A Model Disagreement | Stephen M. Walt

My purpose in writing "Rigor or Rigor Mortis?" was to evaluate the contributions of formal rational choice theory to the field of security studies.¹ I argued that formal theory was useful—but not essential—for developing precise and logically sound arguments, and suggested that the benefits of formalization were not cost-free. I also argued that recent formal work had not produced a significant body of new and original insights, and I sought to show that much of this work was either untested or empirically questionable. Accordingly, I concluded that although formal theory could be a valuable part of the field, it was not intrinsically superior to other well-established research techniques. As a result, I emphasized that the field of security studies should strive to maintain its methodological pluralism. To paraphrase Georges Clemenceau, the study of warfare is too important to be left solely to formal modelers.

The five responses to my article raise many important issues. Lacking sufficient space to address all of them, I focus here on the central points that divide us. I do not believe that the responses cast serious doubt on my original claims, and as I attempt to show below, several of them actually provide additional support for my position.

My reply consists of five sections. The first section considers the issue of logical consistency and precision, which several of the respondents declare to be the most important feature of a scientific theory and the cardinal virtue of formal techniques. The second section examines the question of creativity and originality and shows why the examples of innovative work offered by my critics do not undermine my original assessment. The third section revisits the issue of empirical validity and shows that the responses actually lend further support to my central argument. The fourth section addresses the crucial issue of policy relevance, which is still a major liability of formal work in the field of security studies. The fifth section considers the hegemonic aspirations of the modeling community and reiterates my plea for methodological pluralism.

Stephen M. Walt is the Evron M. and Jeane J. Kirkpatrick Professor of International Affairs at the John F. Kennedy School of Government, Harvard University.

^{1.} Stephen M. Walt, "Rigor or Rigor Mortis? Rational Choice and Security Studies," *International Security*, Vol. 23, No. 4 (Spring 1999), pp. 5–48. Further references appear parenthetically in the text.

International Security, Vol. 24, No. 2 (Fall 1999), pp. 115-130

^{© 1999} by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

Logical Consistency and Precision

The responses to my article make two main claims on the subject of logical consistency and precision. The first claim is that I place little value on this criterion, thereby stacking the deck against formal modeling.² The second claim is that logical consistency is the sine qua non of any scientific theory, a claim intended to demonstrate the intrinsic superiority of formal techniques. The first charge is false; the second merits additional discussion.

Contrary to the first assertion, I do not denigrate logical consistency or precision. As I wrote in my article, social science "requires theories that are logically consistent, precise, original, and empirically valid." I also declared that "other things being equal, theories that are stated precisely and that are internally consistent are preferable to theories that are vague or partly contradictory," adding that "logical consistency is highly desirable and efforts to achieve it are a central aim of science." And I went to some lengths to point out that this was an area where formalization could make a contribution (pp. 8, 12, 17). Like motherhood and apple pie, in short, logical consistency is something that all of us endorse.

Where we differ is in the relative importance of this criterion and the relative performance of formal and nonformal approaches. Several respondents assert that logical consistency is the most important criterion for judging a social science theory. This view is clearest in the response by Bruce Bueno de Mesquita and James Morrow, who write that "logical consistency takes precedence over [creativity and empirical validity]," adding that it enjoys "pride of place among the criteria for judging social science theories."³ I disagree. Although consistency and precision are valuable, they are not the only—or even the most important—qualities that scientists look for in a theory. And as I noted in my article, formalization is neither necessary to achieve consistency nor sufficient to guarantee useful results.

First, logical consistency alone is essentially meaningless, for one can derive any conclusion one wishes if one begins with the right set of premises. It is not surprising, for example, that formal theorists often reach logically consistent but contrary conclusions, and logic alone cannot tell us which one is

See Lisa L. Martin, "The Contributions of Rational Choice: A Defense of Pluralism," p. 82 n. 10; Emerson M.S. Niou and Peter C. Ordeshook, "Return of the Luddites," pp. 85–86; and Frank C. Zagare, "All Mortis, No Rigor," p. 108, all in *International Security*, Vol. 24, No. 2 (Fall 1999).
 Bruce Bueno de Mesquita and James D. Morrow "Sorting Through the Wealth of Notions"

^{3.} Bruce Bueno de Mesquita and James D. Morrow, "Sorting Through the Wealth of Notions," pp. 56–57, ibid.

correct.⁴ For this reason, Einstein declared that "even the most lucidly logical mathematical theory was of itself no guarantee of truth."⁵

Second, I stand by my claim that "although all three criteria are important. . . . originality and empirical validity are especially prized" (p. 13). Bold new theories understandably attract greater attention than subsequent efforts to tidy up the fine details of an argument. A creative new theory is unlikely to last long if it is wholly contradictory, but it will deserve widespread attention even if it rests on as-yet unidentified assumptions or contains causal claims that have to be qualified upon closer inspection. Why? Because a powerful new argument can be useful even when it contains inconsistencies, and because refining the logic of an argument is easier once one has an argument to examine.⁶ Both creating a new theory and refining its logic are useful parts of science, but the first one gets the loudest applause.

Third, formalization is not necessary to make precise, logically consistent arguments.⁷ Although critics like Emerson Niou and Peter Ordeshook regard virtually all nonformal work in social science as "mired in imprecision, vagueness, obscure logic, ill-defined constructs, nontestable hypotheses, and ad hoc argument," I believe many nonformal works of social science are clear, logical, and precise.⁸ Moreover, the underlying logic of many nonformal works is frequently easier to discern than an elaborate formal model, even for those who have acquainted themselves with the latter method.⁹

8. Niou and Ordeshook, "Return of the Luddites," p. 87.

^{4.} Compare, for example, Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence* (New York: Cambridge University Press, forthcoming); and Bruce Bueno de Mesquita and William Riker, "An Assessment of the Merits of Selective Nuclear Proliferation," *Journal of Conflict Resolution* Vol. 26, No. 2 (June 1982), pp. 283–306; or Robert Powell, "Stability and the Distribution of Power," *World Politics*, Vol. 48, No. 2 (January 1996), pp. 239–267; Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives* (New Haven, Conn.: Yale University Press, 1992), pp. 190, 205–206; and James D. Fearon, "War, Relative Power, and Private Information," paper presented at the annual meeting of the International Studies Association, Atlanta, Georgia, March 31–April 4, 1992.

^{5.} Albert Einstein, quoted in Timothy Ferris, *The Whole Shebang: A State of the Universe(s) Report* (New York: Touchstone, 1997), p. 28.

^{6.} This is true even of the most basic research tools. The calculus was a research tool of enormous value from the moment it was invented, but "a century and a half elapsed between the time the calculus was invented and the time [Augustin-Louis] Cauchy successfully gave it a logically acceptable form." See Judith V. Grabiner, *The Origins of Cauchy's Rigorous Calculus* (Cambridge, Mass.: MIT Press, 1981), p. 16.

^{7.} It should be noted that Bueno de Mesquita and Morrow and Martin acknowledge this point. See Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 58; and Martin, "Contributions of Rational Choice," p. 76.

^{9.} Robert Powell argues that nonformal theories suffer from a "lack of transparency" and an "inability to determine what follows from what," citing as evidence the fact that realists like John Mearsheimer and Charles Glaser disagree about certain aspects of international politics. See

A Model Disagreement 107

Fourth, much of the recent formal work in security studies does not devote a great deal of effort to identifying and resolving the (alleged) inconsistencies of nonformal scholarship. Rather, in many cases models are used to identify underlying assumptions or boundary conditions (i.e., the conditions that must obtain if an existing hypothesis is expected to operate). As I noted in my article, this sort of analysis can be useful. But identifying underlying assumptions or boundary conditions is not the same as showing that a well-verified nonformal theory was internally contradictory. And given the importance that my critics place on this criterion, it is noteworthy that they offer at most a single example of a logically contradictory argument that was corrected through formal analysis.¹⁰

Are Formal Modelers Creative?

All of the responses suggest that I underestimated the originality of recent formal work. Niou and Ordeshook emphasize the creativity that modeling requires, and several respondents suggest that I neglected important recent works or mischaracterized the works I did discuss. Both Lisa Martin and Frank Zagare also argue that it is unfair to focus on individual works rather than an entire stream of interrelated models. I consider each of these assertions in turn.

First, Niou and Ordeshook argue that I "cannot see the level of creativity that often goes into a model's design," noting also that "logical consistency is itself a profoundly important creative contribution."¹¹ The issue, however, is not whether the construction of a model involves creative thought; rather, it is whether the model leads to new, empirically valid insights about international security. As described in my article, this is usually not the case.

Second, Bueno de Mesquita and Morrow and Robert Powell challenge my assessment of a number of recent formal works, and offer a seemingly daunting list of "original contributions" made by formal modelers. Space does not permit me to discuss every one of the works they invoke (some of which are

Powell, "The Modeling Enterprise and Security Studies," *International Security*, Vol. 24, No. 2 (Fall 1999), pp. 99–100. In fact, it is quite easy to identify why scholars in the realist tradition reach different conclusions; see Stephen Brooks, "Dueling Realisms," *International Organization*, Vol. 51, No. 4 (Summer 1997), pp. 445–477.

^{10.} Martin points out that Thomas Schelling made somewhat inconsistent statements in two separate books, published six years apart. See Martin, "The Contributions of Rational Choice," p. 79.

^{11.} See Niou and Ordeshook, "Return of the Luddites," pp. 88, 89. Note the priority attached to logical consistency here as well.

still awaiting publication as of this writing), but a brief discussion will show that their list is not as compelling as it might appear.

COUNTEREXAMPLE NO. 1. Robert Powell, "Crisis Stability in the Nuclear Age."¹² This article suggests that Thomas Schelling's famous argument about the reciprocal fear of surprise attack depended on a hidden assumption (i.e., that neither side had the option of simply surrendering the stakes). I would make two points in response. First, this is an example of formal theory being used to identify boundary conditions rather than to make a new and original hypothesis. Second, Powell's argument amounts to saying that first-strike advantages do not make war inevitable so long as either side can avoid war by surrendering. This may be a useful qualification to Schelling, but it is hardly a surprising claim.

COUNTEREXAMPLE NO. 2. James D. Morrow, "Allies and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances."¹³ Bueno de Mesquita and Morrow argue that this article shows why states sometimes form alliances for reasons other than security. The article does not contain a formal model, however, and the central point—that alliances may involve one state sacrificing autonomy for security while its partners gain autonomy by providing it with greater security—does not require a formal presentation. As such, it cannot be invoked to demonstrate the fertility of formal techniques.

COUNTEREXAMPLE NO. 3. Randall Calvert, "The Value of Biased Information."¹⁴ Contrary to Bueno de Mesquita and Morrow's claim, this article does not show that "it is rational for political leaders to surround themselves with 'yes-men.'"¹⁵ Rather, Calvert presents a highly stylized model in which leaders "rationally" place greater weight on advice from those whose views they share than from those they regard as dissenters. The article does not "show" that political leaders *should* surround themselves with yes-men, however, because such a policy can create other dangers (such as a failure to consider a full range of alternatives) that exceed the benefits implied by the model. Calvert's article presents no empirical evidence, and it is worth noting that Alexander George's original work on multiple advocacy (which they portray as contrary to Cal-

Robert Powell, "Crisis Stability in the Nuclear Age," *American Political Science Review*, Vol. 83, No. 1 (March 1989), pp. 61–76.
 James D. Morrow," "Allies and Asymmetry: An Alternative to the Capability Aggregation

^{13.} James D. Morrow," "Allies and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances," *American Journal of Political Science*, Vol. 35, No. 4 (November 1991), pp. 904–933.

^{14.} Randall Calvert, "The Value of Biased Information," *Journal of Politics*, Vol. 47, No. 3 (May 1985), pp. 530–555.

^{15.} Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 62.

vert's model) explicitly warned that the views of those known to be dissenters may be discounted in the policymaking process.¹⁶

COUNTEREXAMPLE NO. 4. James D. Morrow, "Electoral Incentives and Arms Control."¹⁷ This article presents and tests a formal model linking domestic political considerations to negotiating positions on strategic arms control. The basic argument is that U.S. leaders will make more concessions as economic conditions worsen, but only up to a so-called turnover point. The model is largely atheoretical (i.e., the purported links between economic conditions, congressional preferences, and Soviet and U.S. negotiating positions are not well specified), and Morrow concedes that the evidence for his model is not very strong. He describes the statistical results as "suggestive," but admits that "they do not constitute conclusive evidence." He also acknowledges that "key [congressional] votes in future quarters produce mixed results depending on the exact specification." Although he runs a variety of regressions searching for the best fit, many of the regression coefficients do not achieve statistical significance. Thus Morrow concludes that "given the weakness of the statistical results, the question of whether turnover points exist remains open."¹⁸ This article provides evidence of Morrow's honesty (he admits that the model is not well supported), but it is hardly a good example of new and original insights resulting from formal theory.

COUNTEREXAMPLE NO. 5. Joanne Gowa, "Bipolarity, Multipolarity, and Free Trade."¹⁹ Gowa's article argues that the desire to strengthen one's allies can encourage great powers in a bipolar world to adopt policies of free trade. The central point to note, however, is that Gowa's article is not an example of formal theory. She does employ simple 2×2 games to illustrate her argument (including the familiar prisoners' dilemma), but she derives no equilibria and does not deduce testable hypotheses from the formal structure. Interestingly, Bueno de Mesquita and Morrow do not claim that formal theory led to

^{16.} George cites George Ball's opposition to President Lyndon Johnson's Vietnam policies as a case where a dissenter's advice was discounted because he "did not share the top policy makers' premise that Vietnam had become strategically important to the United States." See Alexander L. George, "The Case for Multiple Advocacy in Making Foreign Policy," *American Political Science Review*, Vol. 66, No. 3 (September 1972), p. 773.

^{17.} James D. Morrow, "Electoral Incentives and Arms Control," *Journal of Conflict Resolution*, Vol. 35, No. 2 (June 1991), pp. 243–265.

^{18.} Ibid., pp. 261, 262, 265.

^{19.} Joanne Gowa, "Bipolarity, Multipolarity, and Free Trade," *American Political Science Review*, Vol. 83, No. 4 (December 1989), pp. 1245–1256.

powerful new insights in this case. Rather, they merely assert that "the formal nature of Gowa's argument *facilitated* this theoretically fruitful debate."²⁰

COUNTEREXAMPLE NO. 6. James D. Fearon, "Rationalist Explanations for War."²¹ Bueno de Mesquita and Morrow argue that I mischaracterize Fearon's arguments about the effects of anarchy and private information. I disagree. Standard treatments of anarchy do not deny that states can reach agreements (and even abide by them): the main point is that the commitment problem has long been understood to be a central feature of anarchy. Similarly, I did not argue that there was no difference between "secrecy" and "private information," only that both can foster miscalculation and lead to war via essentially the same mechanism. Furthermore, although some forms of secrecy can be revealed in order to facilitate a deal, opponents are unlikely to take an enemy's revelations at face value. Thus the distinction between "secrecy" and "private information" may be difficult to discern in practice. Finally, my point was not that Fearon's article made no contribution at all, only that its central theoretical claims were not new.

COUNTEREXAMPLE NO. 7. Woosang Kim and James D. Morrow, "When Do Power Shifts Lead to War?"²² Bueno de Mesquita and Morrow argue that this article yields a number of novel insights. What "novel hypotheses" does it offer? First, "risk-acceptant rising states and risk-averse declining states increase the chance of war." Second, "the greater the rising state's dissatisfaction with the status quo, the more likely war is." Third, "the lower the expected costs of war, the more likely war is." These propositions are obvious, and the second and third are almost tautological. The fourth hypothesis—"war is more likely when the parties are roughly but not exactly equal in capabilities"—is less obvious, but hardly new.²³ Kim and Morrow perform a number of statistical tests of these propositions, and the results provide only partial support for their conjectures. (Among other things, the results are quite sensitive to the specific measures used to estimate the variables in the model.) Thus it is not clear how much has been learned from formalizing the problem as they do.

COUNTEREXAMPLE NO. 8. In his response, Powell challenges my assessment of recent formal work by observing that there are subtle differences between

^{20.} Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 64 (emphasis added).

^{21.} James D. Fearon, "Rationalist Explanations for War," *International Organization*, Vol. 49, No. 3 (Summer 1995), pp. 379–414.

^{22.} Woosang Kim and James D. Morrow, "When Do Power Shifts Lead to War?" American Journal of Political Science, Vol. 36, No. 4 (November 1992), pp. 896–922.

^{23.} Ibid., p. 907.

certain recent formal treatments and the work of earlier scholars (e.g., Robert Jervis). In particular, he argues that his definition of the "costs of war" differs from Jervis's depiction of the "offense-defense balance."²⁴

I agree that Powell's formulation of the costs of war in his article on relative gains is "analytically distinct" from some of the ways that the offense-defense balance has been defined in the literature. But the central theoretical argument advanced is not new: when military technology, geography, and so on make warfare more profitable, states will be more fearful of one another and less inclined to cooperate. Reasonable people can disagree over whether this is a major breakthrough or merely a simple refinement, but it is worth noting that Powell himself regards Jervis's article as the "seminal" analysis underpinning his own claim.

Finally, both Martin and Zagare suggest that it is unfair to examine individual articles in isolation, because a cardinal virtue of modeling is the capacity "to generate linked, coherent sets of propositions and insights."²⁵ Maybe so, but they offer no examples of where this capacity led to new and original ideas. Moreover, because any theory rests on a potentially infinite number of assumptions, one can always generate a new model by altering a key assumption of an earlier model, thereby generating a stream of interrelated models and creating the appearance of scholarly momentum.²⁶ Unless some effort is made to summarize and test the potentially infinite number of competing models, however, it is not clear what the overall contribution is. And as Niou and Ordeshook point out, "Despite the proliferation of competing models of deterrence, bargaining, coalitions, threats, and so on, those models are rarely set against each other for competitive empirical assessment."27

Furthermore, just as a chain of weak links will not bear much weight, a stream of models whose individual insights are familiar or unsurprising will not make much of a contribution. Zagare complains that I focused on only one of his many articles and suggests that the corpus of his recent work does yield more powerful new insights. If one looks at his other works, however, important new insights are hard to discern. He claims that it is a major advance to construct a theory of deterrence in which "threat credibility is a variable," but

^{24.} See Robert Jervis, "Cooperation under the Security Dilemma," World Politics, Vol. 30, No. 2 (January 1978), pp. 167-214; and Robert Powell, "Absolute and Relative Gains in International Relations Theory," American Political Science Review, Vol. 85, No. 4 (December 1991), pp. 1303–1320.
25. Martin, "The Contributions of Rational Choice," pp. 74, 76.
26. On this general point, see Donald N. McCloskey, Knowledge and Persuasion in Economics,

⁽Chicago: University of Chicago Press, 1994), chaps. 10, 13.

^{27.} See Niou and Ordeshook, "Return of the Luddites," p. 84.

the variable nature of threat credibility has been recognized as a central element of the theory since the 1960s.²⁸ Similarly, in "Assessing Competing Defense Postures," Zagare and Kilgour find that "a deterrence equilibrium . . . can occur under almost any conditions, provided that the players have an existential fear of escalation." Turning briefly to the real world, they explain the absence of war between the United States and the Soviet Union by arguing that "the Soviet Union, while motivated to expand, was unwilling to fight a costly strategic war to do so and U.S. leaders knew it." They admit that they cannot explain why limited wars did occur at various points in the Cold War (in their words, such conflicts "lie outside the parameters of the present model"), and they conclude by agreeing with Jervis that "a rational strategy for the employment of nuclear weapons is a contradiction in terms."²⁹ Readers with a particular interest in the minutiae of abstract bargaining theory may find these arguments creative and original, but scholars who are interested in the real world are unlikely to find their understanding enhanced by reading these works.

In sum, developing a formal theory can be a creative act, and formal theorists do offer new ideas on occasion. On the whole, however, the production of new ideas and hypotheses is not impressive. Among other things, much of the recent formal work suffers from the diminishing returns common to "normal science." Once an initial theoretical breakthrough is made, formalization is one way to refine, qualify, or extend the results. As I have said repeatedly, these contributions can be useful. But the added value generally declines as scholars pursue increasingly rarified results. I therefore stand by my original statement that "formal theory enjoys no particular advantage as a source of theoretical creativity" (p. 30). And although I share Powell's belief that the best way to judge a body of scholarship is to read it one's self, most readers who take the time to do so will reach the same conclusion that I did.

^{28.} See Zagare, "All Mortis, No Rigor," p. 112; and Thomas C. Schelling, *Arms and Influence* (New Haven, Conn.: Yale University Press, 1966), especially chap. 2.

^{29.} See Frank C. Zagare and D. Marc Kilgour, "Assessing Competing Defense Postures: The Strategic Implications of Flexible Response," *World Politics*, Vol. 47, No. 3 (April 1995), pp. 400, 403–404, 406–407. In another article, Zagare and Kilgour discover that "the strategic position of a defender without a credible first-stage threat is not enviable," which was precisely the critique of massive retaliation offered by critics like William W. Kaufmann in the 1950s. They also find that "if credibility is considered to vary only across issues, the model suggests that confrontations are least likely as the issues becomes less salient to one side or the other." And when considering why nuclear war did not occur in the 1950s (when U.S. strategy was allegedly one of "massive retaliation"), they admit that "the model provides no obvious answer to this question." See Zagare and Kilgour, "Modeling Massive Retaliation," *Conflict Management and Peace Science*, Vol. 13, No. 1 (Fall 1993), pp. 78–79.

A Model Disagreement | 113

Testing 1, 2, 3 . . .

In "Rigor or Rigor Mortis?" I argued that formal modelers have placed relatively little weight on empirical testing. Many formal articles offer no empirical evaluation at all, and some of the more ambitious efforts to provide empirical tests do not achieve a high standard of scholarly rigor. My critics do not offer a serious challenge to this assessment.

First, neither Bueno de Mesquita and Morrow nor Niou and Ordeshook devote much effort to defending the empirical validity of the works I criticized in my article, and the defenses they do offer are not convincing. Bueno de Mesquita does not even try to defend the case studies in War and Reason and now says that they were merely intended "to illustrate the intuition behind the model's logic and not as evidence."³⁰ He defends the quantitative evidence in the book by citing two recent articles (one of them a revised version of a paper cited by me, the other a forthcoming article). If one actually reads these articles, however, it is difficult to see how he can regard them as supporting the claims made in his book. According to one of the authors he cites, "All we can say about the [Bueno de Mesquita and Lalman] model is that the percentage of cases correctly predicted (CP) is between 1% and 41%. As a reference, a null model that always predicted Status Quo would correctly predict 34% of the observations. So, if we were to give the international interaction game the greatest benefit of doubt possible, it would predict 21% better than the modal category." This same author concludes that "there is less support for the international interaction game than Bueno de Mesquita and Lalman claim."31 With "support" like this, who needs criticism?

Niou and Ordeshook, by contrast, appear to back away from the idea of rigorous testing at all. Instead of defending the empirical evidence provided in their book, they counter by declaring that "reality is far too complicated to be accommodated in any straightforward way by any simple tractable model," and that "scientific testing is an imprecise, often informal process." This reasoning is used to justify their heavy reliance on ad hoc arguments in the empirical portions of their book: "To suppose that a formal model can wholly

^{30.} Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 68.

^{31.} See Curtis S. Signorino, "Strategic Interaction and the Statistical Analysis of International Conflict," *American Political Science Review*, Vol. 93, No. 2 (June 1999), pp. 290, 292, 294 (emphasis added). A second paper cited by Bueno de Mesquita and Morrow concludes that "although the data suggest that Bueno de Mesquita's composite measures influence decision making during crises, the results are still weak." See Alastair Smith, "Testing Theories of Strategic Choice: The Example of Crisis Escalation," paper presented at the annual meeting of the American Political Science Association," Boston, Massachusetts, September 3–6, 1998, p. 21.

encompass a complex process . . . without resorting to some ad hoc discussion is ludicrous."³² I agree that all theories simplify reality and that testing can be an imprecise business, but it is hard to be impressed when a model employs assumptions that are wildly at odds with our empirical knowledge and when the fit between theory and evidence *depends* on ad hoc factors that contradict the main elements of the theory. Among other things, the model presented in *The Balance of Power* assumes that war is costless, that a state's entire stock of resources can be transferred costlessly to another state, and that all states have complete information about one another's strength. Given these unrealistic assumptions, it is perhaps not surprising that "the notion of 'ceding' [territory], which is central to the model, is virtually absent from the history" to which it is applied.³³ Their model is not an unavoidable simplification of an admittedly complex reality; it is an artificial creation that bears little resemblance to the empirical world they are attempting to analyze.

Second, several of my critics suggest that I employ a double standard on this issue, noting that I praised nonformal theorists like Kenneth Waltz and Robert Jervis despite the lack of empirical testing in their work.³⁴ In fact, although I admire the work of Waltz, Jervis, Schelling, and others, I believe their work would be better had they devoted greater effort to testing their claims empirically. Some of Schelling's ideas about coercion do not work very well when they are brought into the real world, and those of us who have tried to test Waltz's neorealist balance-of-power theory have found it necessary to modify the theory in order to conform to historical experience.³⁵ Thus there is no double standard at work here.

Third, several of my critics suggest that modelers do not need to test their conjectures themselves, because there can be a division of labor between scholars who derive hypotheses and those who test them.³⁶ I agree that individual scholars have different comparative advantages and that expecting

^{32.} Niou and Ordeshook, "Return of the Luddites," pp. 93-94.

See Emerson M.S. Niou, Peter C. Ordeshook, and Gregory F. Rose, *The Balance of Power: Stability in International Systems* (Cambridge: Cambridge University Press, 1989); and Glenn Snyder, "Alliances, Balance, and Stability," *International Organization*, Vol. 45, No. 1 (Winter 1991), p. 137.
 See Powell, "The Modeling Enterprise and Security Studies," p. 104; and Niou and Ordeshook,

[&]quot;Return of the Luddites," p. 90. 35. On Schelling, see Wallace J. Thies, When Governments Collide: Coercion and Diplomacy in the

Vietnam Conflict, 1964–1968 (Berkeley: University of California Press, 1980); and Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, N.Y.: Cornell University Press, 1996). On Waltz, see Stephen M. Walt, *The Origins of Alliances* (Ithaca, N.Y.: Cornell University Press, 1987), pp. 21–22, 263–264.

^{36.} See Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 71; and Zagare, "All Mortis, No Rigor," p. 114.

many modelers to do careful empirical testing may not be an efficient allocation of their talents. My point, however, was that the modeling community as a whole has tended to place a low value on this criterion. More important, the argument that the scholarly community can rely upon a division of labor is convincing only if one is committed to maintaining a diverse array of scholars in the field. As discussed below, however, it is not clear that this is what most formal modelers seek.

Policy Relevance

Several respondents argue that formal models have made a significant contribution to real-world policy debates. I do not deny that formal modelers have written on policy-relevant topics in a few cases, but as I said in my article, they "have joined in only after the central parameters were established by others" (p. 47). I stand by my original assertion, therefore, and most of my critics do not challenge it. Moreover, I see little evidence that the policy community has paid much attention to recent formal work, or that policymakers would gain much real-world insight if they did. For example, has the vast formal literature on crisis bargaining produced insights or lessons that might actually help someone who was trying to manage a real-world crisis? What novel and practical lessons have been derived from the abstract discussions of nuclear deterrence found in the writings of prominent formal theorists? By contrast, more concrete explorations of deterrence policy by scholars such as Bruce Blair, John Steinbruner, and Scott Sagan have identified important dangers, suggested a variety of useful remedies, and attracted considerable attention in both the academic and policy worlds.³⁷

The boldest challenge to my claims about the irrelevance of much of the recent formal work comes from Bueno de Mesquita and Morrow, who argue that Bueno de Mesquita's applied forecasting model is "a practical tool for policy analysis." They quote testimony from a government official suggesting that the model has an accuracy rate of 90 percent and declare that "the United States government . . . finds the model accurate and . . . uses the model to assist with important foreign policy matters."³⁸

^{37.} See Bruce G. Blair, *Strategic Command and Control: Redefining the Nuclear Threat* (Washington, D.C.: Brookings, 1985); John D. Steinbruner, "Beyond Rational Deterrence: The Struggle for New Conceptions," *World Politics*, Vol. 28, No. 2 (January 1976), pp. 223–242; and Scott D. Sagan, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons* (Princeton, N.J.: Princeton University Press, 1993).

^{38.} See Bueno de Mesquita and Morrow, "Sorting Through the Wealth of Notions," p. 70.

Space does not permit a detailed critique of these claims, but such assertions should be taken with many grains of salt. First, although Bueno de Mesquita has described his model in general terms, the published descriptions of the forecasting model are not sufficiently detailed to permit others to replicate all of his results. Such opacity makes sense from a commercial point of view, but in science, transparency should take precedence over preserving market share.

Second, Bueno de Mesquita's claims to predictive accuracy are questionable. Without access to the full range of predictions made by the model, it is impossible to tell just how accurate it is or how many novel predictions it actually makes. But the published record of the forecasting model is not impressive. One article "predicts" the U.S. victory in the Cold War; unfortunately, the article was written six years after the Soviet Union collapsed.³⁹ Another article attempts to forecast the Middle East peace negotiations, but fails to anticipate either the mechanism by which the process occurred or the final outcome.⁴⁰ Similarly, a coauthored book forecasting the fate of Hong Kong offers a host of familiar generalities (for which no model was needed), along with a number of more specific and controversial predictions that have not fared well thus far.⁴¹

^{39.} See Bruce Bueno de Mesquita, "The End of the Cold War: Predicting an Emergent Property," *Journal of Conflict Resolution*, Vol. 42, No. 2 (April 1998), pp. 131–155. Bueno de Mesquita performs 100 simulations using the model and discovers that the United States wins the Cold War in the majority of them. This result is not surprising, insofar as the United States began the Cold War with three times the gross national product of the Soviet Union and a superior geopolitical position. For a more prescient prediction, consider Kenneth N. Waltz's 1979 forecast that "with half of our GNP [gross national product], [the Soviet Union] nevertheless has to run hard to stay in the race. One may think that question is not whether a third or fourth country will enter the circle of great powers . . . but rather whether the Soviet Union can keep up." Similarly, my own analysis of the Cold War (published in 1987) argued that "the most important causes of security cooperation work to its advantage and isolate the Soviet Union from virtually all of the world's strategically significant states." See Waltz, *Theory of International Politics* (Reading, Mass.: Addison-Wesley, 1979), pp. 179–180; and Walt, *Origins of Alliances*, chap. 8, especially pp. 284–285.

^{40.} See Bruce Bueno de Mesquita, "Multilateral Negotiations: A Spatial Analysis of the Arab-Israeli Dispute," *International Organization*, Vol. 44, No. 3 (Summer 1990), pp. 317–340. Among other things, the model assumes that the Soviet Union is as powerful a player as the United States in 1990 and fails to anticipate the Soviet collapse in 1991. The model predicts that Israeli-Palestinian negotiations would yield "nothing approaching even a semiautonomous state," and that "there appears to be no reason to anticipate more than modest concessions by the Israelis to the interests of the Palestinians in the near future." Ibid., pp. 337, 340. In fact, Israel did make important concessions during the peace process, and the Palestinians are very close to having their own state. 41. For example, the authors claim that "the succession [to Deng Xiaoping] will be clouded by severe infighting" featuring "several rounds of brutal exchanges," and suggest that "the [Communist] party will hold on to some power and perhaps nobody will win." The prediction is not very precise, but signs of severe infighting have been notably absent thus far. They also predict that "there is likely to be a sudden and dramatic collapse of support for market reforms" within a year

A Model Disagreement | 117

Lastly, the evidence that the "U.S. government . . . uses the model to assist on important foreign policy matters" is dubious. Bueno de Mesquita's claims rest on the testimony of one midlevel Central Intelligence Agency official and an *Izvestiya* article recounting a briefing by unnamed U.S. "officials." One midlevel bureaucrat does not equal the "U.S. government," however, and two former directors of the National Intelligence Council (which prepares National Intelligence Estimates) and a former deputy director for intelligence have reported that they were not aware that Bueno de Mesquita's model had any impact on the estimation process or on policy. According to one of these officials, models like Bueno de Mesquita's are primarily useful "to stimulate questions for further research and study."⁴² Based on the evidence to date, this forecasting model is a weak reed upon which to base a claim to real-world relevance.

Methodological Pluralism

A central theme of my article was the importance of methodological diversity. Specifically, I pointed out that security studies has always been "theoretically and methodologically diverse," and "the field as a whole will be richer if such diversity is retained *and esteemed*" (pp. 8, 47–48, emphasis in original).

Given my position, Martin's data showing that formal modelers do not "dominate" the field is beside the point. The issue is not whether a particular group or methodological technique is currently hegemonic; it is whether any group has hegemonic *ambitions*. After all, the time to resist hegemony is before an imperialist movement becomes too strong to resist, not after it has established itself in a position of predominance.

Let us be candid. There is a widespread perception that formal modelers are less tolerant of other approaches than virtually any other group in the field of

or two of the Chinese takeover of Hong Kong, and forecast "that Hong Kong's autonomy will be eroded quickly, only to be restored for a period of a year or so." Neither development has occurred. They predict a dramatic decline in press freedom in Hong Kong (which has not transpired), and suggest that an important bellwether would be the suppression of demonstrations in Hong Kong commemorating the Tiananmen Square uprising on June 4, 1989. So far, demonstrations have been permitted. They correctly predict an economic slowdown in Hong Kong and in China itself, but this decline was the result of the East Asian financial crisis rather than the transition to communist rule. Nor has the value of the Hong Kong dollar declined significantly, despite the pressures created by the financial crisis. For their original forecasts, see Bruce Bueno de Mesquita, David Newman, and Alvin Rabushka, *Red Flag over Hong Kong?* (Chatham, N.J.: Chatham House, 1996), pp. 8–9, 94, 97–98, 126–127, 129–130.

^{42.} Personal correspondence with Joseph Nye, Richard Cooper, and Douglas McEachin.

political science. This is not true of every scholar who uses formal methods, but it is striking how widespread the belief is and equally striking to note that formal modelers are the only group in our profession that is regarded in this way.

Do my critics' responses lend support to these concerns? Sadly, yes. Although several of them pay lip service to the principle of methodological pluralism, their disregard for nonformal approaches is apparent. As noted above, Niou and Ordeshook clearly regard most scholarship in political science as primitive at best, and they deride those who find formal work needlessly obscure as Luddites "who studied French and Plato in college rather than calculus." (For the record, I studied all three). Niou and Ordeshook's contempt is even more apparent when they write that there is nothing in the field of strategic studies that deserves the label of "theory," or when they refer to the scholarly study of real-world problems as "mere journalism, until it can be given the solid scientific grounding that formal theorists pursue."⁴³ Similarly, although Bueno de Mesquita and Morrow endorse my claim that "security studies should welcome contributions from formal theory," their failure to include the rest of the sentence (which calls for the inclusion of "large-Nstatistical analysis, historical case studies, and even the more rigorous forms of interpretive or constructivist analysis," p. 48) is revealing. The question is: Do they share my belief that each of these methods should be "retained and esteemed?" Finally, Martin's claim that "as scientific fields . . . develop, they invariably become more mathematical" betrays a belief that over time, nonformal approaches should be relegated to the dustbin of history (or history departments).⁴⁴ Some branches of science have become highly mathematical, but others (such as biology and geology) retain a large and important qualitative dimension. In any event, the record to date does not suggest that formal models in security studies are superior to other research traditions.

Finally, we should not forget that a professed commitment to pluralism is "cheap talk." It is easy to say one is in favor of other approaches, but the real question is how different scholars react when allocating scarce resources. This is an empirical question, and all members of the field are free to participate in the research project. Over time, we can all keep track of which methodological subfields show signs of imperialist tendencies, consistently favoring their own

^{43.} Niou and Ordeshook, "Return of the Luddites, pp. 87, 93, 96.

^{44.} Martin, "Contributions of Rational Choice," p. 78.

A Model Disagreement | 119

tribe over others.⁴⁵ Needless to say, concerns about the imperialistic tendencies of formal modelers will decline if they are willing to endorse and support scholars whose work uses other well-established techniques, rather than exhibiting a clear and consistent preference for other formal modelers.

Conclusion

Scholarship is a competitive enterprise, and knowledge advances partly through the clashes of competing ideas. Yet the competition that drives progress should be tempered with the recognition that different research traditions can and should coexist. Just as biodiversity is central to a healthy ecosystem, intellectual diversity is an important part of a healthy scholarly community. In the past, security studies has profited by welcoming contributions from a diverse array of historians, political scientists, economists, natural scientists, psychologists, and others. The field has been methodologically and theoretically wide-ranging, but united by a close concern with real-world policy issues. This combination of diversity has enabled scholars with different backgrounds and talents to profit from one another's contributions, thereby allowing the field as a whole to advance more swiftly than it would were any single tradition to extinguish the others.

In short, there are good reasons to encourage a diversity of research approaches within any subfield as important as security studies. Because there is scientific value in each of the established traditions of contemporary social science research, the field of security studies will be impoverished if any single approach becomes hegemonic.

^{45.} Lest I again be accused of a double standard, I offer the following data. During my ten years as a tenured faculty member at the University of Chicago, there were eleven occasions where my department voted to hire, renew, or promote a formal modeler. I voted in favor nine times and voted to oppose twice, a percentage similar to my record on nonformal candidates.