Management and Sustainability of Clubs
Conceptual Foundations of NATO Politics

The past two decades confirmed that NATO remains the most significant military alliance of contemporary world politics, an organization that not only admitted a dozen new members in its political and military structures, but has also made major political and military adjustments from its original 1949 design. While abundant literature explores certain aspects of NATO’s transformation, international relations as a discipline lacks a comprehensive, universally accepted theoretical framework explaining NATO’s management and functioning in the post-Cold War world. The first two sections of this chapter define the important concept of alliances and highlight the distinction between research that deals with the origins of alliances and research that studies their functioning and management. Because alliances represent a specific form of advanced intergovernmental cooperation, special attention will be paid to the literature on intergovernmental bargaining and political integration. The following sections reevaluate the logic of collective goods theory and examine the relevance of club goods for studying contemporary NATO politics. Lastly, the final section discusses the mixed method employed in this study and some of its constraints.

ALLIANCE FORMATION

There is a general agreement among IR theorists that “it is impossible to speak of international relations without referring to alliances.” Alliance is a “formal association of states for the use (or nonuse) of military forces, in specified circumstances, against states outside their own membership.” This
Chapter 1

definition highlights several important features. First, it refers to an association with a military or security purpose that involves at least several separate state units and requires a formal agreement between the participants. Second, alliances differ from alignments, which are based solely on common interests and informal commitments. Third, they are formed in reaction to or against other states outside their membership. Sometimes alliances are defined as both “formal and informal arrangement for security cooperation between two or more sovereign states.” Both definitions highlight several distinctive features of these associations: alliances deal with sovereign states as units of analysis; they address the issues of international security, and, thus exclude any other form of international interaction. Alliances, therefore, are relatively narrowly and explicitly defined forms of cooperation among nations.

By and large, current research on alliances focuses either on their formation or their functioning and management. Realists identify three core variables that shape alliance formation: power, threat, and interests. From the perspective of power politics alliances are seen as “a necessary function of the balance of power operating within a multiple-state system.” Whether or not a nation should pursue a policy of alliances is “not a matter of principle but of expediency.” States can choose between two distinctive policies of alignment—they can either align with a certain power, and therefore bandwagon, or balance against it. The choice of political actors to “balance each other or climb on the bandwagon depends on the system’s structure.” Thus, balancing and bandwagoning are two contrasting types of foreign policy behavior which are induced by the concern of the states to maintain their positions in the system.

Other scholars attribute the choice of great powers and weaker states aligning to the potential threats that these states face. This theory argues that “the forces that bring states together and drive them apart will affect the security of individual states by determining both how large a threat they face and how much help they can expect.” States can either balance a threatening power or bandwagon with the most threatening state. In the balance-of-threat theory balancing is defined as alignment against the source of threat while bandwagoning is alignment with the source of insecurity. The former policy is far more common than that latter. Additionally, states may also choose to ally with stronger states if other states with offensive military capabilities are likely to provoke other nations to form defensive coalitions. When choosing between the policy of balancing and bandwagoning, states are also constrained by their size and power— weaker states are more likely to bandwagon, rather than balance because they can do little to affect the outcome. Alternatively, “strong states can turn the losing coalition into a winning one because their capabilities may play a key difference in the outcome of a con-
flict.” When no allies are at hand, states may choose to ally with the most threatening state in order to survive.

The balance-of-interest theory challenges the postulate that balancing and bandwagoning are opposite behaviors motivated by the same goal to achieve greater security. According to Randy Schweller, however, states have very different reasons to balance or bandwagon. While the aim of balancing is self-preservation and, therefore, is driven by the desire to avoid losses, bandwagoning is usually driven by the opportunity for gain. Thus, alliances are approached as a response both to threats and opportunities.

Aside from systemic features, such as power, threat and interests, the formation of alliances may be linked to various domestic variables. Stephen David makes the case that states form alliances not only to respond to threats emanating from other powers, but also to deal with internal, state-related or regime-related threats. Richard Harknett and Jeffrey VanDenBerg found out that countries experiencing competing allegiances, lack of legitimacy and control of the state apparatus, form alliances as “a reaction not to the presence of an external threat, nor by domestic political compromise, but by the presence of both external and internal security challenges that reinforce each other.” Therefore, the behavior of states is often determined by the presence of internal or interrelated threats. Such a behavior makes the explanation of alignments more complex and difficult. For example, in order to balance against an immediate internal threat, a state aligns with an external power that poses a long term security challenge. Conventional alliance theory will interpret this as bandwagoning while in fact it is balancing against an impending domestic threat. Kegley and Raymond expand this framework and suggest that alliances can also bestow additional “private” benefits. These include the increase the military capabilities of certain states, as well as political opportunities for state’s leader to consolidate support at home by gaining political recognition, economic support, or legitimizing ideology.

Democratic peace theory (DPT) offers an alternative explanation. This approach focuses on democratic governance and studies how it affects the choice of allies. For Michael Doyle “even though liberal states have become involved in numerous wars with non-liberal states, constitutionally secure liberal states have yet to engage in war with one another.” It is not that such a war is impossible, but rather that “preliminary evidence does appear to indicate that there exists a significant predisposition against warfare between liberal states.” Therefore, variables measuring the level of democratic governance such as constitutional law, republican representation, and the separation of powers determine whether states will enter into an alliance relationship with other states that share similar characteristics. Thus, any step toward armed conflict among those liberal and democratic countries would
be much slower and cumbersome because “liberal states do exercise peaceful restraints and a separate peace exists among them. And this peace provides a solid foundation for the United States’ crucial alliances with the liberal powers.” Therefore, after 1990 NATO has embarked on a new “mission of Europe whole and free” reflecting an “evolving concept of security underpinned by a considerable faith in the pacifying effect of shared democratic institutions and values.”

ALLIANCE MANAGEMENT

International alliances are intergovernmental bodies of international cooperation, which involve a great deal of bargaining as means to exercise leverage over different partners. Andrew Moravcsik indicates that that three key variables influence the outcome of intergovernmental bargaining among the EU members—interest, power, and credibility of commitments. He argues that the outcomes of such interstate bargaining are generally Pareto-efficient, but their distributive outcomes vary greatly. In other words, no other bargaining outcome exists which can improve the well-being of one negotiating country without harming another one. When an alternative allocation exists, so that it is possible to improve at least one person’s well being without harming anyone else, then the Pareto optimality requires that the actors choose the alternative. This analysis of intergovernmental bargaining based on the relevance of interests, power, and credibility of commitments, rests on several core assumptions of rationality: First, rational action involves some sort of utility maximization, i.e.,, a situation in which an individual is confronted with an array of options and he or she picks up the one that best serves their objectives. Second, an individual’s available options are rank-ordered, i.e.,, they may consider the two available options unequal or equal. If these options are unequal he or she prefers one to the other and if they are equal, it means that the individual is indifferent. Third, the preference orderings of the rational actors are transitive. Fourth, each individual maximizes the expected value of their payoffs measured on some sort of utility scale. As a result, the maximizing actions of the individuals produce certain collective outcomes. Moravcsik found that, in the context of European integration, certain bargains may favor Germany, while in other situations they may favor France or Britain.

Alliance bargaining includes a large set of issues such as contributions to military preparedness, burden-sharing, common military action, or joint strategy in case of war. In the past allied bargaining was less salient, more sporadic, and with “widely differing intensity.” The introduction of NATO
and the Warsaw Pact on the world stage following World War II increased
the frequency of bargaining among allies, helped shape the image of adver-
sary, and determined the need for a strategy to evaluate adversary’s inten-
tions and capabilities. For example, the existence of diverging assessments
of deterrence in the case of NATO accounted for differences among allies
over strategy resulting in the negotiation of a vague and broad wording of
the strategic concept of Flexible Response. Therefore, the Flexible Response
made it possible to project several strategies based on the variation of threat
perception among member states. As a result, alliance theorists have sug-
gested contending rationalist explanations of alliance bargaining. Some
scholars view it as a part of a broader decision-making process in which in-
dividual members forward their interests and use the alliance’s assets, rules,
norms, and procedures to maximize membership benefits. Others explain
alliance management with the power that individual states exercise in order
to influence the outcome of the bargaining process. Yet others focus on the
importance of the credibility of commitment to secure the cohesion and over-
all sustainability of alliance structures.

Interests

Alliances may be established for various reasons—to maximize power, to
balance against, or bandwagon with the threat, to address the interests of
the individual members, or to form a pacific community of nations. Once
established, however, they require serious efforts on the part of the members
to coordinate their individual interests. Thus, alliances may sometimes serve
as mechanisms to reduce costs “by spreading them among several partners”
and providing benefits that that cannot otherwise be obtained unilaterally.
Alliances can also provide a “medium for exerting leverage over partners,”
where great powers may gain increased influence over other countries’ for-
eign policy decisions. Such influence may flow from the alliance leader to the
smaller partners, as well as from the smaller to bigger allies.

States participate in tacit or explicit bargaining over common and competi-
tive interests. Their primary common interest is to preserve the alliance and
keep the benefits of membership flowing, whereas their primary competitive
interest is to manage allied influence in order to minimize costs and risks.
Earlier studies have established a causal link between the interest of the in-
dividual allies and the size of the alliances that they form. In his 1962 study
on political collations, William Riker ascertained that in social situations
similar to n-person, zero-sum games with side-payments, the participants
have interest to “create coalitions just as large as they believe will ensure
winning and no larger.” International alliances are one variation of such
minimum winning coalitions and can be maintained as long as they have sufficient number of allies to ensure the existence of these coalitions. Mancur Olson and Richard Zeckhauser extended this logic to explain why the United States carried a larger share in the overall burden within NATO during the Cold War. They discovered that a tendency toward disproportionality exists in sharing the burden between large and small nations, whereby the larger nations bear disproportionately bigger shares of the costs associated with the common defense relative to the small ones. That was the case of NATO after World War II when, as a result of the implicit division of labor between the United States and its European allies, the United States shouldered greater military burdens and exercised “correspondingly greater influence in alliance decision-making, including the articulation of NATO military strategy.” In order to prevent such disproportionality, alliances establish institutional agreements that enhance coordination and also introduce arrangements for sharing various costs on a percentage or any other basis.

Scholars of neoliberal institutionalism build on this logic and redefine the notion of state interest by highlighting the joint benefits of international institutions to improve cooperation. Therefore, in this framework, alliances are seen as institutions constructed by governments to “provide information, reduce transaction costs, make commitments more credible, establish focal points for coordination, and in general facilitate the operation of reciprocity.” Institutionalists challenge the logic that states’ concerns for relative gains prevails, and support that the importance of institutions is “conditional on factors, such as the number of major actors in the system and whether military advantages favor offense or defense.” Institutionalists approach NATO as a regional security regime that establishes stable norms and rules which lead to stability in conventional military forces. They argue that the variation in NATO’s institutional adaptation is explained by “variation in relative costs,” and “by whether the rules, norms, and procedures of a given institution enable states to overcome obstacles to cooperation (such as provisions for sanctioning or bargaining).” Therefore, international structures like the United Nations or an enlarged NATO, serve as instruments to improve cooperation and manage effectively security threats.

Power

Realism and neoliberalism are the two main positivist traditions focusing on the relevance of power in the context of alliance politics. While classical realists focus on power as a state-related characteristic, neo-realists and neo-liberals share a common assumption that power is a systemic variable. Currently, the international relations literature distinguishes between defensive
and offensive realist assumptions. When applied to the study of alliances these distinctive sets of assumptions shape varying expectations about state behavior. Defensive realists posit that states may engage in various unilateral efforts or military cooperation, including alliances, in order to balance the power of threatening states. Once the threat ceases to exist, sustaining this cooperation becomes much more difficult. Defensive realists believe that NATO was created to balance against the Soviet threat and that when the latter ceased to exist, they expected the demise of the alliance because the presence of American forces will become an irritant to the European states, whose security is not threatened. In 1993, Kenneth Waltz boldly predicted that “NATO’s days are not numbered, but its years are. Some hope that NATO will serve as an instrument for constraining a new Germany. But once the new Germany finds its feet, it will no more want to be constrained by the United States acting through NATO than by any other state.”

For defensive realism NATO can only sustain itself as a structure of international security and be managed successfully if a threat exists that “can provide sufficient glue to hold the alliance together.”

Offensive realists highlight the distinction between great powers and the other actors in an anarchical international system. Great powers are determined “largely on the basis of their relative military capability,” such as sufficient military assets to fight successfully against the most powerful state in the world. Power is the key variable that drives state behavior, including alliance management. Power is composed of two components—latent and military power. The former is based on the size of the population and the level of its wealth, while the latter is embedded in the army, the air and naval forces that directly support it. Essentially, the theoretical argument of offensive realism rests on two separate pillars. The first one is built on the assumption about the ever present potential for offensive power coupled with uncertainty, where as the international system forces great powers to maximize their relative power because that is the optimal way to maximize their security. The second pillar rests on the logic of water’s stopping power because large bodies of water sharply curtail the power projection capabilities of armies and thus shift the balance of land power in important ways. Hence, land power is the central ingredient of military power, because armies are “the principal instrument for conquering and controlling territory.”

Even though other realist scholars criticize the logical consistency of the argument about the stopping powers of water, they accept the key assumption that NATO is an asset to maximizing power. For offensive realism the United States would not be willing to give up power and could play on the uncertainty of the Soviet Union’s collapse in order to encourage new allies to join the Alliance. Furthermore, Washington continues to dominate NATO
decision making, much the way it did during the Cold War “making it difficult for many of those states (especially Germany) to cause trouble with Russia.”

Thus, realism approaches post-Cold War NATO as a tool for managing power relations between the allies who differ from each other primarily due to the variation of their capabilities, holding that one of them is much more powerful than the others.

Even though majority of realist scholarship on NATO converges on the view that the United States has unmatched power capabilities realism remains divided on United State’s ability to sustain the current distribution of power. The unipolar optimists argue that American hegemony is beneficial for the United States and for the international system as a whole. As a result, this imbalance can be sustained under certain conditions without necessarily resulting in intensive balancing against the emerging hegemon. TV Paul, James Wirtz, and Michel Fortmann suggest a revision of the traditional realist logic beyond the balancing and bandwagoning explanations. They argue that soft balancing exists under “near-unipolarity” that involves a “mixture of cooperative and balancing behavior.” In this strategic environment the functioning and management of alliances like NATO can be attributed to the multi-faceted relations in a near-unipolar world. Alternatively, unipolar pessimists share skepticism about the sustainability of the U.S.-dominated unipolarity and believe that at best it will only last for another decade or two. The U.S. hegemony will trigger a counter hegemonic backlash against the United States and the costs of preserving America’s preponderance will outweigh the benefits. Terrorism, “soft,” and “semi-hard” balancing are only several instances of such balancing behavior. Although realism has dominated IR theory for quite some time, it fails to present a comprehensive theory of alliance management. Scholars like Waltz, Mearsheimer, Layne, and Wohlforth disagree on whether the current distribution of power would be sustained and, therefore, they offer little insight on the process of adding new allies, missions, and capabilities into NATO.

The neoliberal tradition suggests a very different interpretation of power and alliance management. In general, neoliberal approaches assume that, in a world of anarchy, cooperation is possible if states have significant common interests. Unlike realism, however, the rational behavior of the states is not determined by their sheer survival but by the utility maximization of international interaction. States have consistent, ordered preferences; they calculate “costs and benefits of alternative courses of action in order to maximize their utility in view of those preferences.” Neoliberal theories reject relative power as a key variable driving international interaction and argue that states focus on absolute gains that flow from international cooperation. They
accept the possibility of sustained imbalance and attribute successful alliance management to the effect of international institutions, the uniqueness of the American political system, or the pervasiveness of democracy after the end of the Cold War.

Celeste Wallander, Robert Keohane, and Joseph Nye view international alliances not as a mere aggregation of military capabilities, but as security coalitions differing in their purposes and “degree of their institutionalization.” These institutions are more ‘portable’ than the rules and practices of single purpose alliances focused solely on threat. The seamless web of multilateral institutions of which NATO is also a part “allow others to participate in decisions and act as a sort of world constitution to limit the capriciousness of American power” thus providing for transparency, consultation, and incentives for cooperative strategies among members. Therefore, the North Atlantic Alliance is becoming an institution whose military functions will decline as threats diminish and “it should gradually expand to encompass all democratic European states that are committed to peaceful, friendly relations based on territorial status quo.”

Institutional assets and adaptation present one neoliberal explanation for the sustainability of the current distribution of power. Joseph Nye and G. John Ikenberry, attribute the successful alliance management in the post-Cold War world to the hegemonic distribution of power, which combines the uniqueness of the democratic political system with America’s exceptionalism. They explain the stability of the Pax Britannica and Pax Americana with various domestic arrangements and the exercise of “soft power” in the international system that ensures reliable peace and security. Thus, the stability and expansiveness of the post-Cold War world is based on the United States’ unmatched economic and military power today” coupled with its unique capability to engage in “strategic restraint, reassuring partners and facilitating cooperation.”

To sum up, neoliberal theories offer multiple explanations about NATO’s management in the post-Cold War world. Some attribute NATO’s persistence to its assets as an institution of international security, while others highlight the significance of democracy and the uniqueness of American hegemony in managing NATO politics and steering its transformation after the end of the Cold War. However, the problem with these explanations is that they underestimate the fact that alliances are formal associations with military or security purposes. Thus, neoliberal approaches provide abundant explanations as to why NATO stays in business and continues to expand, yet they fail to explain how the Alliance manages its new out-of-the-area missions, deals with the new members, and promotes the development of new allied capabilities.
Credibility of Commitment

Glenn Snyder defines commitments as “arrangements of value that disposes one to act in a certain way.” Commitments emanate from two possible sources: the promise (verbal or written) in an alliance contract and the interests in aiding an ally that exists apart from the promise itself. Interests determine the behavior of the parties in a specific issue about which they are bargaining. Albeit constituting a moral and in some cases also a legal obligation, the contact itself is insufficient to build a strong alliance commitment unless additional political values are attached to it, such as prestige and a reputation for honoring agreements. Furthermore, the degree of commitment, which includes the pledge’s explicitness, determines the precision of casus foederis. The more explicit and precise the commitment, the higher the cost in non-fulfillment and, therefore, the greater the credibility of commitment. In the case of NATO, the credibility of commitment is formally embedded in Article Five of the North Atlantic Treaty, which explicitly states that an armed attack on one or more of the allies in Europe or North America “shall be considered an attack against them all.” The precision of casus foederis ensures that if such an armed attack occurs, each ally will “assist the Party or Parties so attacked by taking forthwith, individually and in concert with the other Parties, such action as it deems necessary, including the use of armed force, to restore and maintain the security of the North Atlantic area.” NATO’s powerful and explicit commitment is unique in world history and no other political and military alignment has ever had the same magnitude of commitment. That is why the alliance has been able to attract almost a dozen new members since its inception in 1949.

The commitment is also linked to each state’s underlying interest in defending its partners. If a state has strategic or intrinsic interests, such as preventing allied resources from coming into the hands of a potential opponent, it is more likely that it will honor its pledge or reciprocate its partner’s expectation. As previously mentioned, the commitment is determined by the power of the allies. Glenn Snyder correctly notes that “a strong state will have a clear interest not only in the existence and independence of a weak partner, but also in acting to protect the partner, since the partner cannot defend itself.” Alternatively, if the ally is strong enough to defend itself, the weaker state may have no stakes in acting to preserve its stronger partner even though it may be interested in preserving the alliance. Thus, the precision of commitment is constrained by the two other variables that shape intergovernmental bargaining—interests and power.

The credibility of commitment also has direct implications for the level of allied dependence measured in terms of net benefits that each state receives from an alliance, compared to the benefits available from other alternative
sources. Allies seek the optimal level of commitment that will not only guarantee their security but also preserve the alliance and maximize the net benefits from it. For this purpose the alliance members are careful not to take excessive responsibilities for others’ security and, hence, they face two opposing fears: the fear of abandonment and the fear of entrapment. Entrapment describes a situation of a state being dragged by its allied commitments into a conflict over an issue (or issues) about which the state is not concerned. The fear of abandonment becomes pervasive when the worry arises that a state has alternative partners and may choose one of them over its present allies. Two components of this fear exist—the probability calculus that this partner will actually act and the cost associated with any such decision. Contemporary alliances such as NATO have been able to establish various formal and informal mechanisms for consultation and exchange of information aimed at undermining this fear by increasing transparency and ultimately minimizing the risk of defecting. Another form of abandonment—the failure to fulfill alliance commitments or to renege when the *casus foederis* arises—has been an issue at NATO in regard to the operation in Afghanistan because some allies are more involved than others in actual combat operations against the Taliban insurgency. By the same token, an example of the fear of entrapment can be viewed in France’s decision to withdraw voluntarily from NATO military structures in 1966. French political leadership, however, realized that this fear was unwarranted and therefore Paris returned to the military command in April 2009.

The explanations based on interests, power and credibility of commitments embrace a rationalist approach which is not universally accepted among political scientists because it fails to explain how integrative processes influence the course of international relations and alliance politics. Alternatively, scholars like Ernst Haas, Karl Deutsch, Emanuel Adler, and Michael Barnett study the effects of political integration and security communities and suggest alternative explanations of the alliance transformation.

**POLITICAL INTEGRATION AND ALLIANCE POLITICS**

Scholars of political integration noticed in the 1950s a novel pattern of political behavior in which “powers and duties being handed over to regional and other authorities for the better performance of certain communal needs.” This process of political integration advanced in spite of old constitutional division among nations. A self-propelling tendency occurred during which new authority gradually emerged “above” the territorial states. This process known as functional integration is driven by organized economic interests
and occurs in several separate steps. It requires a wider coordination that starts with technical and functional integration within the same group and gradually spills over to coordination of several groups of functional agencies of different nations; and, finally leads to a fusion between these functional agencies and other international planning agencies. As a result, single body or multiple advisory supra-national organs emerge that “have some overall political authority” above the interests of the individual states.\(^{52}\)

Yet, there is a distinction between economic and security integration. Initial decisions to integrate economically are made by interest groups that organize and develop their own norms and procedures. These groups exercise pressure on states’ governments “to manage economic interdependence to their advantage by centralizing policies and institutions.”\(^{53}\) The entire process leads to functional and political spillovers, whereby the pressure for integration in certain sectors propels pressure for further integration in others. The same cannot be observed about political integration because military and defense issues “have not displayed a close affinity to integration unless the issue involves a related question of saving and allocating resources for welfare measures.”\(^{54}\)

The effects of political integration are explained through the idea of security communities. These communities are “social groups with a process of political communication, some machinery for enforcement, and some popular habits of compliance.” They possess a capacity to act as a political unit and have “the ability of a unit to control its own behavior and to redirect its own attention.” By merging security and community, new types of interactions emerge where “the members of that community will not fight each other physically, but will settle their disputes in some other way.”\(^{55}\) The degree of integration highlights a key distinction: The amalgamated communities present a formal merger of two or more previously independent units into a single larger unit such as in the cases of the United States and Switzerland. Alternatively, pluralistic security communities retain the legal independence of separate governments as autonomous units where “two separate governmental units form a security community without being merged.” Karl Deutsch recognizes that the “pluralistic security communities turned out to be somewhat easier to attain [and] preserve than their amalgamated counterparts.”\(^{56}\)

The security communities framework can be applied in the context of NATO politics. The Alliance presents an example of such “pluralistic communities of shared values,” where members “perceive each other as peaceful and [. . . ] are likely to overcome obstacles against international cooperation and to form international institutions such as alliances.”\(^{57}\) Several core conditions exist that facilitate the advancement of these communities: (1) compatibility of major values relevant to political decision-making; (2) capacity of the participating political units to respond to each other’s needs, messages,
and actions quickly, adequately and without resorting to violence; and (3) mutual predictability of the behavior required for the members of the community in order for it to function. In addition to these background (or core) conditions, the formation of these communities requires the length of time over which the pattern of integration persists and contributes to their consolidation. Security communities have two important implications for understanding alliance management. First, their formation affects the behavior of interest groups, political parties, and governments acting in an intergovernmental and supranational setting, which seek coalitions at a supranational level based on common ideology. Second, the conceptualization of security communities is important in explaining how political integration affects alliance management through constructing novel types of relations among the different actors of international politics. Third, the formation of pluralistic communities takes some time after alliance structures are established and several different conditions need to be present. These include: precipitating conditions; conditions that facilitate the integrative transactions, organizations, and social learning; and the presence of mutual trust and collective identity. This approach is certainly valuable in explaining change in international politics made through political integration among allies, thus sharpening the debate with realist and neoliberal theories. At the same time, the framework does not render much explanatory power how the allies manage their military and political resources in the new integrative communities. Addressing these questions necessitates a careful look at the existing empiricist research. Alternatively, the collective goods model suggests a parsimonious and yet accurate approach, one that has been somewhat overlooked by international alliance scholars. The distinctive features of the different rationalist theories dealing with the functioning and management of alliances are summarized in table 1.1 below.

THE COLLECTIVE GOODS THEORY

Mancur Olson, a pioneer of collective goods research, argues that each group, whether large or small, “works for some collective benefit that by its very nature will benefit all of the members of the group in question.” Although all members have a common interest in obtaining this collective benefit, they have no common interest in paying the cost of providing the good. Because military alliances are, first and foremost, defensive by nature, their core purpose is to deter potential aggressors. Therefore, alliances are created to provide their members with collective defense, which is a pure public good. Generally, the literature on international cooperation distinguishes between two major characteristics of international goods provided as a result
Table 1.1. Major Theories of Alliance Politics

<table>
<thead>
<tr>
<th>Theories of alliance formation</th>
<th>Independent or Causal Variable(s)</th>
<th>Theoretical Model or Explanation</th>
<th>Relevance for Alliance Politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance of Power</td>
<td>Power in the international system</td>
<td>States form alliances to balance against preponderance</td>
<td>Systemic explanation congruent with the logic of defensive realism</td>
</tr>
<tr>
<td>Balance of Threat</td>
<td>International Threat Perception</td>
<td>States tend to ally against their most threatening counterparts</td>
<td>Balancing and bandwagoning are opposite types of behavior—stronger states choose to balance against threats, weaker states tend to bandwagon with the source of threat</td>
</tr>
<tr>
<td>Balance of Interests</td>
<td>State power to achieve greater security</td>
<td>Alliances are responses to threats and opportunities—states balance to ensure self preservation and bandwagon to take the opportunity to gain</td>
<td>Alliance politics affects states at unit level and systemic level; balancing and bandwagoning are not opposite types of behavior</td>
</tr>
<tr>
<td>Omni-balancing and Omni-alignment</td>
<td>International threats; Internal (state- or regime-related) threats</td>
<td>States form alliances as a reaction to the presence of both external and internal security challenges that reinforce each other</td>
<td>States ally with an external power that possesses a long-term security challenge in order to balance against an immediate internal threat</td>
</tr>
<tr>
<td>Democratic Peace</td>
<td>Regime type democracy vs. non-democracy</td>
<td>Democracies tend to ally with other democracies against non-democratic states</td>
<td>Alliances like NATO have a pacifying effect of shared democratic institutions and values; they contribute for a lasting peace</td>
</tr>
</tbody>
</table>

1. **Intergovernmental Bargaining**
   - Power, interests and credibility of commitment
   - The outcomes of interstate bargaining are generally Pareto-efficient; i.e. there is no other situation when all allies are better off without making the other allies worse off.
   - Certain bargains may benefit some allies, others may benefit other allies. Overall, alliance members are rational utility-maximizers.
Minimum Winning Coalition  The power of the individual nations; size of collations and alliances

Neoliberal institutionalism  Alliances as international institutions (both independent and dependent variables)

Collective goods theory  The size of the allies (in terms of GDP) is linked to its contribution to common defense

Club goods theory  Size of the allies, allied resources (esp. defense spending); country-specific assets

Security Communities  Power is transferred; new authority gradually emerges above territorial states

In n-person, zero-sum games, participants create collations that are just as large to ensure winning and no larger

Institutions provide information, reduce transaction costs, enhance credibility of commitments and facilitate reciprocity

Alliances provide pure public (or collective) goods in the form of collective defense. As a result, there is disproportionality in sharing the burden among different allies

Alliances operate similar to clubs—they provide excludable goods that have both public (collective) and private features.

Powers and duties are handed over to regional and other supranational authorities for the performance of certain communal needs

International alliances are maintained only as long as they have sufficient number of allies to ensure their existence

Alliances like NATO serve as regional security regimes that establish stable norms and rules. They are flexible and adaptable structures enabling states to overcome the obstacles of cooperation

Smaller nations tend to contribute significantly less than their larger counterparts; alliances establish various institutional arrangements to remedy this disproportionality

There are homogeneous and heterogeneous clubs; most of the alliances are heterogeneous, i.e. the size of the club and the identity of its members determines the outcome of intra-club bargaining

Pluralistic security communities are formed; as a result of political integration new types of relationships occur among different allies
of interaction among nations: nonrivalry and nonexcludability. Goods are nonrival when their units can be consumed by someone “without detracting, in the slightest from the consumption opportunities still available to others from the same unit.” Examples of such goods are sunsets, unobstructed views, or pollution-control devices. Those goods whose benefits can be withheld by the owner or provider at a certain cost for others display excludable benefits. Alternatively, benefits that are available to all once the good is provided are termed nonexcludable. Usually most of goods are excludable, such as territory, oil, clothes, food, etc. Additionally, private goods exhibit both rivalry and excludability, such as territory or natural resources, whereas public goods exhibit neither rivalry nor excludability, such as access to radio waves, oxygen, or public parks. Furthermore, a third category of goods with mixed characteristics exists. Depending on whether goods are excludable and whether they exhibit rivalry, these are called coordination goods and common property resources. These fall between the two extremes of purely private and purely public goods, and also provide partially rival or partially excludable benefits as shown table 1.2.

Scholars of international relations use collective security as an example of a purely public good, whereby the international community has at its availability a “self-enforcing international mechanism or body that automatically reacts by rejecting and immediately renouncing an aggression in the international system.” The collective security concept, championed by Inis Claude, signifies an “all-inclusive,” inner oriented system “involved in the safeguarding of all against aggressive assault.” Such a system regulates international behavior by creating institutionalized confidence among participants. Charles and Clifford Kupchan argue that the establishment of a collective security system in Europe led by NATO would provide a collective good and allow Russians to define themselves as members of a European community of nations, rather than as outsiders. In essence, a collective security system is a “pact against war,” signed by all states in the international community, and collective security mechanisms are directed towards “no predetermined or clearly defined enemy, nor can [a collective security system] operate on the basis of a predetermined coalition.”

Table 1.2. Types of International Goods (Barry Hughes, 1993)

<table>
<thead>
<tr>
<th>Excludability</th>
<th>Rivalry</th>
<th>Non-rivalry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private goods</td>
<td>Impure public goods: coordination goods</td>
<td></td>
</tr>
<tr>
<td>Non-excludability</td>
<td>Impure public goods: common property resources</td>
<td>Public goods</td>
</tr>
</tbody>
</table>
tions offers examples from the inter-war period during which the ideas of collective security were implemented to help regulate various territorial disputes such as the plebiscite in Upper Silesia in 1921, the border dispute between Greece and Bulgaria in 1925, and the 1935 referendum in Saarland. Therefore, ideas associated with a collective security system assume that security is a public good, provided not only to the members of a certain region, alliance, or international organization but also to all states in the international system.

Collective goods theory also accepts the microeconomic classification of the goods according to their elasticity of expenditure into four categories—inferior, inelastic, elastic, and superior. The expenditure for superior goods increases by as much or more than the income of the consumers. Inferior goods exist when the expenditure on the good decreases or is unchanged as income increases, i.e.,, income elasticity of expenditure (E) £ 0. A good is inelastic when the expenditure on the good increases but by a smaller percentage than the income; therefore, 0 < E < 1. Alternatively, when the expenditure on the good increases by a percentage that is as great as or greater than the percentage by which income increases but by a smaller absolute amount, then the good is elastic. For elastic goods, the elasticity is in the range between 0 and S0/Y0 (0£E<S0/Y0), where S0 is the expenditure of the good when income is Y0.68

The collective goods model assumes that this microeconomic classification of international goods applies to alliances. As a result, several distinct observations about their management could be inferred. First, this approach is consistent with inter-governmental bargaining logic. This logic assumes that the decision to allocate resources to military alliances or any other international body is determined by the national interest of each country. Second, in the case of NATO, the model found a positive and statistically significant relationship between the size of the members in terms of gross national product and the percentage of their GNP that they allocate for defense.69 As a result, smaller states tend to rely on bigger states for their defense, while bigger states always pay a disproportionately larger share. In other words, the size of the members determines their behavior in the specific alliance setting. Third, the theory identifies one particular situation in which there will be no tendency toward disproportionality, which exists when defense is regarded as a superior good. This situation holds true when any increase in the national income is used to buy only this good or when none of the increase in income that the nation receives in the form of defense is spent on goods other than defense. Peaks of all-out war or exceptional insecurity create a situation in which security is regarded as a superior good and, in this case, alliances will not have any tendency toward disproportional burden sharing. The U.S.-British alliance during World War II is an example of such a situation: as the
United States mobilized, Britain proportionately maintained, if not increased, its military capability from 1941 to 1944, and therefore proportionally shared the burden. Finally, to avoid disproportionate burden sharing, larger states initiate different institutional agreements that regulate the contribution of the individual members convince their smaller partners to accept them. These agreements vary from the coordination of allied efforts to the sharing of certain costs on a percentage basis. In the absence of any institutional arrangements, the tendency toward disproportional burden-sharing is actually stronger than it would be otherwise.70

Nonetheless, these observations do not imply that irreconcilable differences among the allies and fairly considerable divergence of purpose among them would automatically destroy the alliance. While in some cases limited disagreements among members may have a negative impact on alliance politics, in other cases the differences among the allies may actually increase the effectiveness of the alliance. The Operation “Allied Force” in Yugoslavia in 1999 provided an example of differences that occurred among the contributing NATO countries regarding the approach to air war, the sharing of sensitive information, and the introduction of ground forces. These differences considerably strained the alliance and highlighted key differences in the perspectives of the United States and its European allies. Nevertheless, the awareness that alliance unity was at stake in Kosovo helped motivate members to work together because NATO’s credibility was at stake. After the operation’s completion the alliance identified several major areas that required efforts for enhanced cooperation, such as improved command structures, new defense capabilities initiative, an improved consultation process, and a more flexible system for operational decisions.71

Collective goods theory offers several explanations about the disproportionality of burden-sharing among small and big allies. First, alliances are not burden-free for small states. Often when the great powers act, the small states have fewer available options to choose from. Therefore, they are more likely to cooperate, especially in areas that these nations may not have previously considered. Second, the presence of small and big states in an alliance may present economies of scale. Organizations in general strive for a larger membership because they provide certain benefits that increase proportionately as membership increases. In other words, the admission of new allies expands the range of the collective good because ceteris paribus the new members reduce the marginal costs for defense. The production of special types of weapon systems is a good example which illustrates how the expansion of alliances generates economies of scale and enhances interoperability if the allies purchase the same type of military equipment. Nonetheless, if a potential new entrant has disproportionately high security needs, it creates a dilemma
of entrapment wherein all other allies have to bear tremendous costs for its defense as would have been the case of Georgia during the 2008 war with Russia had this nation been a NATO member at that time.

Third, alliances are involved in situations that contain, at least in the near term, a strong element of irreversibility, i.e., once states join an alliance leaving or stepping back on its commitments becomes much more difficult. This does not mean that the states cannot renege on their commitments—in fact history shows that most of the pre-World War II alliances have eventually dissolved—instead, it implies that, at least in the short run, the allies find it considerably more difficult to walk away from such commitments and this situation occasionally creates a sense of entrapment. An example of such a tendency can be found in the dynamic of U.S. defense spending during the Cold War. Albeit it was not always on a continuous upward trajectory, during the Cold War Washington usually spent on average about 4 percent of its GDP on defense. Despite some fluctuations that can be attributed to various domestic factors, the United States could not afford to fall below this critical threshold due to the Soviet threat. The United States also possessed very few tools available to convince its European allies that they needed to contribute more to the allied efforts. Fourth, alliances can offset the effect of the free rider problem or “the exploitation of the big by the small” because they provide not only collective, but also purely national, non-collective benefits to the nations that maintain them. Nonetheless, the main conceptual problem of the collective goods model is that it does not draw any analytical distinction between collective defense and collective security and, therefore, undervalues the private or impure public features provided by alliances such as NATO. The club goods theory offers an important corollary to this model.

THE CLUB GOODS MODEL

Richard Cornes and Todd Sandler championed the club goods theory as an alternative approach to studying international goods. They define clubs as voluntary groups “deriving mutual benefit from sharing one or more of the following: production costs, the members “characteristics or a good characterized by excludable benefits.” Clubs possess several important characteristics: “excludability” (i.e., clubs discriminate between members and non-members), voluntarism, and sharing a certain type of good. Therefore, members choose to join certain clubs because, as rational actors, they anticipate certain benefits from membership. Thus, the utility jointly derived from membership and the consumption of other goods for each of the members must exceed the utility associated with non-membership status. For example,
in the case of NATO the members expect that the benefits of membership outweigh the costs associated with it.

Clubs incorporate the idea of sharing. They may share the use of impure public goods or the enjoyment of the desirable attributes of members. Whether this is a golf club membership, access to the clubhouse, or collective defense, club members inevitably need to share some sort of club facilities or goods. Therefore, the idea of sharing resources and access to goods results in competition over the access to these goods. Because, in most occasions, these resources are limited, the competition results in rivalry for club benefits, causing detraction from the quality of the services previously received. A classic example of such rivalry among nations includes the control of territory, especially if it has valuable natural resources such as fishing whales in the open seas. Usually club members share partially rival public goods that are also excludable such as recreational facilities, tennis clubs, swimming pools, and highways. International security is a good example of an excludable good characterized with rivalry, whereby “the overriding goal of each state is to maximize its share of world power, which means gaining power at the expense of other states.”\footnote{Alliances, on the other hand, illustrate a specific form of rivalry where states compete over maximizing the benefits from collective defense, but also engage minimally in burden-sharing to avoid any unnecessary costs. Therefore, by definition, alliances provide impure public goods or goods with mixed characteristics such as collective defense.}

The club goods framework holds that a direct relationship exists between the size of the membership and the ratio between costs and benefits. As membership expands, both costs and benefits arise which leads to crowding. Costs involve increased congestion due to a higher number of members, while benefits result from cost reductions due to economies of scale that lower the expenses associated with the club good. By adding a cost offset to the benefits derived from expanding the club size, crowding ultimately leads to finite memberships. In other words, every club has a certain capacity before it becomes congested. Congestion exists when costs surpass benefits. In the case of NATO, the issue of congestion was raised in several occasions since 1997 when the alliance invited new members with limited military capabilities to join when the Cold War ended.\footnote{The logic of the congestion argument has been challenged by international institutions scholars who make the case that the persistence of security organizations like NATO is determined by the ability of their specific and general assets, i.e.,, norms, rules, and procedures to match “the kinds of security problems faced by its members.”} The clubs also have important exclusionary mechanisms whereby users’ utilization rates can be monitored and non-members or non-payers can be barred. Without such mechanisms incentives for members to join and pay
dues would not exist. The operation of these mechanisms, however, must be at a reasonable cost, i.e., the costs should be lower than the benefits gained from allocating the shared goods. Also, the club goods framework distinguishes between two different types of costs and benefits: those related to the individual members and those associated with the club as a whole. When analyzing the clubs, it is necessary to distinguish between members with certain user privileges and non-members. In the case of NATO, after the Alliance launched its Partnership for Peace program (PfP) in 1994, most of the East European countries as well as the former Soviet Union joined the program but did not enjoy the privilege of common defense under Article Five nor did they enjoy participation in any substantial decision-making. By the same token, the members of the club need to determine clearly the amount of the shared good. In the case of NATO, the territorial outreach of the collective defense is limited to within the “territory of any of the Parties in Europe or North America” as defined in Article Six of the North Atlantic Treaty.\textsuperscript{77} Thus, while NATO remains committed to further developing existing, and open new individual relationships with nations outside of the North Atlantic Area that currently cooperate with the Alliance under the global partnerships initiative such as Australia, New Zealand and Japan, their prospects of joining the organization are very bleak.\textsuperscript{78} While most of these nations have clearly expressed that they “are not really interested in becoming members,” they have, on numerous instances, indicated a willingness to work with the alliance on its out-of-the-area missions and new allied capabilities, which are sometimes referred to as the “collective security business.”\textsuperscript{79} Therefore, under the current provisions of the North Atlantic Treaty the global partnership initiative incorporates participants with minimal prospects for actual membership rather than a sub-club of potential entrants.

**Homogeneous Clubs**

The club goods theory distinguishes between two types of clubs based on their membership structure—homogenous and heterogeneous. Homogenous clubs include members whose tastes and endowments are identical. If either tastes or endowments differ, then the club is called heterogeneous or mixed. In the economic and business world the majority of clubs are labeled homogenous. In international politics, homogeneity can usually be determined from a specific criterion, such as military power or economic wealth. Examples of relatively similar or homogeneous nations in terms of their wealth include the group of the most advanced industrial nations formerly known as G-7. In the realm of international security other examples of relatively homogeneous clubs include the Austro-German Alliance of 1879 and the Franco-Russian
Alliance of 1891–1894. In both cases these states possessed great power status as well as relatively similar shares of European military resources in terms of population, defense appropriations, or defense personnel.\textsuperscript{80}

The optimal membership size of these clubs is defined as the “core.” The core implies that no individual or set of individuals can improve their situation by forming a different partition.\textsuperscript{81} When the optimum club size is less than the actual members, i.e., when the club has more members than it needs to function effectively, strong incentives exist for club members outside of the core to switch to other undersized clubs. Therefore, there is a natural tendency for clubs to equalize their size toward the most optimal number of members. As a result, a multi-club world provides safeguards against discrimination because members that feel discriminated against have the option to switch between clubs.\textsuperscript{82} Modern alliances include a much larger group of states. Thus, homogeneity becomes an aberrant condition rather than a rule for these international bodies, especially because alliances aggregate the varying capabilities of the states that create them. For this reason, an analysis of contemporary alliance politics needs to take into account variations in the size and individual characteristics of the states.

**Heterogeneous Clubs**

In the case of homogeneous clubs, an evaluation of each member’s own payoffs is easier because club members are very similar and their contribution to the club equals everyone else’s. Studying heterogeneous clubs is, nonetheless, “considerably more elusive” because of variation among individual members.\textsuperscript{83} No uniform approach surfaces in the literature that suggests a method to evaluate whether one member contributes more or less than the others. Some theorists suggest partitioning the heterogeneous population into similar relatively homogeneous groups of members before the average net benefits of each member within the group and of the different groups are evaluated. The criteria to determine homogeneity vary and may include the allocation of resources for defense, the size of the allies, their positions on the level and scope of cooperation within the club, and whether these allies possess distinct military capabilities (e.g., airlift, rapid reaction and deployment, etc). Hence, the total net benefits to the heterogeneous clubs do not depend simply on the number of members but rather on the number and identity of the members (Table 1.3). Thus, the heterogeneous club model presents a flexible and yet adequate framework for studying NATO politics, whereby the alliance comprises several relatively homogenous sub-clubs. The model allows to draw comparisons within the various groups of relatively similar members, and between the different larger groups (or sub-clubs) of allies.
The analysis of club goods dynamic implies that if a club has a monopoly over a certain good, that monopoly automatically leads to sub-optimality. A single-club world may also be optimal if a low elasticity exists for demand of the good provided by the club for its members. From this perspective, NATO may be viewed as an optimal heterogeneous club only if the collective defense that it provides is regarded by some or all its members as an inferior or inelastic good, i.e., such a good that expenditure on it decreases or remains unchanged as income increases or at least the expenditure on the good increases by a smaller percentage than the income. Nonetheless, it is possible to have a situation in which there is only one club that establishes a monopoly and, yet, it is optimal. It is known in the literature as a noncompetitive provision and implies that the mere presence of a single club does not necessarily lead to a monopoly. In part, the monopoly effect can be reduced by other sub-clubs that provide a near substitute of the specific good.

The club goods theory, therefore, departs from the traditional condition for competitiveness in economic literature and argues that monopoly or oligopoly does not necessarily imply large club inefficiencies. Member-owned clubs are examples of club management that may limit the amount of rent that the monopoly or oligopoly can capture. Hence, understanding an optimal monopolistic or oligopolistic club like NATO requires a theoretical model

<table>
<thead>
<tr>
<th>Type of Club</th>
<th>Membership</th>
<th>Example</th>
<th>Relevance for Alliance Politics</th>
</tr>
</thead>
</table>
| Homogeneous clubs| Members with identical tastes and endowment    | Golf Clubs, G-7; the Austro-German Alliance (1879); the Franco-Russian Alliance (1891-94) | 1. The benefits depend on the number of the members.  
2. Multi-club world provides safeguards against discrimination. |
| Heterogeneous clubs | Members with different tastes and endowment | NATO; the Axis powers; the Western allies during World War II | 1. The benefits depend on the number and identity of the members (club “s core).  
2. Sub-clubs provide near substitute and reduce the monopoly effect. |
that somehow corresponds to a member-owned and operated club solution. Although in international security it is difficult to assume that any state is able to establish a monopoly over a certain club, the disparities in the capabilities between the United States and the rest of the NATO allies create a dynamic that may, under certain conditions, be analogous to a member-owned monopoly or oligopoly by the state that has an unmatched resource base and capabilities.

The heterogeneous clubs present a framework of analysis for alliance management in the presence of a non-competitive condition. This condition determines a special role of the United States in dealing with issues of transatlantic diplomacy and efforts to coordinate the outcome of intra-club bargaining on various issues. For example, the active involvement of the U.S. in ISAF has, on multiple occasions, influenced inter-allied bargaining and has resulted in enhancing the operation’s effectiveness.

The Team and the Club Concepts

The club dynamic bears a resemblance to another similar concept in the literature on transatlantic security and diplomacy—that of the teams. The underlying assumption in the team is that different players exist each possessing distinct roles and skills. Therefore, the team’s ability to functioning is based on specialization and interdependence, both of which rest on a combination of a common value set and a common goal. Specialization is the organizing principle of a team. The underlying logic is that players specialize to various degrees depending on the overall capability of the team. Interdependence, on the other hand, is the team’s modus operandi. Interdependence creates difficulty for community members to act independently. Alternatively, a lack of interdependence, allows for unilateral actions when common action is not optimal. The teams have a sense of inherent equilibrium in which the interlocking specialization between the United States and European countries automatically leads to “team entrapment” that appears attractive on both sides of the Atlantic.

The team concept, however, does not take into account several distinctive features of heterogeneous clubs. Despite the distinction between “superstars” and ordinary members of the team, a distinction that highlights the gap in individual members’ capabilities, the concept fails to account for diverging identities and patterns of intergovernmental bargaining among different members of the club. For example, the team assumes that the roles of the individual players are determined in advance and are not contested or challenged by any of team members. In reality intergovernmental bargaining often requires individual partners to adjust positions, which is a long and
burdensome process. The initiation of the 2004 NATO Training Mission in Iraq and the negotiations surrounding NATO’s contingency planning for the 2003 war in Iraq constitute two examples that illustrate the complexity of transatlantic dialogue. Therefore, the idea of the team reflects a desirable outcome of harmonious transatlantic cooperation rather than the actual process of accommodating difference among different allies.

The club goods model and the team concept share, nonetheless, a common understating about partnership being the organizing principle of cooperation. The idea of specialization embedded in the team and the club concepts assumes that specialization utilizes more efficiently their scarce resources in the area of security. Additionally, team logic rests on the comparative advantage theory’s argument that resources allocated to security can be used more cost-efficiently if countries form a team and specialize. Specialization is attractive even for countries with the same or lower levels of efficiency in production of all goods.

Although the clubs and the teams share similar assumptions, the heterogeneous clubs present a broader framework. For example, the criteria for homogeneity within the club varies and may include the size of the allies, the allocation of resources for defense (e.g., defense spending), the scope and size of military capabilities of the individual members (e.g., airlift, rapid reaction and deployment, etc), or their positions on the level and scope of cooperation within the club. The club suggests that the alliance outcome may or may not result in specialization, which depends on specific circumstances. The key focus, nonetheless, is on the relationship between the management of allied resources and the development of specific capabilities. Specialization is only one aspect of this relationship. Lastly, the problem with the team is that it undervalues the effect of bargaining among different European countries. The team vision, nonetheless, provide a valuable analysis about how to manage relations within smaller, relatively homogeneous units. Teams are also particularly useful in understanding the development of niche capabilities, such as multinational teams and force enablers, both of which will be discussed later in the book.

NATO’s Heterogeneity

Earlier studies implicitly identify some features of the heterogeneous club dynamic with regard to NATO politics such as the distinction between public goods and private benefits that the Alliance provides for its members, or the gap between national objectives and the objectives of the alliance as a whole. James Golden indicates that “there is no unity of objectives between national defense contributions and alliance defense contributions.” He suggests that,
in the absence of a compelling mechanism to convince the member states to place the common interest above the private interest, NATO leaders must search for bilateral political, economic, and security agreements among member nations. Golden’s study on NATO during the Cold War showed a causal link between the size of the allies, their wealth in terms of gross national income, and their spending on defense. This means that during most of its existence, NATO has been functioning as a heterogeneous club where the members’ resources for common defense vary significantly. The presence of different sub-clubs of allies with similar features based on their history, geography, and size has always been a distinctive feature of NATO politics, and different allies have contributed disproportionately more or less to the common defense.

Scholars on NATO admit that creating a measure and comparing national defense contributions has always been a major challenge. Consider, for example, the case of Iceland. The country has no defense forces nor defense budget but its key geographic position constitutes a critical element in alliance defense. Additionally, inflation and exchange rate fluctuations tend to overestimate or underestimate a nation’s resources. Currently this issue is less of a concern given the fact that as of January 1, 2011, twelve NATO allies have introduced a single European currency—the Euro—and seven others are expected to do the same within the next four to six years. As a part of this process, the new NATO allies and EU members have stabilized their currencies exchange rates. Although today fewer constraints exist in evaluating the national contributions of NATO members in terms of actual defense expenditure measured in national currencies or in terms of the share of gross national product, there is still no uniform approach in place about how to operationalize defense contributions. The size of member states is certainly important in determining their overall contribution. Alternatively, the elimination of the size factor allows studying the contribution to collective defense per capita, thus illustrating other aspects of the dynamic within NATO sub-clubs, and between them.

While previous research has explored the relationship between defense spending and national wealth in the context of the Cold War, testing the model for the expanded NATO has both theoretical and methodological implications. It provides an opportunity to compare the outcomes of the club dynamic in an enlarged NATO. Also, the model allows for exploration of the intra-club dynamic and the relations between, and among, various allies and sub-clubs. In Golden’s model, the base year is 1980 primarily because of data availability and reliability, while in this study, the base year is 2004 because that year marks the largest expansion in NATO’s history with seven new members. It is important to compare the correlation between defense
expenditure and national income in 2004 with the average values for the period between 1994 and 2004. The decade between 1994 and 2004 was probably the period of biggest change in NATO’s history that fundamentally transformed NATO politics: it began with the historical Brussels declaration adopted in 1994, followed by the initiation of the PIP program the same year, the new operations in Bosnia and Kosovo (1994–1999), the introduction of the Membership Action Plan (MAP), and ended with NATO’s involvement in Afghanistan and the big bang expansion in 2004. This dynamic is shown in Figures 1.1 and 1.3 where the defense expense and GDP indicators for each of the NATO countries have been averaged for the period between 1994 and 2004.

The relationship between defense expenditures and national income for 2004, as well as for the entire period between 1994 and 2004, shows several important tendencies. First, the correlations for 2004 and those for the period between 1994 and 2004 do not differ substantially. The data for 2004 highlight the overall tendency for the whole decade. Second, similar to 1980, the United States remains the highest defense spender in 2004 followed by Greece, Turkey, and the United Kingdom. These four countries clearly surpass the rest of NATO in terms of defense spending. Third, although the U.S. case is excluded from the graphs in Figures 1.1 and 1.2 for technical reasons,
the data show that the gap in defense spending between the U.S. and the rest of the allies has almost tripled since 1980. In the last decade alone this gap has soared about 20–30 percent because of the increased U.S. spending on the war on terror.

The per capita data for 1980 indicate that, aside from the United States, four major groups of allies could be identified. The first group includes allies with substantial military resources: the United Kingdom, France, and the Federal Republic of Germany. The second group comprises the poorer countries in Southern Europe with below average per capita incomes: Italy, Greece, Portugal, and Turkey. The last group includes several relatively small but rich states in the north: Belgium, Canada, Denmark, Luxembourg, the Netherlands, and Norway. The relationship between defense spending and per capita income for 2004 reveals that the picture has not changed substantially for the old NATO members as shown in Figure 1.4: three relatively poor but high defense spenders existed in the south—Turkey, Greece, and Portugal. The other two groups of allies remained relatively unchanged.
France, the United Kingdom, and Italy belonged to the group of major allies with significant military resources and strength. Canada, the Netherlands, Belgium, Denmark, Luxemburg, and Spain form another group of relatively rich countries with low defense spending. The only major difference from the 1980s in this respect is that in 2004 Germany dropped from the group of allies who maintained higher defense spending and moved to the group of small and wealthy West European nations who maintain low levels of spending. Alternatively, with a defense spending of about two percent Italy and Norway switched to the sub-club of big NATO allies with major resources and above average defense spending.

Aside from sheer size and defense spending, when it comes to the formation of sub-clubs, the identity of individual allies along the Atlanticist-Europeanist continuum also must be considered. France, Belgium, and Luxembourough are the staunchest Europeanists and they often seek cooperation with the moderate Europeanists in the south (Spain, Portugal, and Greece), while Germany, Italy and, to a certain degree, Norway are among the moderate Atlanticists. Alternatively, the United Kingdom, Denmark, and the Netherlands are strong Atlanticists seeking close cooperation with the United States and Canada across a large number of issues dealing with transatlantic security. Despite the fact that these identities are relatively stable, variations in individual members’ positions is possible as shown in the two case studies below.

One major difference from NATO’s dynamic in 1980, however, is the emergence of an entirely new group—that of the new allies from Central and Eastern Europe. They have two very similar characteristics: First, their national income is fairly low, comparable only to that of Greece and Turkey. Second, their defense spending (as a percentage of their gross national product) is slightly below or around the average of the alliance. Certainly variations exist within the group—some of the members’ populations constitute only as few as 1.3–2 million people (Estonia, Slovenia, and Latvia), while others like Romania and Poland have a population of about 22 and 38 million respectively. Hungary and the Czech Republic are closest to the median size of the ten new allies at ten million. At the same time, the structure of the armed forces and the pace of military reforms indicate a major degree of uniformity among those countries which is also captured in the relationship between income allocated to defense and national income per capita. Albeit, with certain variations, most of the new allies from Central and Eastern Europe have a strong Atlanticist identity that can be explained by their history, national strategic interests, as well as their “recent real-world experience.” Thus, based on these criteria, the new countries from Central and Eastern Europe emerged as one of the most coherent groups among all member-states (Figures 1.3 and 1.4).
Lastly, a noticeable trend among a number of European allies illustrated in Figures 1.3 and 1.4 is to be found in their increasing cuts on defense spending. For example, while in 2004 eleven of NATO’s twenty-six members met the official goal of spending 2 percent of their GDP on defense, in 2006 only seven allies (two of which are from Eastern Europe) met this goal. Thus, NATO as a whole recorded one of the lowest levels of defense spending of the individual members. These low levels led to immense frustration within the United States at the unwillingness of European allies to contribute to collective defense.

This survey of the relationship between national wealth and defense spending confirmed that the club goods theory is a useful framework through which to study NATO politics. The Alliance has always been a heterogeneous club that brings together diverse members. Based on the size of its members and the amount of resources that they are willing to allocate for common defense, at least several relatively homogeneous sub-clubs have always existed. The latest rounds of expansion in 1999, 2004, and in 2009 introduced a whole new sub-club of members from Central and Eastern Europe who share relatively similar characteristics. The variation of nations’ size and defense spending barely account for NATO’s heterogeneity. The differences of their national
positions require a great deal of compromises and accommodation before an actual decision is being made. Is such a decision optimal and what are the criteria for optimality? Does optimality entail solely ensuring or sustaining Alliance cohesion or does it also include effective implementation of desired goals?

The optimality of the decision-making can be evaluated from two different perspectives—as a process of making a collective decision or an outcome resulting from this decision. As discussed earlier, NATO’s decision-making process outlined in Article Five of the North Atlantic Treaty requires consensus among all allies. If the positions of individual members diverge and a compromise is not possible, the alliance cannot act and an arrangement can only be reached outside of the club. Such a scenario, however, excludes access to key NATO assets; it would not be endorsed by the alliance and would, therefore, hamper significantly the legitimacy of possible use of force. Despite the argument that fewer participants have a higher chance to reach a compromise, decisions negotiated within smaller groups of nations lack widespread
recognition in the international community. As a result, coalition-based decision making within issue-oriented sub-clubs, produces goods whose value is lower compared to consensus-based decision making that involves all club members. Most of NATO’s post-Cold War decisions to intervene in conflicts were preceded by an authorization of the United Nations Security Council in order to ensure legitimacy of international action and inclusive participation of all members; the use of force against Yugoslavia in 1999 being a notable exception in this respect. Heterogeneous institutions like NATO have some inherent virtues—not only do they provide geographic advantages, and access to current and potential military resources, but also they serve as a tool to legitimize multilateral action. Therefore, NATO needed to accommodate the diverging identities of its members to avoid possible unilateral action or formation of coalitions outside of NATO that would be, by default, sub-optimal.

Often times when a decision is put forward at the North Atlantic Council, the members have contrasting strategic preferences. Additionally, allies may have different perceptions of threat and may choose not to share the burden of costs and risks equally. Similarly, their dependency on allies and alliance structures may also vary—those who are more dependent on others “feel greater compulsion to defer to the wishes of their benefactors.” Therefore, when a proposal is put on NAC’s table, the allies can either approve it as it is; modify the proposal if one or several allies disagree or when an individual member disagrees but does not want to resort to its veto power; or it can abstain and then refuse to participate in the collective effort. Even though the latter scenario may influence adversely the size of allied contributions, it is often preferred to a possible use of veto. In theory, optimal decision making entails not only consensus building but also efficient management of allied resources and enhanced capabilities to conduct successfully out-of-the-area operations. Nonetheless, the process does not follow strictly this rule as illustrated in the case of NAC’s decision in 1997 to put on hold Slovenia’s bid for membership. Slovenia was considered a strong and well prepared candidate whose application was received well in Europe and in the U.S. Senate, but the Clinton administration decided that an invitation of three candidates that excluded Slovenia would be easier to accept internationally and at home. Despite some notable exceptions, for the most part consensual decision-making increases the likelihood to engage all members and hence improves organizational efficiency.

An assertion that reaching a consensus which takes into account national differences creates the best outcome may be problematic especially if optimality is measured in terms of outcome and overall impact of the decision for the Alliance. Consider for example NATO’s decision in the case of Kosovo in 1999. During the NAC deliberations it was important to find a mutually
acceptable decision and sustain the cohesion of the Alliance during the entire campaign against Yugoslavia. However, such lowest common denominator policy was criticized because it unveiled multiple operational constraints that curbed NATO’s overall effectiveness. The Alliance was able to successfully complete its mission primarily because Serbia was a fairly small albeit an obstinate adversary. In more difficult instances, however, a lowest common denominator policy is not always effective.

NATO’s involvement in Afghanistan is a poignant example in this respect. Soon after ISAF’s inception in 2002, the Alliance began to experience constant problems securing sufficient troop commitment from its members. Various operational problems after 2005 made multinational cooperation more difficult. With the exception of six contributing nations, most of the troops in Afghanistan operated under certain national caveats that imposed various constraints regarding military activities in which allied troops were allowed to participate while carrying out their missions. When NATO agreed to take over the mission in Afghanistan, with the exception of the United States, most of the allies were not equipped with capabilities to participate in counter-terrorism operations. Additionally, many participating nations imposed various caveats on their troop involvement to ensure themselves against a backdrop of domestic and international criticism. These decisions restricted significantly the scope of their involvement in combat operations as well as the capacity of their troops to respond adequately. Different groups of contributing nations emerged due to NATO’s heterogeneity: allies such as the United Kingdom, Denmark, Canada, and the Netherlands joined the United States in provinces with high-intensity conflicts, while others like Germany, France, and many of the new allies chose to stay in areas with low-intensity of combat operations where the death toll is much lower. Therefore, the United States stepped in and unilaterally provided much needed combat troops and capabilities. Some scholars saw this “crippling Americanization” of the conflict in Afghanistan as an indication that NATO has become an organization with marginal and “largely supportive” role in which some allies preferred to coordinate counter-terrorism efforts “through bilateral activities or loose coalitions.” Although a separate study is devoted to this case in chapter six, it is necessary to note that the lack of consensus and further commitment of the allies naturally reflects Alliance’s limited capabilities to meet the specific operational demands of the mission. Nonetheless, there is a major constraint from an academic standpoint to evaluate alternative explanations in the cases of Kosovo and Afghanistan in which a different format of international involvement, such as a unilateral action or a broad coalition, might have been more efficient. A study of a would-be unilateral or coalition-based action is simply not feasible methodologically.
RESEARCH METHOD

Despite the abundance of literature that grapples with different aspects of NATO politics, the literature lacks a comprehensive study that integrates all different aspects of NATO’s transformation in a comprehensive explanatory framework. This book attempts to fill in the existing theoretical vacuum and offer a holistic approach to studying NATO’s transformation in terms of new allies, missions, and capabilities. By focusing on the link between available resources and the development of new capabilities aimed at supporting new allied missions, this study answers the question of how NATO functions in the post-Cold War and how allied power is being managed in the course of its transformation.

The selection of methods and research procedure is constrained by the fact that NATO constitutes a *sui generis* political and military alliance. With its sixty years of existence, it has proven to be one the most sustainable, long-lasting alliances with very clear and unprecedented commitments embedded in Article Five of the North Atlantic Treaty. Not only did NATO incorporate a dozen new members after the Cold War, but it also undertook major tasks beyond its traditional collective defense responsibilities in places like Bosnia, Kosovo, the Mediterranean, Afghanistan, and Iraq. Therefore, the single case study may be the most appropriate method for several reasons. First, this particular case contributes to identifying common patterns of alliance politics and suggests broader generalizations often referred to as “building block studies” that pose tough tests for existing theories of alliance management. Developing a model that studies the management of military resources and allied capabilities poses such a tough case because it studies key ingredients of power politics in a new contextual environment in which NATO undergoes major transformation. Second, alternative methods, such as the comparative method, seem hardly applicable because they require at least two or more instances of alliance transformation that resemble each other in every respect but one. Instead, the study employs variable-oriented within-case analyses that seek causal relationship(s) of a specific variable by comparing how it performs in different instances within a particular case. These comparisons compensate for the limits of statistical and comparative case studies and are critical to the “viability of the small-n analysis.” Third, the study employs ‘process-tracing’ which identifies a causal path between two types of relationships: (a) the effect that the available resources (independent variables) have on the advancement of allied capabilities (the dependent variable); and (b) how the demand for new missions influences the outcome of these capabilities.
Hypotheses

The interconnectedness of the three patterns of NATO’s transformation—the new allies, missions, and capabilities will be explored through three separate hypotheses. The underlying logic is that after the end of the Cold War the Alliance has been able to transform itself by efficiently using the available resources of its member states in order to boost much needed allied capabilities. The three hypotheses are summarized in Table 1.4.

The first hypothesis studies the evidence for a relationship between military resources and specific allied capabilities. It is tested through a multi-method approach—quantitative study and a survey of NATO’s capabilities that are needed for peacekeeping, crisis prevention and response, and stability and reconstruction missions. The hypothesis is tested against three groups of nations: old members, new members, and four non-NATO European Union members (Austria, Finland, Ireland, and Sweden).

The second hypothesis explores to what extent and how the demand for new missions stimulates the advancement of new allied capabilities. This hypothesis is important because, if confirmed, it presents evidence for a causal relationship between the demand for new missions and the development of new capabilities. The hypothesis will be tested primarily through a qualitative approach.

Table 1.4. Hypotheses Tested

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Statement/ Proposition</th>
<th>Nations or Cases against which the hypothesis is tested</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis One</strong></td>
<td>There is a relationship between the available military resources and specific allied capabilities.</td>
<td>The old NATO members (Iceland excluded)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The new NATO members</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Four non-NATO members of the European Union</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peacekeeping Missions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crisis Management Missions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-proliferation Missions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and Stabilization Operations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>International Security Assistance Force in Afghanistan</td>
</tr>
<tr>
<td><strong>Hypothesis Two</strong></td>
<td>The demand for new out-of-the-area missions contributes to the development of specific allied capabilities.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hypothesis Three</strong></td>
<td>NATO membership stimulates the advancement of allied capabilities.</td>
<td>The old NATO allies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The new NATO allies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Four non-NATO members of the European Union</td>
</tr>
</tbody>
</table>
survey of NATO’s three core missions in post-Cold War international security—peacekeeping, crisis prevention, and reconstruction—and how they are related to the initiation and development of specific capabilities (e.g., combined joint task forces, rapid reaction forces, and multi-national battalions for chemical, nuclear, biological, and radiological defense).

The third hypothesis is related the first one and studies to what extent the process of expansion is related to the advancement of NATO’s new capabilities. The study surveys whether the prospect for membership has stimulated a more efficient use of available resources as well as development of new capabilities for countries that have recently joined NATO as opposed to the non-NATO members who are interested in joining NATO, as well as the non-members who do not plan to join the alliance in the near future. This hypothesis is tested primarily through a survey of patterns in the process of NATO expansion from the mid-1990s to early 2000s.

**Data Collection**

As indicated earlier, the research combines quantitative and qualitative data to survey the military transformation patterns of NATO members in the 1990s and 2000s. The goal of the mixed method approach is to test the validity of the argument from two different methodological perspectives and ultimately develop a theoretical explanation that rests on the logic of heterogeneous clubs and the concept of complementarities.

The quantitative component of the study identifies indicators that best represent the theoretical concepts intended to measure military power, peacekeeping, crisis prevention and response, and the capacity to stabilize the area of conflict. The quantitative method is also valuable toward identifying and operationalizing new variables and hypotheses, such as the relationship between military resources and allied capabilities. Most of the quantitative data have been collected from two main sources: the Military Balance (MB) Yearbook and Stockholm Peace Research Institute (SIPRI). Occasionally, due to the rotation of troops or the shift of missions, the MB database tends to underestimate or overestimate the actual values of the data through recording both the new and old troops or recording none. In order to avoid data spuriousness, the MB and the SIPRI data have been compared and, in cases of discrepancy between the two databases, the data has been checked against other alternative sources, such as national governments’ statistics, to confirm actual values.

The case study approach yields several advantages compared to the large-
feasible within a single case to use a variety of methods in order to test the consistency of a larger number of alternative explanations. For example, the qualitative part of the study includes a survey of primary sources (official NATO documents, military doctrines, press releases and leaders’ statements) and interviews with decision-makers. The interview component of the data collection provided a trade-off among theoretical parsimony, explanatory richness, a manageable number of cases, as well as generation of valuable original data. The semi-structured format of the interviews with open-ended questions gave freedom to probe and formulate follow-up questions. The interviewees were determined based on two criteria: First, they represent the national elites of NATO members, i.e., people in decision-making or leadership roles and; second, the respondents are experts about the topic at hand. Two types of elites meet these criteria: (a) government representatives, such as national delegations and permanent representations in Brussels, or the foreign and defense ministries of these countries; and, (b) representatives from non-governmental organizations and think tanks, such as the Brookings Institution, the Wilson Center, and the George C. Marshall Associations in Eastern Europe. The sample included over a dozen representatives of the political elites in the United States, Western Europe, and Eastern Europe. These interviews were conducted in 2006 in Brussels, Belgium; Sofia, Bulgaria, and Washington, D.C. The interviewees varied from scholars and researchers on U.S. foreign policy and transatlantic relations to high-ranking policy makers in the United States, diplomats from the national delegations at the NATO Headquarters, and think tank representatives.

There is little doubt that the focus on NATO’s post-Cold War transformation poses numerous methodological challenges. The single-case mixed methods approach may not be the only way to examine the patterns of NATO’s post-Cold War transformation, but it may be the most appropriate. Alternatively, a comparative study of several similar historical or contemporary cases would expand the theoretical findings. The problem of the comparative approach is that it is very hard to find historical or contemporary analogy to NATO’s scope, durability, and levels of commitment that could render basis for such a comparison.

NOTES

35. Christopher Layne argues that Mearsheimer’s argument is based on two separate theories because the stopping power of water is essentially a defensive realist concept. For further detail see Layne, *The Peace*, 19–20; also “The Debate Continues: American Hegemony in Theoretical and Historical Perspective: Christopher Layne’s The Peace of Illusions,” 47th International Studies Association Annual Convention, San Diego, California, March 22–25, 2006.
62. Barry Hughes calls the goods that exhibit excludability and non-rivalry *coordination goods*, while those goods that exhibit rivalry and non-excludability are referred to as *common property resources*, see Barry Hughes, “The EC and the Evolution,” 49.
64. For detail see Inis L Claude, Jr., *Power and International Relations* (New York: Random House 1962), 114.


69. For details about collective goods model see Olson and Zeckhauser, *An Economic Theory*, v.


72. Miroslav Nincic and Thomas Cusack explain the variation in U.S. defense spending during the Cold War with the perceived utility of such spending in stabilizing aggregate demand, the perceived economic effects arising out of such spending and the pressure of various institutional constituency-related factors. For details see Miroslav Nincic and Thomas Cusack, “The Political Economy of US Military Spending,” *Journal of Peace Research* 16, no. 2 (1979), 101–15.


75. Concerns about the utility of NATO expansion and skepticism about the incorporation of new members were first expressed by political scientists and policy makers in the mid-1990s. For further details see Michael Mandelbaum, “Preserving the New Peace,” *Foreign Affairs* (May/June 1995); Ted Galen Carpenter and Barbara Conry, *NATO Enlargement: Illusions and Reality* (Cato Institute, Washington D.C., 1998); and Zoltan Barany, *The Future of NATO Expansion* (Cambridge University Press, 2003).


80. Despite some variation in terms of population, energy consumption or defense personnel, the composite resource indexes of the individual allies were very similar. For details see Snyder, *Alliances*, 82–83; and Paul M. Kennedy, “The First World War and the International Power System,” *International Security* 9, no. 1 (Summer 1984), 7–41.
81. Cornes and Sandler, _The Theory of Externalities_, 199.
82. Waltz, _Theories_, 93.
87. Comparing the inflation rates during the 1990s is not as challenging for the old NATO allies as it is for the new NATO members. Although the economies in the old allies (with exception of Turkey) have shown a steady growth and low inflation, that has not been the case in Central and Eastern Europe, where the economies experienced high levels of inflation and substantial exchange rate fluctuations in the early and mid-1990s. As a result, there are significant fluctuations in the total defense expenditure, the national product and even the share of the national product spent on defense. Also, the inflation effect is strengthened by the major economic and financial crises that some of the Eastern European countries experienced during their transition in the 1990s such as Poland, the Czech Republic, and Bulgaria.
88. The convergence criteria (also known as the Maastricht criteria) determine the conditions under which the European Union member states can adopt the Euro. The countries in the Euro-zone need to meet four criteria: (a) the inflation rate should not exceed 1.5 percent of the three best-performing member states of the EU; (b) the government deficit must not exceed 3 percent of the gross domestic product (GDP) at the end of the preceding fiscal year and the ratio of gross government debt to GDP must not exceed 60 percent at the end of the preceding fiscal year; (c) the member states should not have devaluated their currency for two consecutive years and; (d) the long-term interest rate must not exceed 2 percent of the three best-performing member states in the Euro zone.
89. James Golden developed a matrix to evaluate national contributions in the context of alliance objectives and this article employs his methodology; see Golden, _The Dynamics_, 25–51.
91. Golden, _The Dynamics_, 44–45.
94. Slovenia is the only country with a noticeably high national income (about the same as Portugal) but still lower than NATO’s average. In the mid-2000s most
of the countries from Central and Eastern Europe spent below or around 2 percent on defense, Bulgaria and Romania being the only exception with about 2.4 percent.

95. These seven members were Bulgaria, France, Greece, Poland, Turkey, the U.K., and the U.S. Source: Stockholm International Peace Research Institute database, http://www.sipri.org/, 3/16/2009.


101. George and Bennett admit that such a control is very difficult to achieve and list various alternative methods such as Mill’s “method of agreement” and “method of difference,” Ragin’s Qualitative Comparative Analysis and King, Keohane and Verba’s “assumption of unit homogeneity.” For details on these approaches see George and Bennett, *Case Studies*, 153–71.