulation of these interests, policies, and actions in the form of a treaty of alliance appears to be redundant" (1948 [1973], 182). For another example, see Morrow (1991, 906).

[7] Paul Schroeder (1976) has argued that states also enter alliances in order to control the behavior of their allies. This interest is not distinct from those discussed in the text, however; presumably the ultimate objective of such control is to win or deter a war.

[8] As Lalman and Newman have pointed out, however, an alliance can also have an unintended, opposite effect. Because a state's entry into an alliance reveals "which nations are friendly enough for alliance partners, a nation also reveals which nations are not as close. This new information may prove risky to a nation's efforts to deter threats" (1991, 242).

[9] This is precisely the problem that plagued NATO during the Cold War. Despite the declared commitment of the United States to its European allies, it was clear to them that their interests and those of the United States did not coincide completely. For example, that the United States would elect to risk nuclear war with the U.S.S.R. in the event of a Soviet attack on Western Europe always seemed less than certain to the European allies of the United States.





have significant effects either on deterrence or on the probability of victory.[10] Under these circumstances, no state will have a strong incentive to ally with any other.

That alliances are neither uncommon nor omnipresent in the real world suggests, however, that they are neither inevitable *nor* worthless by-products of anarchic international systems. Even if the assumption that the alliance entry game is a PD is maintained, Glenn Snyder's conclusion does not necessarily follow. There is nothing inevitable about the outcome of an infinitely repeated PD game or of a finite game in which incomplete information exists. Under either condition, strategic interaction between or among states engaged in such a game *can* produce an alliance-free international system.[11]

If the assumption about PD preferences is not maintained, it is much more difficult to sustain the claim that alliance-formation is inevitable. State preferences with respect to alliances, for example, might conform to those that define the game of Harmony. As figure 3.1 makes clear, alliance abstention (cooperation) is a dominant strategy in this game. Each state is better off if it does *not* ally, whether or not another state chooses to do so.

There are several situations that might induce this preference ranking. Among them is one in which all great powers are satisfied with the international status quo and their preferences are common knowledge. The Concert of Europe seems to be an example, even if relatively brief, of such a situation.[12] As suggested by recent developments within the former Soviet Union and between it and the United States, a dominant

[10] The empirical evidence is mixed about the extent to which states honor their commitments to their allies in the event of war. Siverson and King argue, for example, that although "many alliance partners joined in war participation (67 or 23.1 percent) [between 1815 and 1965], many more (223 or 76.9 percent) did not" (1980, 2). Bruce Bueno de Mesquita, however, argues that, of nations which were attacked, "76 percent of the allied nations received fighting support from *some* of their allies, while only 17 percent of the nonallied states found anyone fighting alongside them" (1981, 113, emphasis added).

Note that Bueno de Mesquita's calculation addresses only the issue of whether a belligerent received support from *some* of its allies. It does not address what seems to be the more relevant question; that is, the percentage of a belligerent's allies that honored their commitment to enter the relevant war.

[11] It is important to note that iteration does not ensure that the Pareto-superior equilibrium outcome will emerge. The only equilibrium outcome of the stage game (that is, PP) is also an equilibrium outcome of the repeated game. The effect of iteration on PD games is explained in more detail later in this chapter.

[12] For an incisive analysis of the operation of the Concert, see Jervis (1985).

| | | Column: Cooperate | Column: Defect |
|---|---|---|---|
| Row | Cooperate | 4,4 | 2,3 |
| Row | Defect | 3,2 | 1,1 |

FIGURE 3.1. Harmony

strategy equilibrium of alliance abstention might also emerge as a consequence of the disintegration of one or more great powers.[13]

Thus, the argument of G. Snyder and others that alliance-formation is inevitable is not persuasive. Neither is the argument that, because alliances are necessarily "cheap talk," alliance entry is irrational. If, for example, information is incomplete, and if costly signaling accompanies alliance formation, alliances may not be inherently incredible bluffs.[14] Stein himself is sensitive to this possibility. Even if alliances do not actually influence the behavior of the allied states, he observes, they "may still fulfill a useful role as devices to signal to third parties intentions about contingent future behavior that might otherwise not be presumptive" (1990, 153n7).

Reputational effects, for example, can induce a state to enter a war in which its ally is a belligerent even if that state expects to incur short-run costs in doing so *and* the outcome of the war will not directly affect its interests. Although a state might be better off in the short run if it reneges on its commitment to its ally, it may be better off in the long run if it does not. If it reneges, a state may acquire a reputation as an unreliable ally.[15] Thus, the long-run reputational costs which a state

[13] States can also abjure alliances if their preferences conform to those that define an Assurance Game; that is, $R > T > P > S$.

[14] "To reveal private information where there are incentives to misrepresent it," as James Fearon observes, "signals must be costly in a quite specific way. For example, in crisis bargaining an action that is equally costly to take whether the state has high or low resolve should not lead a rational adversary to change its prior belief about the state's true willingness to use force." Rather, he adds, "only actions that are potentially *more* costly for 'low resolve' than for 'high resolve' states will convey willingness to use force on the particular issues in dispute" (1992, 38; emphasis in original).

[15] In a footnote, Stein does refer to the possibility that a state might honor its commitment to an ally if it is concerned about its reputation (1990, 164n44). Conybeare also notes that unreliable allies "would suffer reputational and possibly other sanctions, leaving them dangerously isolated" (1992, 65n8).





incurs if it reneges may exceed those of entering the war. Treaties, as Charles Lipson observes, put "reputation at stake." As a result, he notes, they "add to the costs of breaking agreements, or rather, they do so if a signatory values reputation" (1991, 533).

The existence of incomplete information about either state types or interests, however, is not necessary to explain alliance formation in a balance-of-power world. Rational states may ally with each other even if their interests are wholly congruent *and* this congruence is common knowledge. They might do so, for example, if the formation of an alliance facilitates either the exploitation of economies of scale with respect to defense and deterrence or the ex ante coordination of strategies that is essential to maximize the probability of victory in the event of war.[16]

The North Atlantic Treaty Organization seems to be an example of such a case. NATO has enabled its members to exploit the scale economies that inhere in nuclear weapons, as well as coordinate their wartime strategies ex ante.[17] A less obvious example is the Anglo-French Entente. The pre–World War I accord between Britain and France allowed its members to coordinate their forces ex ante insofar as the accord entailed an implicit division of labor with respect to naval power in the event of war with Germany.

The formation of alliances is, in short, more problematic than much of the international relations literature suggests. Decisions to ally are not inevitable, even if state preferences conform to those that characterize a PD game. Nor are alliances necessarily a credible deterrent to war. Nonetheless, a rational state may adopt a strategy of alliance entry for any of several reasons: to allow it to signal its interests to other states in situations of incomplete information; to exploit economies of scale with respect to deterrence and/or defense; or to effect the ex ante coordination of forces that might influence the outcome of any war that does occur.[18]

[16] The presence or absence of potential welfare gains from explicit coordination may answer one question Keohane raises in his review of Walt's book: "Under what conditions do alliances become formalized?" (1988, 175).

[17] This is obviously not meant to imply that the members of NATO have identical stakes in all situations, nor that economies of scale exist or have been exploited with respect to all the military forces of NATO's members.

[18] Altfeld (1984) argues for and provides an empirical test of the hypothesis that ra-

Decisions to ally, therefore, *can* be understood as the product of strategic interaction among rational states. This suggests that it makes sense to assess the impact of alliances on trade within the context of an analytic framework based on the same assumption: that is, that states act rationally. The next section of this chapter focuses explicitly on the question of the conditions under which rational states will trade more freely within than across alliances.

## Power, Trade, and Tariffs

The play of great-power politics affects the incentives of states to trade freely with each other because of the security externalities associated with trade. These external effects arise because the source of gains from trade is the increased efficiency with which domestic resources can be employed. The resulting increase in real income frees economic resources for military uses. Thus, trade increases the potential military power of any country that engages in it (D. Baldwin 1985, 216; Hirschman [1945] 1980, 14; McKeown 1982; Root 1984, 75; Srinivasan 1987, 352).

The anarchic structure of the international system implies, in turn, that a state cannot be indifferent to the potential power of either its allies or its adversaries. Inherent in the structure of the international system is the ability of any state to use or threaten to use military force to achieve its goals. Whether any state will do so depends in part on its power. The latter depends in part on its real income. As a result, any increase in the real income of its adversaries decreases the security of a state. Conversely, increases in the income of its allies increase its own security.[19]

Thus, the real income gains that motivate free trade are the source of the externalities that can either impede or facilitate it: Trade with

tional states will join alliances if, for example, the opportunity costs of doing so are lower than are those involved in a unilateral increase in military spending. Morrow (1991) argues and finds empirical support for the proposition that states trade off security gains against autonomy losses when they ally.

[19] As Srinivasan points out, whether a state will actually use its gains from trade to increase its defense budget depends on its social-welfare function (1987, 356–57). This does not affect the analysis here because, if conditions change, increased GNP will allow a state to increase its military power more easily than it would otherwise have been able to do.





an adversary produces a security diseconomy; trade with an ally produces a positive externality. In either case, trade creates a divergence between private and social costs: The costs of trade to the nation differ from those that accrue to the individuals involved.[20] This implies that government intervention in trade can increase the welfare of the nation as a whole.

### *Government Intervention in Trade: "Acupuncture with a Fork?"*

The argument I advance here joins a relatively large set of arguments for government intervention in trade that are based on national security. This argument is *not,* however, vulnerable to the logical flaw that afflicts many members of this set. In most cases, careful analysis leads ineluctably to the conclusion that intervention in trade is only rarely, if ever, the optimal response to situations in which trade adversely affects national security.

The situations can be diverse. For example, they can involve the impact of free trade on the industrial structure of a nation, on the composition of its labor force, or on its pattern of consumption. The diversity of situations does not change the well-established conclusion that intervention in trade imposes higher costs on national welfare than do other policy alternatives.

The logic of the argument that leads to this conclusion is straightforward. Suppose, for example, that open international markets result in a pattern of domestic production that is inimical to national defense. More specifically, suppose that free trade leads to a socially suboptimal level of production in certain industries. A market failure results: Private and social costs of production diverge. As a consequence, government intervention can correct this distortion and thereby increase national welfare.

Intervention in trade, however, is not the most efficient response to a market failure of this kind. A tariff can raise output in the affected industries to the socially optimal level. Because it does so by increasing domestic prices, the tariff simultaneously distorts consumption. A production subsidy is a more efficient response. It raises output to the optimal level without influencing prices. An analogous argument would apply if, for example, consumption were either above or below the socially optimal level. A consumption tax or subsidy would correct the

[20] See Chapter One, note 9, for an explanation of why externalities create a divergence between private and social returns.

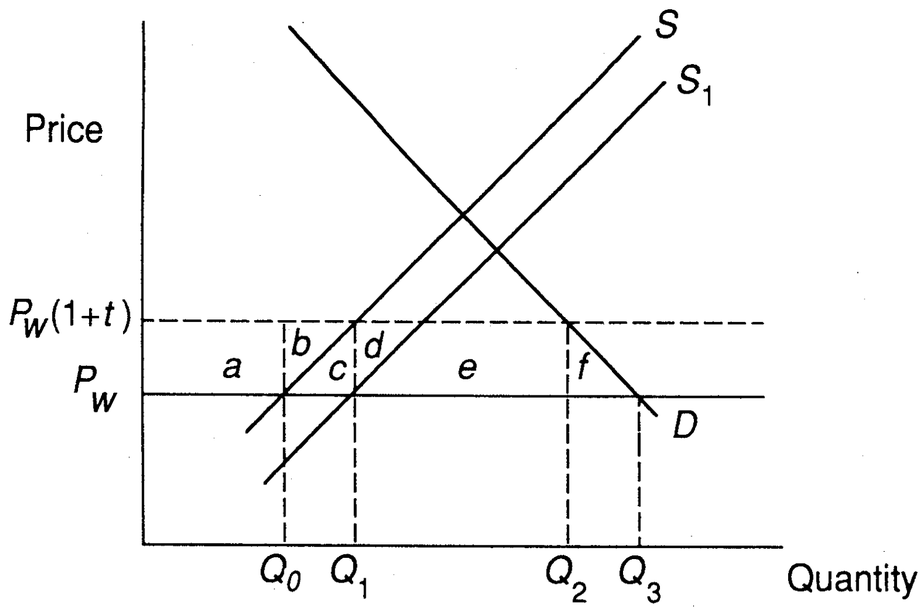


FIGURE 3.2. A Subsidy vs. a Tariff
*Source:* Deardorff and Stern 1987, 35.

market failure without the adverse effect on production that would result from a tariff.

Figure 3.2 illustrates this argument. The figure depicts the domestic market for a good that can also be imported at the world price, $P_w$. Free trade, however, leads to a socially suboptimal level of domestic output, $Q_0$, in a particular industry. $S$ represents the private marginal cost, and $S_1$ the social marginal cost of producing the good. The social benefit that would accrue from an increase in output is measured by the area between the two curves. In addition to its usual effects, therefore, a tariff that increases domestic output from $Q_0$ to $Q_1$ also generates social benefits equal to the area of $b + c$. If $b$ is greater than $f$, a tariff increases welfare.

A production subsidy, however, achieves the same result at less cost. The subsidy costs the government $a + b + c$. Part of this subsidy, $a + b$, is a transfer to producers; but part of it is a social gain, $b + c$. The net gain, therefore, is equal to $b$. Because $b > b - f$, a production subsidy is welfare-superior to a tariff. The general principle, as Alan V. Deardorff and Robert M. Stern observe,

> is that trade intervention, by introducing two distortions rather than one, may succeed in solving one problem but only at the same time that it causes another. Trade policy is like doing acupuncture with a fork: no matter how carefully you insert one prong, the other is likely to do damage. (1987, 39)





Thus, in most cases in which free trade adversely affects national security, the first-best policy is not trade intervention but a policy that deals with the divergence directly (Corden 1986, 86)—production subsidies if domestic production is judged inadequate, consumption taxes or subsidies if consumption is above or below socially optimal levels, wage subsidies if market forces lead to socially suboptimal levels of employment for workers who possess skills perceived as essential for national security.

A first-best case for tariffs as a response to market failure remains, however, where the distortion originates in trade itself. An embargo is perhaps the most obvious case. If the probability of a trade embargo is an increasing function of the volume of imports *and* this relationship is not recognized by private traders, a trade distortion exists: The social cost of imports is higher than is their private cost. In this situation, the optimal policy choice can be a tariff (Bhagwati and Srinivasan, 1976).[21]

Even a credible embargo threat requires government intervention if and only if private markets do not anticipate and respond to the threat. However, strong incentives to anticipate an embargo and to react appropriately confront individual traders in the market concerned. They can, for example, use "furnaces that can switch from oil to gas or vice-versa, cultivate a more diverse clientele for exports, maintain other sources of supply (including stockpiles), and so forth" (Dixit 1987, 267). As a consequence, even a credible embargo threat does not always create a role for preemptive action by the government.[22]

In sum, it seems difficult to defend successfully any claim that intervention in trade is the first-best response to situations in which free trade adversely affects national security. Whether the market failure involves an inadequate level of domestic production, insufficient or excessive domestic consumption, or an inadequate supply of workers with particular skills, avoiding "acupuncture with a fork" implies correcting the distortion at the source rather than intervening in trade. With the possible exception of an embargo that does not prompt action

[21] For a dissenting view about the relationship between import volume and the probability of an embargo, see Helpman (1987, 372).

[22] When it does, in addition, several policy alternatives exist that are less costly than import restrictions (for example, supplying information to private traders about prospective embargoes, accumulating stockpiles, imposing consumption taxes, and providing production subsidies) (Dixit 1987, 265).

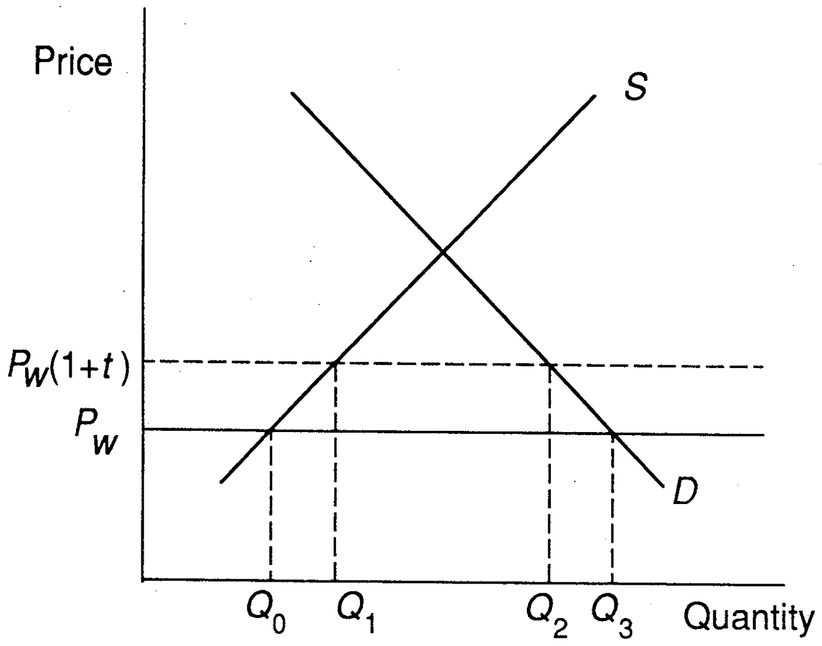


FIGURE 3.3. Trade with an Adversary

by private markets, trade intervention can be justified only as a second-best response.

This conclusion does *not* apply to the argument of this book, that trade between adversaries creates security diseconomies. In this case, the market failure does not originate in a domestic distortion; it originates in trade itself. In addition, unlike an embargo, this situation does not create strong incentives for private markets to respond appropriately. An individual trader will have an incentive to take these social costs into account only in the unlikely event that his trade is very large relative to the nation's as a whole. As a consequence, only government intervention can correct the divergence between private and social costs that arises as a consequence of the security externalities associated with trade. Moreover, because the marginal divergence arises as a direct consequence of trade, intervention in trade is a first-best policy. Figure 3.3 illustrates this situation.

The figure shows the domestic market for a good that can also be imported at its world price, $P_w$, which is the private marginal cost of imports. In the absence of government intervention, domestic output is $Q_0$; domestic demand is $Q_3$; and the amount imported is the difference between them ($Q_3 - Q_0$). Because of the external diseconomy that trade with actual or potential adversaries generates, the marginal social cost of imports is higher, shown as $P_w (1 + t)$. This market failure can





be eliminated by a tariff, $t$, which increases domestic prices to $P_w (1 + t)$; raises domestic output to $Q_1$, lowers domestic demand to $Q_2$, and reduces imports to $Q_2 - Q_1$. In this case, "acupuncture with a fork" *is* the treatment of choice. Restricting trade both by raising the price of imports and by expanding domestic production is optimal.

### *Optimal Tariff Games*

Successful intervention in trade, to correct the security externalities associated with it, depends on the ability of a country to affect the real income of the state which is its target. Although any tariff will distort resource allocation in the target country and therefore decrease its real income, it will do so without imposing net costs on the home country only if the latter can affect its terms of trade (i.e., the relative price of its exports on world markets).

One of the principal lessons of standard trade theory is that the ability of a state to influence its terms of trade depends on its size in global markets. Standard theory also makes clear, it should be noted, that the relevant size is not that "of the country as a whole but rather its share of world trade in markets in which it exports and imports" (Deardorff and Stern 1987, 37). If a country is large enough to exercise market power, then, it can use a tariff to increase its own real income at its adversary's expense.[23] In doing so, it will narrow the gap between the private and social costs of trade. Thus, a tariff can be welfare-superior to a policy of free trade for a state in an anarchic international political system.

Under the assumptions of standard international trade theory, an optimal tariff is the *only* alternative available to a country that seeks to reduce the income of its trading partner without incurring a net reduction in its own.[24] However, strategic trade theory, developed recently, suggests that other instruments of trade policy can achieve the same end.[25] It does so because, unlike its predecessor, strategic trade theory does not assume that perfect competition prevails in all markets

[23] In order to do so without imposing costs on its allies, a country must be able to use discriminatory trade policies. This is not easy to accomplish. For detailed analyses, see Willett and Jalalighajar (1983–84); Bayard, Pelzman, and Lopez (1983); and Kaempfer and Lowenberg (1992).

[24] For a complete explanation of the assumptions that standard trade theory relies upon, see Yarbrough and Yarbrough (1991, 280–81).

[25] The seminal article in this area is Brander and Spencer (1984).

Instead, strategic trade theory assumes that some industries are more valuable than are others to the nation as a whole. If such industries exist, government intervention in trade can increase national welfare. As in the case of standard trade theory, such intervention will decrease the real income of the target country.

The logic of strategic trade theory is straightforward. The basic assumption is that some industries generate excess profits or rents—that is, "a receipt in excess of the opportunity cost of a resource" (Ekelund and Tollison 1981, 18).[26] A nation can therefore increase its real income to the extent that it can capture a larger share of these profits. Under carefully specified conditions, government intervention in the form of an export subsidy will enable it to do so.[27] The subsidy allows the home industry to credibly commit to producing a larger output than it would otherwise, thereby forcing its foreign rival to contract its output.[28] Because the increase in home industry profits that results exceeds the amount of the subsidy, national welfare rises.[29]

Government intervention in trade can also be welfare-enhancing in the case of industries that generate externalities.[30] When an industry

[26] These returns "can be thought of as scarcity values of the restricted positions or slots in the industry. . . . [Thus,] economists call them monopoly rents" (Dixit 1986, 290).

[27] For example, although the existence of excess profits is a prerequisite for a welfare-enhancing export subsidy, it is difficult to establish whether industries are actually generating rents. Concentration among firms in an industry is not sufficient by itself to imply that excess profits exist, because competition may be fierce even if the number of firms is small (Dixit 1986, 292). In addition, supranormal profits at the end stage of competition may follow subnormal profits at the initial period, thus making average returns normal (Dixit, cited in Richardson 1987, 287).

Moreover, the existence of excess profits is necessary but not sufficient to ensure that an export subsidy will increase national welfare. Among other necessary conditions are: (1) the industry itself cannot capture these profits (Dixit 1986, 294); (2) the competition between the targeted industry and its rival is Cournot rather than Bertrand (that is, the strategic variable is output rather than price); and (3) the number of domestic firms in the relevant industry is small. If price competition exists or the number of domestic firms is large, an export subsidy will decrease rather than increase national welfare: An export tax is indicated instead (Eaton and Grossman 1983).

[28] More precisely, the subsidy empowers the home industry to act as a Stackelberg leader. As such, the home industry can determine where on its reaction curve its rival will operate.

[29] The impact on consumer surplus is not at issue here. The Brander–Spencer model assumes that industry output is not consumed at home. As a result, the effect on consumer surplus does not enter the welfare calculations of the subsidizing country.

[30] Again, intervention on the basis of externalities will enhance welfare only if certain conditions are met. For example, externalities must be country-specific. Externalities, however, often cross borders (for example, via the diffusion of knowledge or trade intermediate goods) (Krugman 1987, 231). Indeed, if an externality is international, optimal policy is to free-ride on the subsidies of *other* nations (Dixit 1987, 266).

In addition, external economies are unpriced. As a result, Krugman contends that governments can identify industries that are potential targets of welfare-increasing intervention only by combining "detailed knowledge of the industry with a heavy reliance on guesswork" (1986, 17). Others are somewhat more sanguine. Richardson argues that "indicators" of the value of externalities can be observed, such as "the renewal pattern of license fees for patents, the different valuations of a firm by the stock market and a takeover suitor, and the subtly conditioned covariation in productivity growth across different economic activities" (1990, 119).

[31] Chapter Two provides a more detailed explanation of the optimal tariff argument.

[32] See, for example, Deardorff and Stern (1987, 34).



generates positive externalities, its output is socially suboptimal. Although its operation does not necessarily produce supranormal profits in the form of direct returns to factors employed in the industry, "capital and labor in the sector . . . will yield high returns to society because in addition to their own earnings they provide benefits to capital and labor employed elsewhere" (Krugman 1986, 13). One source of externalities, for example, inheres in the incomplete appropriability of new knowledge: Frequently, it is difficult to prevent others from benefiting from the creation of knowledge; as a result, private markets tend to underinvest in it.

For several reasons, I base the analysis in this book upon standard rather than strategic, trade theory. First, it is standard trade theory that underlies much of the analytic literature in the subfield of international political economy. Second, strategic trade theory was not developed until after 1980, although it is, of course, conceivable that policymakers had an intuitive appreciation of strategic trade theory that preceded its formalization. Third, the advent of strategic trade theory does not affect the logic of the argument I advance here. Its only effect is to expand the range of policy instruments to which it can be applied. For these reasons and in order to simplify the presentation, I base the following analysis on standard trade theory.

Thus, I argue that countries can use their ability to affect their terms of trade to correct the market failure produced by trade in an anarchic system. This is, of course, a variant of the traditional optimal tariff argument.[31] As such, it might elicit the same skepticism that greets this and other suggestions about using trade policy as a rent-seeking instrument.[32] In addition, the fatal weakness of the conventional optim

tariff argument seems to afflict the argument presented here. If retaliation ensues, both states will be worse off: the volume of trade will decrease, but the terms of trade will not change.[33] The optimal tariff game, in short, is a PD game.

The novelty of the argument in this book inheres in its conclusions about the PD games that result when states simultaneously attempt to exert power over their terms of trade. The discussion that follows demonstrates that tariff games between allies differ systematically from those that actual or potential adversaries play. It also suggests that the probability that alliances will evolve into free-trade coalitions varies across international systems.

## THE STANDARD OPTIMAL TARIFF GAME

The payoff matrix of the standard optimal tariff game is shown in figure 3.4. In this variant of the PD game, $T$ represents the payoff that accrues to a state that unilaterally deploys an optimal tariff. $R$ is the payoff to each state if both trade freely with each other. $P$ is the return that accrues if both use optimal tariffs. And $S$ is the payoff to unilateral free trade.

In any finitely repeated game, defection is a dominant strategy. As a result, the inevitable outcome of any such game is bilateral tariffs.[34] In contrast, in a game in which states assume at every period that there is some positive probability that the game will continue, a variety of Pareto-improving outcomes, including that of free trade, can be realized.[35]

This can occur, for example, if both states adopt a "grim" strategy: that is, begin with free trade (i.e., cooperation). Then, trade freely in the succeeding period if the other cooperated in the previous period; otherwise, deploy an optimal tariff (i.e., defect).[36] This strategy will sus-

[33] This assumes that states possess identical degrees of market power. If they do not, it is possible for one state to be better off even after the cycle has been completed than if it had pursued free trade. See Johnson (1953–54, 142–53).

[34] As noted above, if incomplete information of a specific kind exists in a finite PD game, it is possible for mutual cooperation to emerge on at least some plays of the game. See Kreps et al. (1982).

[35] As the folk theorem states, if the horizon is infinite and the actors are very patient, *any* individually rational outcome can be sustained (Fudenberg and Maskin 1986).

[36] In a context in which deviations are perfectly observable, maximal punishment makes sense, because it never occurs on the equilibrium path. It is, therefore, costless (Tirole 1988, 252). Moreover, unlike Tit for Tat (Axelrod 1984), a grim strategy is



| | | Column: Cooperate | Column: Defect |
|---|---|---|---|
| Row | Cooperate | $R,R$ | $S,T$ |
| Row | Defect | $T,S$ | $P,P$ |

FIGURE 3.4. The Standard Optimal Tariff Game
*Note:* $T > R > P > S$; $R \geq (T + S)/2$

tain cooperation in a PD game if the discounted sum of cooperative payoffs, $R/(1 - \delta)$, is greater than is the sum of the one-shot gain from defection, T, and the discounted sum of punishment payoffs, $\delta P/(1 - \delta)$.[37] Thus, cooperation can be sustained by a grim strategy if:

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

or

$$\delta \geq \frac{(T - R)}{(T - P)} \tag{3.1}$$

If this incentive-compatibility constraint is satisfied, a grim strategy will enable free trade to emerge as the outcome of an infinite-horizon optimal tariff game. The success of the General Agreement on Tariffs and Trade suggests that this is possible not only in the abstract but also

subgame-perfect; that is, it is rational for the players to follow the specified strategies if any defection occurs (Rasmusen 1989, 91).

It is not renegotiation-proof, however. That is, if one player defects and the players have the opportunity to renegotiate, they will abandon the punishment strategy in favor of one that confers higher payoffs. A Pareto-perfect or weakly renegotiation-proof strategy is a "penance" strategy: Begin with cooperation; if a deviation occurs, the deviator cooperates and the other player defects; the players then revert to cooperation (Fudenberg and Tirole 1991, 179–82). I use a grim strategy here for two reasons: (1) it is more accessible and, more importantly; (2) the results of the analysis do not change if a penance strategy is used.

[37] The discount factor, $\delta$, reflects both the rate of time preference, $\rho$, and the probability that the game will end, $\theta$. The discount factor is $(1 - \theta)/(1 + \rho)$ (Rasmusen 1989, 90).

| | | State $j$: Cooperate | State $j$: Defect |
|---|---|---|---|
| State $i$ | Cooperate | $R_i$-$w_{ij}$ $R_j$, $R_j$-$w_{ji}$ $R_i$ | $S_i$-$w_{ij}$ $T_j$, $T_j$-$w_{ji}$ $S_i$ |
| | Defect | $T_i$-$w_{ij}$ $S_j$, $S_j$-$w_{ji}$ $T_i$ | $P_i$-$w_{ij}$ $P_j$, $P_j$-$w_{ji}$ $P_i$ |

FIGURE 3.5. The Prisoners' Dilemma: Adversaries
*Note:* This payoff matrix omits second-order effects.

in the real world. Established in 1947, the GATT was created to (and did) enable states to avoid a replay of the tariff game that stymied trade among them in the interwar period.[38]

### TARIFF GAMES BETWEEN ADVERSARIES

The payoff matrix of the standard optimal tariff game assumes that private and social returns to interstate trade do not differ. Life in an anarchic international system, however, effectively alters the payoffs of the standard game. Because interstate trade in such a system produces externalities, this trade creates a divergence between private and social returns that any utility-maximizing state will take into account when it calculates its payoffs from trade.

Thus, the payoffs a state assigns to the outcomes of any given trade game differ from those of the standard optimal tariff matrix. In the case of trade with an adversary, a state incurs a marginal social cost that the standard matrix does not reflect. This transforms the standard matrix as shown in figure 3.5. In each cell, the marginal social cost that a state, $i$, incurs is represented as a fraction, $w_{ij}$, of the payoff that accrues to its adversary, $j$.[39] Thus, $w_{ij}R_j$, for example, is the marginal social cost or security diseconomy that accrues to state $i$ when it trades freely with state $j$.

[38] Even though much progress has been made, completely free trade remains elusive. Nonetheless, GATT has clearly facilitated the process of lowering trade barriers among its member countries. Yarbrough and Yarbrough argue that it is neither grim nor any other self-enforcing strategies but third party (U.S.) enforcement that has deterred states from defecting from GATT agreements (1992, 42).

[39] Here, $w$ is assumed to be a constant in order to simplify the presentation. The effect of allowing $w$ to vary across trading partners is discussed in Chapter Six.





This cost is represented as an increasing function of the adversary's gains from trade. The representation is based in part on what Robert Powell has described as a "very simple, highly stylized assumption about the nature of warfare." That is, the "stronger a state is economically, the more likely it is to prevail in war" (1991, 1,312). Although it is not conclusive, of course, the experience of both twentieth-century world wars suggests that this assumption does not depend on abstract reasoning alone. In each case, the aggregate economic output of the belligerent states was a powerful determinant of the outcome of the war (Kennedy 1985).

The functional dependence of social on private returns from trade also has a microeconomic, or sector-specific, component. The increase in domestic resource efficiency that results from trade frees economic resources for military uses. Whether resource diversion to the military will actually occur depends upon the income elasticity of demand for military spending. If military goods are "normal" goods, the income elasticity of demand for them will be positive. Whether or not they are normal goods obviously depends on state preferences. If, for example, a state engaged in an arms race has preferences that conform to a Deadlock game, it is very probable that its arms will increase as its income does. PD preferences make this less likely but by no means impossible.[40] Historically, however, income and military spending have been directly related.[41]

I assume, then, that the marginal social cost that state $i$ incurs when it trades with its adversary, $j$, is directly related to the gains from trade of the latter. Although it is not strictly necessary to the analysis, I also assume that $w_{ij}$ is less than 1; that is, I assume that the marginal social cost associated with trade is only a fraction of the gains from trade that accrue to the adversarial state.[42] I do so because military portfolios reflect the diversified interests of great powers. Typically, any great power has a range of security interests, only some of which affect the interests of another great power. Its portfolio of military weapons will reflect

[40] For an incisive, extended discussion of this issue, see Downs and Rocke (1990, 68–79).

[41] Klaus Knorr has observed that "absolute amounts of military spending are, on the whole, closely correlated with GNP" (1977, 184).

[42] The effect of allowing $w$ to equal 1 would be to eliminate trade between adversaries, because the resulting game would no longer be a PD. Thus, the logic of the argument would not be affected, but the findings would change. See note 44 below.

these interests. Because military resources are not completely fungible, not all additions to the military power of an adversary will disadvantage its trading partner. Thus, even if an adversary did divert all of its gains from trade to the military sector, its action would not necessarily inflict a social cost on its trading partner equal to the incremental addition to its military power.

For example, many resources that the United States deployed to prosecute the Vietnam War would not have been useful in either a conventional or a nuclear war with the Soviet Union. Similarly, Britain could not transfer without any loss in efficiency to its prosecution of World War I the resources it used to protect its colonial empire during the late-nineteenth century. Even if the military sector were the sole beneficiary of state $j$'s gains from trade then, $w_{ij}$ would still be less than 1.

When the returns from trade are adjusted to reflect the trading states' marginal social costs, the net effect is to make tariff games between adversaries more difficult to solve than is the standard optimal tariff game. If the game is infinite, free trade can be sustained in an optimal tariff game between adversaries if:

$$(R_i - w_{ij}R_j) / (1 - \delta_i) \geq T_i - w_{ij}S_j + \delta_i(P_i - w_{ij}P_j) / (1 - \delta_i)$$

or

$$\delta_i^* \geq \frac{T_i - w_{ij}S_j - (R_i - w_{ij}R_j)}{T_i - w_{ij}S_j - (P_i - w_{ij}P_j)} \qquad (3.2)$$

This condition is more difficult to satisfy than is the condition that emerges from the standard optimal tariff game. In a PD game with payoffs of (3, 2, 1, 0), for example, the minimum discount factor that would support cooperation in the standard game is .50. In the transformed game (with $w_{ij} = .4$), the analogous figure rises to .75.[43] Thus, the utility functions characteristic of adversarial states make any given PD more "severe" or more difficult to solve. In addition, as $w$ increases, the severity of the game also increases.[44]

[43] This is adapted from Snidal (1991).

[44] The game will remain a PD as long as $0 \leq w < 1$. If $w$ equals 1, the mutual reward and mutual punishment payoffs will both equal 0, and the payoff to unlateral defection will be positive. There will not, as a result, be any incentive to cooperate even if the game is repeated.



| State $i$ \ State $j$ | Cooperate | Defect |
|---|---|---|
| Cooperate | $R_i + w_{ij}R_j, R_j + w_{ji}R_i$ | $S_i + w_{ij}T_j, T_j + w_{ji}S_i$ |
| Defect | $T_i + w_{ij}S_j, S_j + w_{ji}T_i$ | $P_i + w_{ij}P_j, P_j + w_{ji}P_i$ |

FIGURE 3.6. The Prisoners' Dilemma: Allies
*Note:* This payoff matrix omits second-order effects.

## TARIFF GAMES BETWEEN ALLIES

The gains of one state impose social costs on another only if the former is an actual or potential threat to the security of the latter. When a state trades with an ally, it realizes both private and social benefits from doing so. In the case of allies, then, the external effect of trade is positive. As in the case of trade between adversaries, a utility-maximizing state will take account not only of private but also social returns in calculating its payoffs from trade with allied states.

The standard matrix does not, therefore, apply to this case, either. The transformed PD matrix is shown in figure 3.6. In each cell, the marginal social benefit a state, $i$, realizes as a consequence of trade with an ally is represented as a fraction, $w_{ij}$, of the payoff that accrues to its ally, $j$.[45] This effectively internalizes the positive security externality associated with free trade between allies. The functional form is based on reasoning analogous to that which applies to the case of trade between adversaries.

The applicable incentive-compatibility constraint then becomes:

$$(R_i + w_{ij}R_j) / (1 - \delta_i) \geq T_i + w_{ij}S_j + \delta_i(P_i + w_{ij}P_j) / (1 - \delta_i)$$

or

$$\delta_i^{**} \geq \frac{T_i + w_{ij}S_j - (R_i + w_{ij}R_j)}{T_i + w_{ij}S_j - (P_i + w_{ij}P_j)} \qquad (3.3)$$

[45] If $w$ were to equal 1 in the case of allies, and the condition that $2R \geq T + S$ is maintained, the game would be one in which no state had an incentive to defect from the free-trade equilibrium.

This condition is more easily satisfied than is the condition that emerges from either the standard PD or the transformed game between adversaries. Using the same values assigned in the example above, free trade can be sustained among allies when the discount factor is .13. In addition, if $w_{ij}$ exceeds the larger of $(T - R)/(T - S)$ and $(D - S)/(T - S)$, free trade becomes the dominant strategy equilibrium. Thus, tariff games between allies are easier to solve than are tariff games between actual or potential adversaries.[46]

## The Effects of Polarity

While the analysis above suggests that all alliances will influence trade, casual empiricism suggests that this effect has been much stronger after than before 1945. Inspection of the incentive compatibility constraint that applies to the intra-allied tariff game suggests one explanation of this difference.[47] All else being equal, the extent to which allies trade freely with each other depends upon the discount factor, $\delta$. The discount factor is inversely related to the probability, $\theta$, that the game will end.

In the context of this analysis, $\theta$ can be interpreted as the risk of exit. This risk is the threat that a member will exit or abandon an existing alliance to join an alternative one. For three reasons, it is higher in multipolar than in bipolar systems.[48] First, while alliances in a bipolar world are the products largely of system structure, alliances in a

[46] The use of modified PD supergames is not original to this book. See, for example, Taylor (1976); Grieco (1988, 1990); and Snidal (1991). Powell (1991) assigns a low score to their use, because: (1) the games do not allow explicitly for the use of force; and (2) they do not permit payoffs to change over time.

Powell constructs and analyzes a model which he believes overcomes these problems. His analysis generates a clear and striking result. Whether states trade freely with each other depends on the costs of war. If they are high, states do trade freely; if they are low, they do not.

Powell's effort to build a better mousetrap is not a complete success, however, for two reasons. First, his analysis is knife-edged, in the sense that the gain from unilateral defection in the trade game is sufficient to ensure the victory of the defecting state if war ensues. Second, he assumes that complete information about the net benefits of war is available to states ex ante.

[47] I do not intend to imply that this is the *only* explanation. For other possibilities, see Chapter Six.

[48] Empirical analyses support this assertion. See, for example, Duncan and Siverson (1982).





multipolar system reflect choices among several possible alternatives (G. Snyder 1984, 415). Second, in a bipolar system, realignment is impossible for either pole. Third, each pole in a two-power system bears exclusive responsibility for alliance stability. Neither can expect any other state to prevent the defection of an ally from within its bloc (G. Snyder 1984). In a multipolar system, however, interest in preserving alliance stability can be distributed across more than one pole. As a consequence, alliance stability can become problematic, as each pole seeks to transfer the burden of maintaining the alliance to another.

Because the risk of exit is lower in a bipolar than in a multipolar system, the security externalities of any free-trade agreement are more likely to remain internalized within the alliances of the former than within those of the latter. As a result, allies in a multipolar system tend to assign a higher value to $\theta$. Thus, they will discount the future benefits accruing from open markets among them more heavily than will their bipolar counterparts. In contrast, the greater stability of bipolar coalitions allows the value of future benefits to approximate present benefits more closely. Thus, for any given structure of payoffs, free trade is more likely to emerge within the alliances of a bipolar than of a multipolar system.

Two testable hypotheses emerge from the discussion in this chapter: (1) free trade is more likely within than across alliances; and (2) it is more likely within the alliances of a bipolar than of a multipolar world. The next chapter systematically tests these hypotheses.