
**NOTICE**
This material may be protected by Copyright Law (Title 17 U.S. code)

---

**BALANCES AND BALANCING**

CONCEPTS, PROPOSITIONS, AND RESEARCH DESIGN

**Jacks Levy**

Over a decade ago, in the context of a review essay on the causes of war, I began a discussion of balance of power theory by writing that

the balance of power is one of the oldest concepts in the literature on international relations, but also one of the most ambiguous and least tractable. . . . The central concepts . . . of balance, power, equilibrium, stability, are rarely defined in any rigorous manner. . . . Ambiguity is increased further by the tendency to some to equate balance of power theory with realist theory or with any theory utilizing power as a central organizing concept. . . . There is no single balance of power theory, but instead a multiplicity of theories. . . . The confusion is all the greater because balance of power theorists cannot even agree on what it is they are trying to explain. (Levy 1989, 228-229)

Nearly all of these ambiguities in conceptualization and causal specification continue to plague the literature on the balance of power. There are still many variations of balance of power theory, many variations of realist international theory, and considerable confusion about how best to conceptualize the relationship between them. Although most balance of power theories are realist theories, not all realist theories are balance of power theories. In this paper I focus on balance of power theories and in particular on their hypotheses about balances and balancing, not on realist theory as a whole. I attempt to identify a common set of propositions shared by nearly all balance of power theorists, with the aim of developing some generally acceptable empirical tests that would help shift the debate from the level of abstract argument to the level of empirical evidence, a shift that is necessary if we are to resolve some long-standing debates about the empirical validity of balance of power theories. In the process I highlight a number of conceptual issues that must be resolved before hypotheses about balancing can be tested empirically, and I try to identify the scope conditions under which key balance of power propositions are most likely to hold.

On one level, finding some common ground should be an easy task. Balance of power theorists have long argued that states demonstrate a strong tendency toward balancing behavior, and that as a result it is extraordinarily rare for any one state to achieve a position of hegemony or dominance over the system. Rarely, however, do these theorists define what constitutes balancing, power, or hegemony, or specify who balances against whom, in response to what kinds of threats, in what systems.

More recent theorizing associated with the balance of power tradition has helped to refine some aspects of the argument and generate additional hypotheses. Walt (1987) questioned whether Waltz (1979) was right that states balance against the strongest power in the system, and argued instead that states balance against the greatest threats to their interests, defining threats as a product of perceived intentions, ideology, and distance, as well as capabilities. Others posed more fundamental questions about balancing and asked whether states balance against power threats or bandwagon with them (Waltz 1979; Labs 1992; Kaufman 1992; Walt 1992b), and Arquilla (1992) and Schroeder (1994) argued that bandwagoning with the strong could expedite their rise. Chubb (1989) argued that status quo states may balance to preserve their security while revisionist states often bandwagon with the strong in order to secure economic gains and otherwise expand their influence. Barnett (1996) meanwhile, suggests that states sometimes balance against threats to national identity rather than against more narrowly defined security threats.

Shifting the debate to a different level, Vasquez (Ch. 2, herein; 1997) examined the evolution of the neorealist research program on balancing behavior. He concluded that this aspect of neorealism was degenerating rather than progressive (Lakatos 1970) because it involved a series of ad hoc attempts to resolve theoretical and empirical anomalies in the balancing hypothesis without generating "novel facts" that could be tested over new empirical domains. Proponents of the balancing hypothesis responded to Vasquez's criticisms, with the debate focusing on the utility of Lakatosian metatheory as a normative standard for scientific progress, on the question of what counts as balancing, and on the historical record (Waltz 1977; Christiansen and Snyder 1997; Elman and Elman 1997; Schwebell 1997; Walt 1997; Chs. 3-7, herein). Still, there is no agreement on what might constitute a reasonable test of the balancing proposition, and certainly nothing that resembles a systematic (quantitative or comparative historical) empirical test.

The balancing/bandwagoning debate has potentially important implications for contemporary American foreign policy as well as for theory. If states tend to balance,
arguments about the need for the United States to intervene around the world to stop regional aggressors and falling dominoes are less compelling, because states in the region will form new coalitions to restore an equilibrium of power and thus increase the anticipated costs of conquest. If states tend to bandwagon, however, aggression can pay for itself by generating new resources and new allies, and external intervention becomes more imperative. Applied to the United States itself, some have argued that if balancing works, a new coalition will soon form in response to American hegemony in the world. If states tend to bandwagon, however, the United States can continue to extend its influence without much worry about generating a counterbalancing coalition; at the same time, however, it must anticipate that blocking coalitions will be slow to form against any rising military power in the future (Arquilla 1992). One of my primary arguments in this essay is that historically the leading states in the system have balanced against dominant continental states but not against dominant economic and maritime powers, and consequently we should not necessarily expect great power balancing against the United States in the early twenty-first century.

The question of whether or not states engage in balancing is very complex. There remains considerable confusion about what balancing is and what the theory predicts, and further conceptual clarification is necessary before ongoing theoretical debates can be fruitfully engaged at the empirical level. I am neither a committed proponent nor a committed critic of balance of power theory and its balancing hypotheses, but I think that it is imperative that debates about balancing be informed by the evidence. In this study I tackle some of the conceptual issues that plague the study of balancing, identify and refine some key hypotheses on balancing that are shared by most balance of power theorists, and identify the scope conditions for these hypotheses, with the aim of facilitating a shift in debates about balancing from the theoretical to the empirical level. Thus my focus is more on the level of theory specification and research design than on meta-theory, and I leave aside the question of the progressive or degenerative nature of the realist research program as a whole or of its theory of balancing in particular.

What Does Balance of Power Theory Predict?

This is not the place for a thorough review of balance of power theory (see Gulick 1955; Claude 1962; Morgenthau 1967; Aron 1973), but some background is necessary in order to identify the key propositions of the theory. There is no single balance of power theory but, instead, a variety of balance of power theories. Each begins with the hard core assumptions of realist theory—the key actors are states, who act rationally to maximize power or security under constraints in an anarchic international system—and then adds empirical content through the operationalization of power and other key concepts and through the specification of additional assumptions. As a result, different balance of power theories generate conflicting propositions about the actions and interactions of states. One example is the well-known debate about whether bipolar systems are more stable than multipolar systems (Waltz 1979; Deutsch and Singer 1969). Tests of the polarity hypotheses might constitute a test of a central proposition of Waltzian neorealism, but they do not necessarily bear on the validity of many classical balances of power theories.

Some balance of power theorists have argued that the purpose or function of a balance of power system is to maintain the peace (Wolfers 1962, Ch. 8; Claude 1962, 55). Let us leave aside the analytical problems involved in postulating the purposes or goals of a system, which not only attributes to systems the properties of individuals but also confounds state preferences with international outcomes that are the joint product of the behavior of two or more actors. The basic problem is that in most formulations of balance of power theory, states generally rank other goals higher than peace and concord of war as an acceptable instrument to advance those ends, if only as a last resort. Such goals include the maintenance of the independence of states (Gulick 1955; Organski 1958; Waltz 1979; Wagner 1986; Jervis 1997, 131), the avoidance of hegemony (Morgenthau 1967; Blaney 1973, 112; Sheehan 1996), or perhaps the general maintenance of the status quo (Kissinger 1964).

While balance of power theorists disagree over the relative importance of these goals, these objectives are in fact interrelated and can be conceived as a nested hierarchy of instrumental goals. The primary aim of all states is their own survival, defined in terms of some combination of autonomy and territorial integrity. There is a hierarchy of instrumental goals for achieving that end. The first is the avoidance of hegemony, a situation in which one state amasses so much power that it is able to dominate over the rest and thus put an end to the multistate system. There are several other goals that are seen as instrumental to the avoidance of hegemony. One instrumental goal is maintaining the independence of other states in the system, or at least the independence of the other great powers, which facilitates the formation of balancing coalitions against potential hegemons. Another is maintaining an approximately equal distribution of power in the system, defined in terms of some combination of individual state capabilities and the aggregation of state capabilities in coalitions.

None of this is meant to suggest that individual states aim to limit their own power or objectives for the sake of the system. They often seek to maximize their own power and/or security, but maintain an equal distribution of power among others; pursue whatever aims they choose, but induce others to pursue more limited aims; promote the status quo, but only if it serves their own interests. They engage in balancing strategies, not for the primary purpose of maintaining a balance, but because limiting the power of others is necessary to maintain their own security and independence. That is, the maintenance of the "system" is the unintended consequence of the actions of many states, each of whom attempts to maximize its own interests under existing constraints.

There are two causal paths leading to the absence of hegemony in balance of power theory, two reasons why the balancing mechanism almost always works successfully to avoid hegemony: (1) potential hegemons anticipate that expansionist behavior will lead to the formation of a military coalition against them, and refrain
from aggression for that reason; or (2) potential hegemons pursue expansionist policies and are defeated in war by a blocking coalition. The first results in peace but the second does not, so that the outbreak of war, even major war, cannot necessarily be taken as evidence against balance of power theory or the balancing hypothesis. Balancing hypotheses predict either state strategies of balancing or an outcome of balance (or nonhegemon), not peace (Levy 1994).

It is true that the power parity hypothesis predicts that an equality of power between two states is likely to lead to peace, or at least that parity is more likely than preponderance to lead to peace. But the power parity hypothesis is a dyadic-level hypothesis that assumes that alliances play no role, while balance of power theory is a systemic-level theory in which alliances are central and in which the outcome of any particular dyadic-level balance of power between two states is theoretically indeterminate. For this reason, Waltz's (1987) conception of bandwagoning to include the appeasement of threats in a dyadic relationship is an excessively broad conception of balancing, and Schweller (1994, 83) rightly criticizes Waltz for providing a theory of how states respond to threats rather than a theory of alliances, which is an integral part of balancing.

Another point of disagreement among balance of power theorists, noted above, is whether states balance against the strongest power in the system (Waltz 1979) or against the greatest threats to their interests (Waltz 1987). Without trying to resolve that debate here, let me emphasize that one thing both sides of this debate would agree on is that states, and particularly great powers, will balance against a state that threatens to achieve a position of hegemony over the system, because any state strong enough to threaten hegemony will in most cases be the greatest single threat to the interests of any other great power.

In emphasizing that balance of power theory generates predictions of both strategies of balancing and outcomes of balance, I acknowledge that some scholars, most notably Waltz (1979) and some other structuralists, argue that his neorealist theory predicts only outcomes, not state strategies or foreign policies. He predicts that balances of power naturally occur but leaves open the question of how they occur. For Waltz, outcomes of balanced power do not necessarily require deliberate balancing behavior by states.

It is important to note that Waltz is not always consistent in his argument that neorealism predicts balanced outcomes but not balancing behavior. At times he makes unambiguous statements about balancing behavior. He argues, for example, that “faced by unbalanced power, states try to increase their own strength or they ally with others to bring the international distribution of power into balance” (Waltz 1997, 915). He goes on to talk about “the reactions of other states to the drive for dominance of Charles I of Spain, of Louis XIV and Napoleon Bonaparte of France, of Wilhelm II and Adolph Hitler of Germany...” (Charles I became Charles V when he became emperor.) In his allusions to balancing behavior, however, one thing that Waltz does not emphasize is the specific motivations for that behavior or the degree of rationality underlying it. His arguments best fit Claude’s (1962) “automatic” conception of the operation of the balance of power.

It is clear that balanced outcomes and balancing strategies are analytically distinct, and that it is possible in principle to have one without the other. Certainly it is possible in principle to have balancing strategies that fail and end up in hegemonic outcomes. It is also possible in principle to have nonhegemonic outcomes without deliberate balancing strategies by states. This would be the result if no state had hegemonic ambitions. But if this outcome were the result of the fear of balancing, then balancing strategies play an important causal role even if states do not actually resort to them. Another possible explanation for the absence of hegemony might be that even in the absence of constraints no state prefers hegemony. While this is possible, it seems to run against the basic thrust of most realist theories. Even Kant argued that “It is the desire of every state, or of its ruler, to arrive at a condition of perpetual peace by conquering the whole world, if that were possible” (in Wight 1986, 144). Waltz is free to argue that balancing is not a necessary condition for balanced outcomes, but it is incumbent on him to specify the alternative causal mechanisms through which nonhegemonic outcomes repeatedly (or always) arise, and to do so in a way that is consistent with the basic assumptions of neorealist theory.

Specifying the causal mechanisms leading to balanced outcomes is important for another reason. A theory that predicts only that balances will form, or that multistate systems will not be transformed into universal empires, has far less empirical content, and can explain far less variation in the empirical world, than a theory that predicts state strategies as well as international outcomes. The loss of empirical content is particularly serious for neorealism, which generates relatively few testable propositions as it is (Keohane 1986c). In Lakatosian terms, a theory that predicts not only outcomes but the causal processes through which they occur is a scientific advance over one that predicts outcomes alone.

Because most balance of power theorists predict strategies of balancing as well as the absence of hegemony; because that is the stronger form of the theory (in terms of amount of empirical content, degree of falsifiability, and hence analytic power, but not necessarily empirical accuracy); and because the burden falls on those who do not hypothesize about balancing to specify the alternative mechanisms through which balances occur, I focus on versions of balance of power theory that predict strategies as well as outcomes.

Conceptual Problems in the Analysis of Balancing

I have suggested that the two propositions that nearly all balance of power theorists would accept is that (1) hegemonies do not form in multistate systems because (2) perceived threats of hegemony over the system generate balancing behavior by the other leading states in the system. States with expansionist ambitions are either deterred by the anticipation of blocking coalitions or beaten back by the formation of such coalitions. Most of the literature on balancing behavior focuses on the second proposition. Before hypotheses on balancing can be investigated empirically, however, we must be able to recognize balancing behavior when we see it. This raises a host
of analytical problems that have either not been acknowledged or not been resolved in the literature. These are related to the degree and timing of balancing, preventive war and other forms of strategic interaction between the balancer and the target of balancing, and endogeneity and case selection. Let us consider each in turn.

**The Degree and Timing of Balancing**

Waltz (1979) raised the question of whether states balance against stronger power or bandwagon with it. Most scholars have accepted this basic formulation of a dichotomy between balancing and nonbalancing, though some have specified numerous forms of nonbalancing behavior. These include the overlapping concepts of "shirk- ing," "bystanding," or "noninvolvement/distancing" (Arquilla 1992) or "hiding" (Schroeder 1994a), perhaps because of "buck-passing" (Christensen and Snyder 1990). Schroeder (1994a) also suggests that states sometimes attempt to "transend" threats by creating new international institutions and norms that will absorb or deflect those threats (see also Schroeder, Ch. 9, herein).

Critics have questioned Waltz's (1979) balancing hypothesis and argued that bandwagoning and other forms of nonbalancing are historically more common than balancing (Arquilla 1992; Schroeder 1994a). 26 This literature has been extremely useful in drawing attention to apparent historical violations of balancing hypotheses and in stimulating further research. Before we can make statements about the relative frequencies of balancing and nonbalancing behavior, however, we need rigorous operational criteria for distinguishing between the two, because balancing comes in different degrees and forms. The literature, which has to date been primarily qualitative in orientation, has yet to address this question of how to draw the line between balancing and nonbalancing behavior for the purposes of classifying behavior.

Two important issues concern the kind of behavior that constitutes balancing and the timing of balancing. Both proponents and critics of balancing agree that Britain balanced against Germany in World War I because Britain entered the war fairly quickly against Germany. But in the years leading up to the war, and even during the July crisis, Britain failed to make an unequivocal commitment to intervene in defense of France in the event of a German attack. This failure may have contributed in significant ways to the German and Austrian decisions for war (Frischer 1967). 21 Similarly, Britain engaged in appeasement in the late 1930s but eventually entered the war. As Vasquez asks (1997, 907; Ch. 2 herein, 37), "Were Britain, France, and the USSR passing the buck in the late 1930s, or were they just slow to balance?" Similar questions can be raised about the behavior of the United States, which intervened in each of the world wars but which made no effort to balance early in an attempt to deter the aggressor.

Intervention against the stronger state after the outbreak of war should not be excluded as a form of balancing. We know in retrospect that the course of certain wars led to overwhelming military victories by one side and subsequent intervention by third parties to limit or roll back the extent of those victories. Military outcomes are not always anticipated in advance, however, and intrawar balancing to maintain or restore equilibrium is an important form of balancing. Indeed, intrawar balancing may be more common than prewar balancing.

Still, the phenomenon of intrawar balancing leads to the question of how late a state might be in entering the war and still be classified as engaging in balancing behavior? Would we still classify Britain as a balancer if it had not entered the war in early August 1914 but instead waited until German armies had occupied Paris? Would we classify the United States as a balancer if it had not intervened until Hitler's armies had conquered nearly all of Europe? It seems to me that late balancing is still balancing, but it suggests the need to distinguish the fact of balancing from the timing of balancing, and to include the latter as a separate variable.

There are two qualifications. First, to qualify as balancing, military intervention must be directed against the strongest power or the greatest threat. It would not be balancing if a third party enters the war against the expansionist state only after it has become clear that the aggressor would be defeated in war (the Soviet Union's late entry into the Pacific War against Japan in August 1945, for example).

Second, it would not be balancing if war is forced on the potential balancer by a direct military attack by the aggressor. As Schweller (1994, 83) argues, balancing involves a situation in which "a state is not directly menaced by a predatory state but decides to balance against it anyway to protect its long-term security interests." During the Napoleonic Wars, for example, the great powers of Europe generally did not balance against France, and when they did resist it was only because they were attacked by France (Schroeder 1994a, 135; Rosecrance and Lo 1996) Similarly, Stalin was forced into the war with Hitler when German armies invaded the Soviet Union, and Vasquez (Ch. 2, herein) is correct to argue that this should not be classified as balancing.

Questions also arise as to whether military alliances in peacetime (against stronger states) constitute balancing. Presumably formal, written defense pacts calling on all signatories to intervene in the event of a third party attack on any of them would be balancing, but nonaggression, neutrality, and entente pacts (as defined by the Correlates of War Project [Singer and Small 1968]) would not. The fact that Britain and France closely coordinated naval planning in the decade prior to the war and structured their respective war plans around those agreements (Williamson 1969) is a reminder, however, that balancing comes in degrees and that we need a separate dimension for the degree of balancing.

There are also questions relating to the intentions and motivations of the balancer and the perceptions of targets. States do not always live up to their commitments, and states sometimes intervene despite the absence of a formal alliance. Would it be balancing if a state signs a defense pact directed against an expansionist state but has no intention of honoring its commitment in the event of war? What if that alliance successfully deterred an external attack? Would it be balancing if a state does not form a formal alliance but has every intention of intervening in a war that would put an expansionist state in an unacceptably strong position (Britain and Germany in World War I, for example)? There is a wide range of ways in which states signal their intentions with respect to intervention (Morrow 2000), and we need to incorporate this fact into our conception of balancing.
Another question is how to classify behavior involving a combination of strategies, some of which may be consistent with balancing but others inconsistent. How do we aggregate these into an overall judgment on balancing? Churchill (1948) writes that the strategy of appeasement Germany was not necessarily a mistake as long as it was combined with one of the time purchased by appeasement to undertake a major military buildup in preparation for the inevitable war. In fact, Carr (1939) called on England to pursue such a strategy, and Christensen and Snyder (1997, 921; Ch. 5 herein, 69-70) use this to argue that appeasement and balancing are not necessarily alternative strategies but in fact may be complementary ones. This might be true in principle, but whether it was true for Britain prior to World War II is open to question, as Britain made little effort to prepare for war while it attempted to appease Hitler (Watt 1989). Britain pursued a strategy of appeasement and, only after it failed, shifted to a strategy of balancing.

The Soviet Union adopted nearly the reverse strategy. It first tried to balance against Germany, and in fact went to great efforts to form an anti-German alliance with France and Britain. When Stalin could not come to terms with the Western powers, however, he reversed course and bandwagoned with Hitler by agreeing to the Nazi-Soviet Pact (Read and Fisher 1988). We talk about state strategies of balancing, and that makes sense in the context of intervention in an ongoing war, but prewar balancing through the formation of defensive alliances involves the actions of two or more states, not a single state. Intentions of balancing do not always materialize.

As this discussion and various historical examples and counterfactual possibilities make clear, balancing is a multidimensional concept that cannot easily be captured by a simple dichotomy of balancing and nonbalancing. While this dichotomy may be useful for some theoretical purposes, for many other purposes it is too limiting, and we need to think in terms of a range of balancing behavior that is characterized by different dimensions and degrees. We should distinguish the scale or magnitude of balancing (absent, weak, moderate, strong), the timing of balancing (early, late), and perhaps also the extent of balancing, defined in terms of the number or proportion of states (or of great powers) that engage in balancing behavior. There was great-power balancing against Revolutionary and Napoleonic France and against Nazi Germany, but in these wars states balanced at different times and with different intensities; some states switched from balancing to bandwagoning; and some states did not balance at all. We need, for many theoretical purposes, a more nuanced set of categories to help capture this variation in behavior.

PREVENTIVE WAR

In the last section I noted that Vasquez and others argue that military resistance against an expansionist state does not count as balancing if it is a response to a direct attack. I accept this argument, but it raises an interesting analytical question and more difficult historical issue that we need to deal with in classifying balancing in particular cases. The interactive nature of decisions for war requires that we look not only at the behavior of the aggressor as well as that of the hypothesized balancer, but also at how the aggressor attacks in the first place. It may be that the stronger state attacks the weaker one not because of blatantly expansionist motives, but because it anticipates balancing behavior by the defender and decides to take preemptive or preventive action (Levy 1987).

Frederick the Great initiated an attack on Austria in 1756, for example, but this is widely interpreted as a preventive war in anticipation of the formation of an offensive alliance and war against Prussia, which Frederick knew was planned for the following spring (Anderson 1966, 34; Dorn 1940, 312-314). If this interpretation is correct, then we would conclude that Austria and its allies intended to balance and Frederick acted in anticipation of balancing, so that balancing played a major role although Frederick initiated the war by attacking Austria.

The possibility of preventive or preemptive attack in response to the anticipation of balancing raises a very difficult problem for the analyst who wants to differentiate between balancing and a response to military attack. It forces the analyst to make a judgment about the motivations of the attacker, and consequently to engage the historical literature on this question. This is rarely conclusive, given the conflicting interpretations of historians, but some judgments are more defensible than others.

Consider Vasquez's (1997) argument that the Soviet Union did not balance but was attacked by Germany in 1941. An alternative explanation is that Hitler assumed that Stalin would eventually enter the war against Germany, recognized that German power relative to the Soviet Union was peaking in 1941, feared that within two years Stalin would be a much more potent adversary, particularly in the context of a two-front war, and took advantage of a closing window of opportunity to invade the Soviet Union (Copeland 2000). Balancing plays a role in the second interpretation but not in the first. While one can find some evidence in support of the preventive war interpretation, there is enough written on the role of Lebensraum in Hitler's foreign policy and on planning for eastward expansion to suggest to me that Hitler planned this all along (Weinberg 1994). On this basis I would tentatively conclude that the preventive motivation affected only the timing of Hitler's move against Russia, that this was an attack by an expansionist state, and consequently that the Soviet Union did not engage in balancing in summer 1941.

ENDOGENEITY AND THE PROBLEM OF CASE SELECTION

The possibility of preemptive or preventive action by an expansionist state against a prospective balancer raises an interesting problem of interpretation. If A attacks B in the anticipation of balancing by B, then balancing has a causal impact even though we do not actually observe B's balancing behavior. We encountered this problem of unobserved balancing before in our discussion of the two causal paths that lead to the absence of hegemony in realist theory: (1) restraint by potential hegemons because they anticipate that aggression would lead to a blocking coalition and a costly and possibly fatal war; and (2) the lack of restraint, aggressive expansion, the formation of a blocking coalition, and (usually) the defeat of the expansionist state. Balancing is central to each of these causal paths, but is observed only in the second. Behavior that is off the equilibrium path has no less a causal impact than behavior that we observe.
The possibility of unobserved balancing raises some important questions of empirical case selection, questions that have been mostly ignored in the literature (but see Jervis 1997, 137–139) and that seriously complicate the assessment of the relative frequency and causal importance of balancing. The problem with nearly all of the recent literature that tries to bring historical evidence to bear on the balancing proposition is that it focuses on the wars that have occurred and asks whether states balance or not. It does not look at the wars that have not occurred and ask whether the absence of war might be due to the anticipation of balancing by the potential aggressor. By ignoring strategic behavior by the potential aggressor and its impact on what we observe, existing studies of balancing systematically underestimate the causal importance of balancing in international politics.

It is conceivable that scholars' observations of a substantial number of cases in which states fail to balance against strong and/or threatening states reflects the possibility that balancing behavior is not particularly common. But the possibility of strategic selection leads to an alternative explanation. If states frequently balance in international politics, but if there is some uncertainty as to whether states balance and under what conditions and in response to what kinds of threats, and if potential aggressors recognize this, aspiring hegemons will engage in expansionist behavior only in those situations in which they have some reason to believe that their potential enemies will not balance or will balance ineffectively. These are the only cases that we observe, and consequently it is not surprising that we find a fair number of cases of nonbalancing. This is a biased sample, however, and it results in a systematic underestimation of the causal importance of balancing in international politics.

Analytic problems like these are now familiar in the literature, as researchers have become increasingly sensitized to problems of endogeneity and strategic selection (which makes it all the more surprising that this problem has not been given much attention in the literature on balancing). Consider the question of the relative success or failure of economic sanctions. Most of the empirical literature finds that economic sanctions frequently fail to accomplish their objective (Hufbauer, Schott, and Elliot 1990). The problem is that, until recently, the literature generally looked at sanctions and asked whether they succeeded or failed, but ignored the problem of strategic selection. If targets can anticipate that threatened sanctions would be successful, they will alter their behavior before sanctions are actually applied. Consequently, successful sanctions will seldom be observed. If, on the other hand, targets anticipate that sanctions will not be effective or that the cost of compliance would be too great, targets will not alter their behavior, sanctions will then be applied, and they will frequently fail (Morgan and Miers 1999).

There is thus a close parallel between the study of the relative success and failure of economic sanctions and the relative frequency of balancing and nonbalancing in international politics. Analyses of economic sanctions based on cases in which sanctions are actually applied will systematically underestimate the causal impact of sanctions by ignoring cases in which sanctions succeed by inducing “anticipatory compliance” with expected sanctions. Analyses of balancing that are based on cases in which balancing actually occurs will systematically underestimate the causal impact of balancing by ignoring cases in which balancing works because potential aggressors’ anticipation of balancing leads them to strategically alter their behavior.

While there is no doubt that an empirical study of balancing that focuses on observable wars or grand coalitions underestimates the causal importance of balancing by ignoring unobserved balancing, what is not clear is the magnitude of this bias or how to estimate it. The basic problem is that the relevant cases are nonevents resulting from the potential aggressor's calculated restraint. It is hard to identify what this population of nonevents looks like or what criteria we should use for the selection of cases, because the potential population of such nonevents is extraordinarily large. One thing that is clear, however, is that the common approach of selecting cases in which wars occur and then examining the presence or absence of balancing behavior in those wars is a form of “selection on the dependent variable” that has the potential to generate very biased results (King, Keohane, and Verba 1994).

It is far preferable to focus on the independent variable, the conditions or actions that are hypothesized to lead to balancing. Proponents of Waltz's theory, which hypothesizes that states balance against power, should examine cases of high concentrations of military power and analyze how frequently states balance in response to that power. Those who hypothesize that states balance against threats rather than against power should devise a way of measuring the intensity of threats, generate a database, and then see if balance of threat hypotheses are consistent with the evidence. Because balance of power theory and balance of threat theory lead to similar predictions if one state begins to approach a position of military dominance in the system, selecting all such cases (or a sample from them) would provide the basis for a test of both sets of hypotheses.

The Scope Conditions of Balance of Power Theory

Our earlier review of the literature on balance of power theory suggests that despite the many different versions of balance of power theories, there are two core propositions that most balance of power theorists would accept: (1) if a state becomes strong enough to threaten to gain a position of hegemony over the system, other leading states in the system will balance against it, and consequently (2) hegemonies will rarely if ever occur. I think that most critics of balance of power theory would accept these propositions as correct inferences from most versions of balance of power theory. This brings us closer to an empirical test that both proponents and critics of balance of power theory would accept as a “fair” test of the theory. There are some additional ambiguities in the meaning of the central concepts of states, system, hegemony, and power, however, and further clarification and operationalization of these concepts are necessary before we can think seriously about testing balancing propositions empirically.

First consider the domain of the proposition that states balance against hegemonic threats. Balance of power theorists speak very loosely about “states” balancing, but implicit in most versions of balance of power theory is a strong sense that it is the
great powers that are expected (descriptively and perhaps normatively as well) to do the balancing. While both great powers and lesser states prefer that the power of an aspiring hegemon be limited, only the former are strong enough to make a difference. The weak know that they can have only a marginal impact on outcomes, and given their vulnerability and short-term time horizons, they will sometimes balance and sometimes bandwagon, depending on the context.

The great power bias in the balancing proposition is reflected in the long tradition of balance of power theory. As Claude (1989, 78) argues, "balance of power theory is concerned mainly with the rivalries and clashes of great powers and—a whole host of . . . world wars." This bias is shared by most (but not all) traditional realist theories and most diplomatic histories (Levy 1989). Waltz (1979, 72–73) is explicit in saying that any theory of international politics must necessarily be based on the great powers, for they define the context for others as well as for themselves. Formal theorists create stylized balance of power models consisting of just a handful of actors (Wagner 1986; Niou, Ondesheek, and Rose 1989; Powell 1996a). Theories of hegemonic decline, power transition, and hegemonic war clearly focus on the causes and consequences of the behavior of the leading powers in the system (Organski 1968; Gilpin 1981; Kennedy 1987; Thompson 1988; Rasler and Thompson 1994).

The various concepts of balance of power or equilibrium in the literature generally refer to a balance or equilibrium among the great powers, not among states in general. Many balance of power theorists, including those involved in debates over the relative stability of bipolar and multipolar systems, use stability to mean the absence of war, by which they mean the absence of war between the great powers, not the absence of war in general. Balance of power theorists do not regard wars between the weak, or between the strong and the weak, as constituting instability in the system. On the contrary, some balance of power theorists argue that imperial wars on the "periphery" of the system serve as a "safety valve" and play a stabilizing role.

This conception of stability in terms of the absence of major war rather than the persistence of key structural features of the system is reflected in the fact that most balance of power theorists regard the periods of 1618–1659, 1672–1714, 1914–1918 as extremely unstable precisely because of the severity of general wars in those periods, even though the multipolar distribution of power in the system or other key structural properties of the system did not change. The collapse of the Soviet Union, the end of bipolarity, and the emergence of a single "hyperpower" was a fundamental structural change around 1989, but this process is not viewed as constituting instability because it was not accompanied by a major war.

Some recent formulations and applications of balance of power theory, and the balancing hypothesis in particular, have focused on smaller states in regional systems (Waltz 1987, 1992). We need to be careful, however, in applying balance of power hypotheses drawn from the experience of the great powers to regional systems in which the assumptions underlying balance of power theory might not fully hold. In particular, the assumption of anarchy—the absence of any mechanism for enforcing agreements—is not fully satisfied in regional systems, where smaller states can sometimes appeal to actors outside of the system. Great powers in the larger system do not have this option. As a result, behavior in systems that are nested within larger systems may differ from behavior within the larger system.

Under some conditions, a state might choose to forego balancing against the strongest regional power (or greatest regional threat) if its survival and security are guaranteed by an external great power. How often this occurs, and under what conditions, remains to be theorized. In terms of outcomes, while hegemonies rarely form in larger great power systems, they occasionally form in local systems, often because such local hegemonies are not contrary to the interests of more powerful states outside the region. This does not mean that Walt (1997) and others are wrong to apply balance of power theory to regional systems, only that scholars must be very careful in doing so and must acknowledge the different structural contexts and their possible consequences.

Thus when most balance of power theorists talk about states balancing power, there is generally an unstated assumption that it is the most powerful states in the system (which might include the strongest regional states in regional systems) who do the balancing, not states in general. In addition, while balance of power theorists disagree about exactly what it is that great powers balance against, they agree that at a minimum the great powers balance against states that threaten to achieve a position of hegemony in the system.

This hypothesis is not yet fully specified because the concepts of hegemony and system remain ambiguous. As I have argued elsewhere (Levy 1985, 1994), a key limitation of most hypotheses about polarity and the balance of power more generally is that scholars have failed to identify (1) the system under consideration and (2) the basis of power in that system. Until this is done, we cannot identify hegemonies or hegemonic threats and consequently we cannot specify testable propositions about balancing behavior.

A reading of balance of power theorists who have had the greatest impact on the international relations literature over the past several decades (for example, Morgenthau 1967; Gillick 1975; Claude 1962) suggests that there is a strong Eurocentric orientation in balance of power scholarship. Most of the literature is written by Europeans (especially the British) and Americans (whose security outlook was primarily Eurocentric until the late twentieth century). This literature focuses on the modern European great power system going back to 1648 or perhaps to 1494/5 or the Italian city-state system (Mattingly 1955), and until the twentieth century the leading powers in that system have all been European. Even if we were to focus on the "modern world system" over the last five centuries, the "system leaders" have all been European (Wallerstein 1974; Thompson 1988).

There are numerous manifestations of the Eurocentric bias in balance of power theory. The concept of a "balancer" or "holder of the balance," while it can be generalized in principle, is nearly always equated with Britain's role in maintaining an equilibrium of power on the European continent by shifting its weight to the side of the weaker coalition (Dehio 1962; Morgenthau 1967). The hypothesis that the system tends to be most stable if there is a colonial frontier into which the dominant actors in the system can expand their power and influence without directly
threatening the vital interests of other dominant actors also reflects a Eurocentric orientation (Morgenthau 1967, Ch. 14; Gulick 1955, Ch. 1; Hoffmann 1968; Wright 1965, Ch. 20).

Even the central claim that the balance of power mechanism works effectively to prevent hegemonies from forming (or the weaker claim that for whatever reason hegemonies have not formed) reflects the European experience, which has been characterized by the continuation of a multistate system and the absence of a sustained period of dominance of a single great power over the European continent over the last millennium.

While Waltz (1979) and many other balance of power theorists imply that the same patterns of balancing and balance hold in any system of two or more states, this was not true for the Chinese multistate system at the end of the Spring and Autumn period and Warring States period.34 At this time a succession of military conquests transformed the multistate system in China into universal empire (within China) under the hegemony of the Qin Dynasty by 221 B.C.35

This raises the interesting questions of why a universal empire was established in China but not in modern Europe (Hui 2001), whether the European or the Chinese experience is the more common one (and thus whether the European experience is relatively unique), and the extent to which balance of power theories arising from the European tradition are applicable to (or are valid in) other historical systems. The answers to these questions have important implications for the extent to which key balance of power propositions can claim to be universal in their applicability, or whether even the most powerful of those propositions (hegemonies do not form) are themselves temporally and spatially bounded.

It is also clear that most international relations theorists and historians writing about the balance of power have implicitly conceived of hegemony in terms of dominance over the European system. For most of the literature on the balance of power, the foundation of potential hegemonic power is land-based military power based on large armies, supported in recent times by armor. Thus balance of power theorists talk about balancing coalitions against the Habsburgs under Charles V in the early sixteenth century, Philip II at the end of the sixteenth century, and the combined strength of Spain and the Holy Roman Empire in the Thirty Years War; against France under Louis XIV and then Napoleon; and against Germany under Wilhelm and then Hitler (Gulick 1955; Dehio 1962; Claude 1962; Hoffmann 1967; Aron 1973; Kennedy 1987). It is revealing that even Waltz (1997, 914; Ch. 3 herein, 54), who speaks in broader theoretical terms about balancing and balances in the international system as a whole, illustrates his arguments with examples of balancing against Charles I, Louis XIV, Napoleon, Wilhelm II, and Hitler.

Although it is common to refer to the Pax Britannica of the nineteenth century and to treat Britain as a hegemon or leader during much of that period,36 that is a view associated with hegemonic stability theory (Keohane 1984), which generally conceives of power in terms of financial and commercial strength; power transition theory (Organski 1968), which operationally defines power in terms of gross national product; or leadership long cycle theory (Thompson 1988), which defines power in terms of naval capability and dominance in leading economic sectors. This is not to say that global economic power is unimportant, only that it is not a central concern of balance of power theories. One of the most important trends in the international system over the last five hundred years is globalization, but as Sheehan (1996, 115) notes, "the voluminous balance of power literature is almost completely silent on this issue."

Balance of power theorists, on the contrary, conceive of power in terms of military power and potential, and their identification of the leading threats to hegemony over the last five centuries (Charles V, Philip II, Napoleon, Wilhelm II, and Hitler) make it clear that military power is land-based military power. For balance of power theorists, it was Germany, not Britain, that was the leading power in the "system" by the end of the nineteenth century.37 If Britain was a hegemon, it was a hegemon in finance, trade, and naval power on a global scale, not a military hegemon over Europe. Similarly, to the extent that the United States was a hegemon after World War II, it was because of American dominance in the world economy; in terms of military power, the situation was one of parity between the United States and the Soviet Union.38

From a balance of power perspective, it is not surprising that a blocking coalition did not form against Britain at the peak of its global economic and naval strength in the 1870s, but that one did form against Germany, the leading power on the continent, in the period leading up to World War I. Nor is it surprising that a great power balancing coalition formed in the late 1940s against the Soviet Union, the primary military threat to the major states of Europe, and not against the United States, by far the leading power in the world in terms of economic strength and naval air power.39 Similar logic explains why a balancing coalition did not form against the Netherlands in the seventeenth century despite its dominance in world trade, finance, and (until midcentury) naval strength (Israel 1989), but instead against Louis XIV and his massive armies (Levy 1985, 368).40

Maritime powers, as compared to continental powers, have fewer capabilities for imposing their will on major continental states; fewer incentives for doing so, and a greater range of strategies for increasing their influence by other means. One explanation for the lower levels of threats posed by maritime powers, at least to other great powers, the majority of which have been European over the last five centuries, derives from the fact that effective military power significantly diminishes over distance, especially over water (Mearsheimer 2001).41 Large armies massing on borders, threatening to mass on borders, or simply having the potential to mass on borders threaten the territorial integrity of other states in a way that strong naval power or financial strength does not. Whereas European hegemons threaten their neighbors by virtue of their very existence, maritime hegemons do not.

Another explanation for the lower threat posed by maritime powers derives from the recent literature on territory and international conflict, which suggests that a disproportionately high number of wars involve territorially contiguous states; that unsettled territorial disputes are an important predictor of war; and that rivalries are significantly more likely to escalate to war if they involve territorial disputes (Vasquez 1993, 1996b; Huth 1996; Hensel 2000; Vasquez and Hensel 2001). The
absence of territorial contiguity removes both a direct path for conquest and a source of many of the disputes that escalate to war, and hence removes an important source of threat.

Maritime powers differ from continental states in their interests as well as their capabilities, and those interests lead to different strategies. The goals of increasing commercial, financial, and naval power on a global scale do not require military or political control on the continent, and this leaves them less threatening to the major European states. Maritime powers may impose their will on smaller states and other actors, but that is not directly relevant for testing balancing hypotheses or balance of power theories more generally. Even with respect to small states, however, maritime powers often exert their influence through means other than overt military force. The “imperialism of free trade” (Robinson and Gallagher 1953) was as potent as military force in establishing British dominance in far corners of the globe.

Thus global maritime powers define their interests on a global scale and have fewer incentives to expand their influence in Europe, appear less threatening to European great powers, and are consequently less likely to trigger balancing coalitions. This does not imply that they have no stake in what happens on the European continent. They often perceive that their overall interests require that no single state achieve a hegemonic position in Europe, for fear that such a position might provide the resources for mounting a challenge to the dominance of the leading global maritime power on a worldwide scale. Thus the leading maritime power often plays a central role in balancing coalitions against potential European hegemons. It is not an accident that the global leader in economic and naval power plays the role of the “balancer” in balance of power theory.

There is another reason why most balance of power theorists have implicitly conceived of the absence of hegemony as the absence of European hegemony and of the balance of power as the balance of military power on the continent rather than the balance of naval and economic power on a global scale. That is because much of the literature that we read on the balance of power is British or American in national origin. Britain has long defined its interests in terms of pursuing a balance of power on the continent but a preponderance of naval and colonial power, and the restriction of the balance of power concept to the continent provided a useful rationalization for those interests. The same can be said for American interests in American global dominance and a balance of power in Europe.

The Vienna settlement is often interpreted in terms of the balance of power (Gulick 1955; Kissinger 1964), but the system emerging from Vienna constrained France and possibly Russia while doing nothing to limit British naval or colonial power. As Bullen (1980, 15) notes, “the concept of the balance of power was hardly ever used except by British governments. The continental powers certainly did not consciously seek to uphold it.” To the contrary, continental statesmen and scholars were often quite skeptical of the concept and the uses to which it was put. As the Duc de Choiseul pointed out, “the English, while pretending to protect the balance on land which no one threatens, are entirely destroying the balance at sea which no one defends” (quoted in Sheehan 1996, 115).42

The European orientation inherent in balance of power theories, as well as their potentially self-serving basis, is recognized by Sheehan (1996, 115):

The balance of power concept for some 200 years after its confirmation as the basis of the European state system remained a purely European phenomenon. Its logic was not applied beyond the boundaries of the European continent. This may have been because the strongest proponent of the theory, Britain, had the most to lose from such a development. It may also have been related to the fact that the European balance of power idea was, in terms of its origins, part of a peculiarly European solution to the problems afflicting the European imagination.43

Although there is no necessary link between state policy interests and the scholarship conducted within its borders, it is no coincidence that scholars from the global maritime powers have generally emphasized a balance of power on the European continent while remaining silent on balance of naval, colonial, or economic power, while many on the continent advocate a balance of power on the seas and in the colonies as well as on land. This need not adversely affect the scientific evaluation of balance of power hypotheses, however, if we accept the view that the logic of confirmation is distinct from the logic of discovery (Popper 1959). Nevertheless, a sensitivity to the origins of an intellectual tradition can alert the researcher to the normative biases and scope conditions underlying those propositions and facilitate efforts to minimize the effects of those biases in empirical tests.

The great power, Eurocentric bias of most balance of power theories has important implications for American global hegemony and reactions to it in the early twenty-first century. Zakaria (2001) notes the puzzle of “why is no one ganging up against the United States? Throughout modern history countries have regularly resisted a rising global power. The world mobilized against Napoleon’s France, imperial and Nazi Germany, the Soviet Union.” There is an inconsistency between Zakaria’s theoretical statement about responses to global power and his historical examples, which all deal with responses to land-based military powers. In terms of the implicit analytic and perhaps normative biases of balance of power theory, the absence of balancing against the United States is not a puzzle at all.

Nor is the absence of balancing against the United States a puzzle from the perspective of hegemonic stability theory, power transition theory, leadership long cycle theory, and other hegemonic perspectives, which emphasize the tendency for coalitions to form around the leading state to create a stable global political economy and security system—assuming the leading state is committed to a liberal international economy. Far from being a source of threat, liberal leaders can be the source of reassurance and stability. Thus Charles Maier (1987) speaks of “consensual hegemony,” and Geir Lundestad (1986) speaks of an American “empire by invitation.” From this perspective it is possible to have a “benign hegemon” (Friedman 1999, cited in Waltz 1999, 694). A balance of power theorist would never say this (but see Schweller and Wohlfarth 2000, 74). For them, hegemonic powers are threatening not because of what they do but for what they are capable of doing. From a
balance of power perspective, therefore, the failure of a grand coalition to form against the United States, if it continues, should not be taken as disconfirming evidence for balance of power theory's central proposition that great powers balance against hegemonic threats, just as the absence of balancing against Britain in the nineteenth century or the Dutch in the seventeenth century should not be taken as evidence against the balancing hypothesis.\(^\text{44}\)

It is true that Waltz (1997, 915–916; Ch. 3 herein, 54; 1999) argues that the world will begin to balance against American hegemony, and implies that the absence of such balancing would constitute evidence against his neorealist balance of power theory.\(^\text{45}\) Waltz is to be commended for specifying with such precision the kinds of evidence that he would take as falsifying his theory (though he does not make a similar argument about Britain in the nineteenth century), because the more precise we are in specifying what observable behavior and outcomes would lead us to abandon our theories, the greater the analytic power of those theories. Regardless of how Waltz wants to treat neorealism, however, any failure to balance against the United States should not be taken as evidence contrary to other formulations of balance of power theory, most of which have a strong Eurocentric orientation.

I have spoken of the implicit biases in balance of power theories, but let me frame this in a slightly different way. Waltz (1979) argues that his neorealist theory and its key propositions about the stability of multistate systems and the strong tendencies for hegemonies not to form are derived from a very limited number of assumptions: the existence of an anarchic system of two or more states who wish to survive. My argument is that Waltz's formulation of neorealist theory's balancing proposition as a universal law is probably incorrect. Waltz's theory—and indeed all theories of balancing behavior and balanced outcomes—is much more plausible if we add a scope condition that limits the domain of the theory to continental systems. Whether hypotheses of balancing and balances are valid for all continental systems, and not just for the European system over the last five centuries, is an important question that requires much more attention in the literature.

**Conclusions**

The question of whether states balance or bandwagon is extraordinarily and deceptively complex. There is considerable confusion about what balancing is and what it is not, who balances against whom and how, how to identify balancing when it occurs, and the conditions under which propositions about balancing and balances are expected to hold. The critics of balance of power theory and its balancing proposition have made an important contribution by demonstrating that historically states do not always balance against the strongest states in the system or the greatest threats to their interests, but sometimes adopt other strategies. They have convinced me that if the question is whether states nearly always balance, the answer, based on the historical record, is almost certainly no. But if the question is whether states balance more frequently than they do not, or the conditions under which states balance, then we need a much more refined conceptualization of balancing and a method for measuring different dimensions of balancing before we can come up with an empirically based answer to that question. One of my primary objectives in this study has been to add some conceptual clarity to debates about balancing.

My other objective, not unrelated to the first, has been to identify some key propositions that are shared by nearly all balance of power theorists, in order to help structure an empirical test that would be accepted by most balance of power theorists and their critics as a good test of the theory. I have argued that the core propositions shared by most balance of power theorists is that the emergence of any single state with the potential to dominate the international system will generate a blocking coalition of other great powers, and that consequently hegemonies over the system will not form. I have also argued that it is necessary to qualify these propositions by specifying their scope conditions: Balancing propositions apply to continental systems, and particularly to the European state system over the last five centuries, but not necessarily to global systems. We know that the European multistate system has not been transformed into a single empire, which leaves the question of whether European great powers have systematically adopted balancing strategies whenever a single state amasses a disproportionately large share of power in the system.

In terms of the logic of comparative inquiry, this is the "most likely" case for the balancing proposition (Przeworski and Teune 1970; Meckstroth 1975). If the balancing proposition does not hold here, our confidence in its more general validity would be significantly diminished, and in this sense this constitutes a good test of the balancing hypothesis.\(^\text{46}\) Evidence that such balancing occurs would provide some support for the theory, but supporting evidence would increase our confidence in the theory less than disconfirming evidence would diminish it, given the asymmetry in most-likely research designs. Because the United States is a global maritime power rather than a continental power, it falls beyond the scope conditions of the theory, and the failure of a general coalition to form against the United States in the twenty-first century would not necessarily contradict the assumptions underlying most formulations of balance of power theory.

**Editors' Commentary**

**Major Query**

Which hypothesis or hypotheses would best assess balance of power theory, and how should it or they be tested?

**Major Points**

Levy attempts to identify the key problems that arise when trying to test propositions associated with balance of power theory, to provide solutions to these problems, and ultimately to come up with an empirical test.
He begins by pointing out that there is no one balance of power theory, but several. Accordingly, a prerequisite for testing is to determine what each theory predicts and what common propositions they share. Once this is done, it is still necessary to set out explicit criteria for determining when balancing and nonbalancing occur, so that it can be observed. Each of these tasks, which seem simple enough, turns out to be more complicated than meets the eye.

In terms of what balance of power theories predict, Levy documents that early theorists, going back at least to the Italian Renaissance, originally saw a balance of power system as a way of maintaining peace (see also Bueno de Mesquita, Ch. 12, herein). The problem, however, is that other balance of power theorists posit that states will go to war to balance power and/or prevent hegemony, and so they deny the connection between balance of power and peace. Still another disagreement is whether states balance against power (i.e., the strongest state in the system) or against threat. Then there are dyadic versions of the theory, which Levy sets aside because some see alliances as playing no role at this level. However, Levy argues that despite these differences "most balance of power theorists predict strategies of balancing as well as the absence of hegemony." The absence of hegemonic outcomes and the tendency for threats of hegemony to generate balancing behavior are the two key tests that most balance of power theorists should accept.

However, there is still the problem of determining when balancing behavior exists. Levy discusses several major problems. One of the more important is the timing of balancing. How late can balancing occur in order for it still to be seen as balancing? Recall in the previous chapter, Schroeder sees the failure of states to choose an initial strategy of balancing as evidence against Waltz. Levy agrees with Vasquez and Schroeder, as well as Schweller, that responses to a direct attack are not balancing. Nevertheless, he is prepared to admit some cases of late balancing as evidence in favor of the theory: for example, balancing that occurs during a war would count if it involves balancing against the stronger state or if it limits or rolls back the victories of that state.

Levy also analyses several more subtle questions, such as whether only certain kinds of alliances can be considered balancing (defense pacts for sure, but not nonaggression pacts or ententes); whether the intentions of states are relevant to the analysis of observed or unobserved balancing; and the role of backup strategies, or pursuing a combination of strategies such as using appeasement to buy time in order to balance more effectively later.

In addressing these questions, all of which are nicely illustrated by historical examples, Levy makes the point that balancing is a multidimensional concept that cannot be treated as a dichotomy (balancing or nonbalancing). This implies that a proper test will need to measure different aspects of balancing. These include: the scale or magnitude of balancing (whether it is absent, weak, moderate, or strong), the timing of balancing (when it occurs—early or late); and the extent of balancing (the number or proportion of great powers that balance).

Such measurement will go a long way in helping scholars marshal evidence for and against the balancing proposition. However, there are two instances where such measurement alone does not get at important evidence—preventive war and unobserved balancing. Levy points out that it is analytically possible that an expansionist state might initiate a preventive attack before this balancing behavior can be actualized. Here, there would be no observable evidence of balancing on the part of other states, but the anticipation of balancing nonetheless affected the initiator's behavior. According to Levy, these cases of "dogs that do not bark" should count in favor of balance of power theory.

Preventive war is a special case of this more general problem of unobserved balancing. Levy maintains that the most effective form of balancing may be when a potential aggressor abandons expansionist goals in anticipation of balancing. Such unobserved cases never make it into the test sample, yet they constitute relevant evidence. A sample of cases that leaves out these noncases can underestimate the evidence consistent with the proposition. Problems due to unobserved balancing are difficult to handle statistically, since they involve intentions and counterfactuals. Often they have to be discussed on a case-by-case basis, and Levy's analysis illustrates the kind of specific detailed discussion of historical cases that will need to take place to resolve these sorts of problems.

Levy concludes his chapter with a discussion of the scope of balance of power theory; i.e., to what domain it is applicable and to what it is not applicable. Based on the literature, he maintains that it applies to great powers and that it is especially relevant to Europe. In a more original vein, he goes on to argue that the balance of power propositions apply to military rather than economic power, and specifically to concentrations of land rather than naval forces. This leads him to conclude that it is primarily the attempt to prevent a hegemonic threat to a continental system containing major states that is the most natural domain for balance of power theory (which means for the modern state system—the European continental system). He points out that states did not balance against British maritime and economic hegemony in the nineteenth century or against Dutch maritime and economic hegemony in the seventeenth century, and he argues that this is not inconsistent with the logic of balance of power theory. He also engages current debates about responses to American hegemony, and using the implicit assumptions of traditional balance of power theory, he argues that the absence of great power balancing against the United States today is not surprising.

For Levy, a global naval power does not pose a threat to European land powers and therefore is less likely to evoke a balancing response. He states that Waltz's formulation of the balancing-of-power proposition as a universal law requiring only anarchy and a desire to survive "is probably incorrect," but a balance of power theory with a narrower scope might not be.

Levy suggests that most proponents and critics of balance of power agree that the theory predicts balancing against hegemonic threats in Europe (Waltz would predict balanced outcomes but does not insist that it is necessarily balancing foreign policy behavior that leads to those outcomes). If the main propositions of balance of power theory do not apply to the European continent, then it is unclear where they would apply. The European continent is the "most likely" case where it
should hold. If it did not pass such a test, the credibility of balance of power theory would be severely undermined. Conversely, if it passed it would provide only limited evidence in favor of the theory, because it would be regarded as an easy test. In sum, Levy succeeds in stipulating a test that balance of power theory must pass, if it is to be taken seriously.

**KEY TERMS**

*most likely* case  A case in which a theory is expected to hold; therefore, an easy test for the theory to pass. Failure to pass an easy test would lead many scholars to conclude that the theory is false and should be rejected.

non-events  Events that do not result in observable behaviors when measured using the theory's concepts and operationalized variables. For example, a would-be hegemon may anticipate a countervailing balancing coalition, and so moderate its actions. The case will appear to contain neither expansion nor balancing, when in fact it provides strong support for balance of power theory.

**NOTES**

2. Realist international theory includes both hegemonic realism and balance of power realism, which generate diametrically opposed predictions about the consequences of extreme concentrations or fragmentations of power in the system (Levy 1994, 2002a). Balance of power realism includes both classical theories as reformulated by Morgenthau (1948), Gillett (1955), and Claude (1962), and the more systematic structural realism of Waltz (1979). Hegemonic realism includes power transition theories, hegemonic stability theory (Gilpin 1981; Keohane 1984), and leadership long cycle theories (Thompson 1988; Rasler and Thompson 1994). This distinction is obscured by standard classifications of realist theory that focus on the differences between various forms of classical realism and contemporary structural realism (Frankel 1996a; Doyle 1997; Schwebler and Priess 1997; Rose 1998; Brooks 1997). Waltz’s (1979) neorealist propositions about the balance of power, for example, have far more in common with classical realist balance of power theories than with Gilpin’s (1981) structural realism.
3. This is an important theme in Vasquez’s introductory essay in this volume. Distinguishing between balance of power theory and realist “theory” is a necessary step toward the construction and testing of falsifiable theories, because the contrasting predictions of different versions of realism mean that nearly any behavior or outcome is consistent with some version of that paradigm, leaving realism as a whole nearly immune to empirical falsification.
4. Alliance formation in pursuit of economic gain is not limited to revisionist states, as Barnett and Levy (1991) demonstrate in their study of the domestic political economy of alliance formation in the Third World.
5. This raises the question of whether there are strategies by which the United States might diffuse any potential opposition and prevent blocking coalitions from materializing. For a theoretical discussion of how leading states construct an institutional order that helps to co-opt their potential adversaries, see Ikenberry 2001.
6. For my own views of the utility of Lakatosian metatheory for the analysis of international relations, and for an application to the research program on power transitions, see DiCicco and Levy 2002.
7. On the varied uses of the balance of power concept, see Hass 1993, Claude 1962, and Sheehan 1996. I treat the balance of power as a theory of state behavior and international outcomes in anarchic systems consisting of two or more states. The use of balance of power in a descriptive sense to refer to a particular distribution of power in the system (an equal balance, a favorable balance, or any distribution) is confusing, and it is better to use the term distribution of power. The concept of a "balance of power system" generates additional conceptual baggage and provides no value-added over a view of the balance of power as a theory of behavior that includes variations in system structure as key independent variables. Systems are not real; they are analytical constructions that theorists use to describe and explain reality.
8. It might be more accurate to say that most balance of power theories are sets of discrete hypotheses rather than logically interconnected theories. More formal attempts to model balance of power theory include Wagner 1986, Niiou, Orsenigo, and Rose 1989, and Powell 1996a. For a formalized argument that key predictions of Waltzian neorealism cannot be derived from the theory's assumptions, see Bueno de Mesquita’s essay in this volume.
9. This goes back at least to Giovanni Botero, who wrote in 1589 that balances produce peace and that by balancing power Lorenzo de Medici had kept Italy at peace for many years (Sheehan 1996, 33).
10. States might prefer peace over war, yet find themselves locked in a structural dilemma (such as a prisoners’ dilemma) that leads to war.
11. The term independence of states is ambiguous in balance of power theory. It is necessary to distinguish between one’s own state, other great powers, and other states. Maintaining the independence of one’s own state is an irreducible national value. Maintaining the independence of other great powers is a means to that end, not an end in itself. Maintaining the independence of weaker states in the international system as a whole is not generally regarded by most balance of power theorists as being among the highest priority goals of states, particularly of great powers. In traditional formulations of balance of power theory, for example, the partitioning of weak states is often mentioned as a possible strategy for maintaining equilibrium in the system (Gillett 1955).
12. This is Claude’s (1962) "automatic" conception of the balance of power, formalized by Waltz (1979); see also Jervis 1997, 132. An alternative view of the balance of power argues that a stable balance of power system requires some degree of restraint by individual actors (Gillett 1955, 33; Kaplan 1987) or a "moral consensus" (Morgenthau 1967, 208–213) as to the legitimacy of the system. For a more systematic argument regarding the role of norms of restraint and conceptions of self-interest in terms of the security of the broader community, see Schroeder 1993.
13. Waltz’s (1979) neorealist theory also predicts that bipolar systems are more stable than multipolar systems, but Waltz defines stability in terms of the persistence of key structural characteristics of the system (anarchy, number of major actors), not the relative frequency of war or peace.
14. Most quantitative empirical research demonstrates that at the dyadic level peace is associated with a preponderance of power (Kugler and Lemke 1996), not a parity of power.
15. Waltz (1979), Mearsheimer (2001), and most "offensive realists" argue that because under conditions of anarchy security is scarce and the intentions of others are always uncertain and subject to change, power is inherently threatening, there are no benign hegemons, and consequently states balance against power. Waltz (1987) and most "defensive realists" argue that anarchy does not necessarily imply insecurity, that states judge security threats in terms of others’ intentions as well as capabilities, and that consequently they balance against the greatest perceived threat rather than against the greatest power. On the distinction between offensive and defensive realism, see Frankel 1996a and Levy 2002a.
16. Arguments over whether or not states consciously and deliberately balance power is reminiscent of Claude’s (1962) distinction between automatic, semi-automactic, or mutual conceptions of balance of power systems. Balances form without the conscious calculations of states in the first conception, only with conscious and calculating balancing strategies by all states in the third conception, and with deliberate balancing by one state (the "balancer") in the second conception.
17. This is what game theorists refer to as behavior "off the equilibrium path."
18. On the debate over whether neorealism incorporates a theory of foreign policy, see Elman 1996 and Waltz 1996.
19. I use deterrence in the broad sense of an actor being dissuaded from taking an action for fear that the expected costs and risks of that action exceed the expected benefits, either because of successful defense on the battlefield or, particularly in the nuclear age, threats of punishment away from the battlefield. While the deterrence concept was not commonly used before the nuclear era, political leaders and scholars would have clearly understood its meaning.
20. Similar to Schwebler (1994, 99) argues that contemporary realist theory has underestimated the frequency of bandwagoning.
21. Thus Britain's failure to balance against Germany before August 1914 contributed to the outbreak of the war while its balancing after the outbreak of war helped avoid a German hegemony over Europe.

22. It must also be a substantial, not token, intervention. The context is different, but would we say that Western states "balanced" against the Bolshevik threat in the Russian Civil War, given the limited nature of their intervention?

23. Similarly, Louis XIV declared war against the Holy Roman Emperor in 1688, beginning the War of the League of Augsburg. This war involved balancing, however, because a defensive coalition had already formed and Louis may have acted preventively out of a defensive strategy to maintain the status quo in Europe rather than an offensive strategy to expand his influence (Lynn 1999).

24. An argument that Stalin had planned an aggressive war against Nazi Germany but was preempted by the German invasion is made by Suvorov (1990) and persuasively rebutted by Gorodetsky (1999).

25. But remember that the Soviet Union tried to balance earlier but could not reach an agreement with the West on the terms of an alliance.

26. This is similar to the problem of general deterrence. Did American military might and the NATO alliance successfully deter the Soviet Union from invading Europe, or did the Soviets have no intentions of invading in the first place?

27. A secondary hypothesis might be that higher concentrations of power lead to stronger balancing, though this is more difficult to analyze because many instances of strong concentration of power will be unobserved because balancing will take place at lower levels of concentration.

28. The difficulty of identifying and measuring threat perception independently of the behavior that is hypothesized to follow from those threats is one reason why it is far easier to test Waltz's (1979) hypothesis of balancing against power than Waltz's (1987) hypothesis of balancing against threats. Schroeder (1994a) goes further and argues that balance of threat hypotheses have little meaning because they cannot be empirically falsified.

29. Recall that individual states seek the absence of hegemony not as an end in itself but rather as a means to self-preservation.

30. The majority of Western diplomatic historians have followed Leopold Von Ranke (1833/1973) in conceiving European history as the history of great power relations. A. J. P. Taylor (1954, xxvii), for example, argues that "the relations of the great powers have determined the history of Europe."

31. Waltz (1979) is in the minority in defining stability in terms of the persistence of key structural characteristics of the system.

32. This is one basis for the common argument that in any given period wars are either frequent but limited, or infrequent but serious, but not both frequent and serious (Morgan and Levy 1990).

33. Similarly, Jervis (1997, 133) notes that "although the overall balance of power system has never failed, local ones have."

34. The fact that the multistate system was transformed, for a time, into a hegemonic system does not necessarily imply that states did not adopt balancing strategies. The absence of balance does not necessarily imply the absence of balancing.

35. Though this lasted less than two decades, a universal empire was reestablished under the Han Dynasty.

36. For different assessments of the influence of British leadership in the nineteenth century, see Ingram 2001 and Thompson 2001.

37. As I have argued elsewhere, balance of power theories and "hegemonic realist" theories generate diametrically opposed predictions regarding the consequences of various distributions of power, which is right at the heart of realism (Levy 1994, 2002a). Whereas balance of power theories see hegemony as ture, hegemonic theories see it as quite common. Whereas balance of power theories suggest that the probability of a major war is quite low when concentrations of power are high, hegemonic theories predict that the probability of a major war is quite low under conditions of high concentrations of power. Whereas balance of power theories imply that the probability of a major war is relatively low when there is a fairly equal distribution of power among the individual great powers and among great power coalitions in the system, hegemonic realists see that as the point of the highest probability of a major war, when the declining leader is being overtaken by the rising challenger. For an argument that balance of power theory and power transition theory are complementary, see Schmitter and Wohlforth 2000, 73).

38. My arguments about the differences between the European and global systems emerged from my critique of Thompson's (1988) leadership long cycle theory and its primary focus on the global system. In part because of the influence of Thompson's work, however, I now emphasize the interactions between the European and global systems.

39. There is another sense in which the NATO alliance is consistent with my argument regarding the fear of concentrations of continental power and the absence of fear of concentrations of global power. In the often-quoted words of Lord Ismay, the purpose of NATO was to "keep the Russians out, the Germans down, and the Americans in."

40. There are enough cases of balancing against naval power to suggest the desirability of more research on the theoretical conditions under which this is most likely to occur and how these differ, if at all, from balancing against land-based power on the continent. France, Spain, and the Netherlands entered the War of the American Revolution against Britain, for example.

41. This "loss of strength gradient" (Boulding 1953) is to a certain extent a function of technology, but for much of the past five centuries of the modern system even if a nation's bodies of water significantly diminished the ability to project military power. Palmerston's mid-century observation that before the steam engine, the English channel was "impassible by a military force" (quoted in Brodie 1941, 49) is quite revealing, and most observers agree that Palmerston overstated the impact of steam on British vulnerability to a cross-channel invasion. Even today analysts debate whether China has the military capability of mounting a successful military conquest of a much weaker Taiwan.

42. The argument that the territorial balance of power in Europe should be combined with a balance of colonial and maritime power was made by the Germans as well as the French. In fact, Sheehan (1996, 137) argues that German writers made a distinctive contribution to balance of power thinking through their willingness to discuss the concept in global terms, which, he argues, the British rarely did. Of course, the German argument was also a convenient rationalization for their repeated charges that their own efforts to acquire territories and influence beyond Europe were being blocked by Britain and other global powers.

43. While one can find occasional reference to balance and the balance of power going back to ancient Greece, Sheehan (1996, Ch. 2) argues that the systematic use of the concept originated in the intellectual uncertainty and crisis of authority that followed the Renaissance and Reformations that led to a search for intellectual order and harmony.

44. For another perspective on the absence of balancing against the United States, see Brooks and Wohlforth (2002, 24–27).

45. Keegly (1993) argues that the emergence of a multipolar system from the present unipolar system would constitute vindication of realist theory. Waltz (1999, 915) agrees, and goes on to say that "to all but the myopic . . . [multipolarity] can already be seen on the horizon." This is consistent with Layne's (1993b) argument that new great powers always rise in response to unipolarity.

46. Whereas least-likely research designs are based on what I call the "Sinatra inference" (if I can make it there I can make it anywhere), most-likely research designs are based on the inverse Sinatra inference—if I cannot make it there I cannot make it anywhere (Levy 2002b).
REFERENCES

References


References


