

The Perils of Predicting Proliferation

Alexander H. Montgomery
Department of Political Science

Reed College, Portland

Scott D. Sagan

Department of Political Science

Stanford University, Stanford

Studies of nuclear proliferation share five serious problems. First, nuclear programs' initiation and completion dates are ambiguous and difficult to code, but findings are rarely subjected to sufficient robustness tests using alternative codings. Second, independent variables overlook important factors such as prestige and bureaucratic power and often use poor proxies for concepts such as the nonproliferation regime. Third, methodologies and data sets should be tightly coupled to empirical questions but are instead often chosen for convenience. Fourth, some findings provide insights already known or believed to be true. Fifth, findings can ignore or gloss over data crucial for policy making and wider debates. This article reviews new quantitative research on nuclear proliferation, noting improved analysis and lingering problems. It highlights the 1999 Kargil war to explore dangers of relying on stock data sets and the need for research on statistical outliers. It concludes with a future research agenda aimed at correcting problems and a cautionary note regarding hasty application of quantitative results to policy making.

Keywords: *nuclear weapons proliferation; nonproliferation treaty; robustness; Kargil; policy relevance*

The systematic study of the causes and consequences of the spread of nuclear weapons is at an early stage of development. The essays in this special journal issue (Gartzke and Kroenig 2009) signal a renaissance of interest in quantitative empirical studies of nuclear proliferation, an approach that was mostly dormant after early attempts at predicting proliferation in the 1980s (Kegley 1980; Meyer 1984). A new generation of political scientists, however, has recently taken up the charge of assessing the causes and consequences of the spread of nuclear weapons (Singh and Way 2004; Asal and Beardsley 2007; Jo and Gartzke 2007; Fuhrmann 2008). The authors in this journal issue continue that important effort by creating new data sets to widen the scope of inquiry, using new and innovative methodologies and testing a variety of theories to

explain how and why states acquire nuclear weapons and to predict the effects of the spread of nuclear weapons on the likelihood and outcome of military conflicts.

This article reviews and critiques this second wave of quantitative literature on nuclear-weapons proliferation. These new studies provide an important counterpart to the mainstream proliferation literature, which generally has focused on in-depth individual case studies or small-*n*, cross-case variation rather than attempting to test hypotheses in a systematic way across all states during long periods of time. The articles in this issue demonstrate the advances made since early attempts at quantifying the causes and effects of proliferation, significantly improving our understanding of these phenomena. This new research on the causes of proliferation usefully focuses on the causes of nuclear assistance from other countries and estimates the effects of that assistance on nuclear-weapons acquisition. It demonstrates convincingly that general nuclear-technology assistance is affected by patterns of enmity and alliance and that receiving sensitive nuclear assistance (help with uranium enrichment, plutonium reprocessing, or nuclear-weapon design) significantly increases the chances of acquiring nuclear weapons (Fuhrmann 2009, this issue; Kroenig 2009, this issue).

The articles in this special issue also offer innovative insights into the nuclear optimism–pessimism debate: does the proliferation of nuclear weapons make states more cautious and reduce the likelihood of war, or does it lead to nuclear-armed states' seeking additional gains and risking catastrophe? The findings are mixed. States that have nuclear weapons are more likely to initiate militarized interstate disputes (MIDs); this tendency increases with nuclear experience (Horowitz 2009; Gartzke and Jo 2009). However, one article argues that the increase in MID initiation results from endogeneity (Gartzke and Jo 2009), while another demonstrates that as states become more experienced with nuclear weapons, they become less likely to reciprocate military challenges or have their challenges reciprocated (Horowitz 2009). The first two findings bolster the pessimism argument, while the latter two support the nuclear optimists' position.

These articles also offer insights into the effects of nuclear weapons on the outcomes of crises and conflicts. Here, the authors find that dyads in which only one state has nuclear weapons are more likely to experience all levels of force in MIDs, with the nuclear state's winning more often and achieving a settlement more quickly in international crisis behavior (ICB) crises. Dyads in which both have nuclear weapons are much less likely to experience the maximum level of force (war) in a MID but are no more or less likely to win or to experience a shorter conflict in an ICB crisis (Beardsley and Asal 2009; Rauchhaus 2009). Thus, nuclear states seem to have incentives to be belligerent against both nuclear and nonnuclear states (which helps the pessimists' case), refraining from escalating to war only against other nuclear states (which helps the optimists' case).

These new studies of nuclear proliferation, however, despite using the most up-to-date data sets and techniques available, do not resolve five serious problems that have

also plagued earlier quantitative studies of nuclear proliferation. First, it is inherently difficult to have accurate coding of the dependent variables regarding whether states are exploring, pursuing, or have acquired nuclear weapons: some governments' civilian nuclear-power programs reflect internal ambivalence about whether the state should pursue a bomb option, and the intense secrecy surrounding weapons programs has meant that earlier published data sets left out many cases of covert programs or the creation of "hedging" options, a problem likely to remain persistent for the foreseeable future. While the articles in this issue use a common data set for nuclear-weapons possession, increasing the comparability of results across this set of articles, their use of a common data set decreases the robustness of their findings because, as we will demonstrate, other reasonable codings on when states initiated or completed nuclear-weapons programs produce very different results. Second, the coding rules used for existing independent variables in nuclear-proliferation analyses are also problematic. Too often, quantitative researchers measure what is easily measured rather than find ways of accurately capturing the concepts that our theories suggest should be important. Regarding the causes of proliferation, for example, important factors that have been discussed in historical case studies of proliferation—such as leaders' psychology, bureaucratic power, and military autonomy and the desire for prestige—are often excluded altogether or measured poorly in statistical studies. The problem also exists for the quantitative literature on the consequences of proliferation, in which the relevant independent variable is assumed to be a state's first nuclear weapon, despite a long-standing debate in the historical literature about whether states are deterred by virtual nuclear weapons, a single bomb, a small arsenal, or a second-strike nuclear capability. Instead of developing new data sets to test for the effects of such factors, the literature uses proxy variables based on pre-existing data sets, making the studies particularly vulnerable to the "looking for the key under the lamppost" charge, as we will demonstrate below. Third, methodologies and data sets need to be tightly coupled to the empirical questions being investigated. Including a broader set of states in a data set without proper controls allows states that cannot have any influence on processes or outcomes to skew results, as we will also demonstrate. Fourth, some quantitative proliferation articles have findings that are trivial, providing us with insights that we already knew or at least believed to be true. While it is always valuable to provide tests of commonly held views to see if they are indeed accurate, many of these studies have failed to produce counterintuitive insights that would add more significantly to our understanding of proliferation. Fifth, and finally, statistical findings can ignore or gloss over individual data points that are crucially important for policy making and wider scholarly debates. We note in this article, for example, that the 1999 Kargil war between India and Pakistan—the only undisputed case of a war between nuclear-armed states—is simply ignored by these studies, as it does not appear as a separate dispute in the MID database and is not listed as a war in the Correlates of War (COW) or ICB databases.¹

It is valuable in this light to compare the study of nuclear proliferation to the study of the “democratic peace” in international relations. In both of these research areas, the scholarship has slowly expanded from qualitative theoretical and historical inquiry to include more game-theoretic approaches and systematic quantitative testing of hypotheses. The development path of research in the democratic peace has progressed significantly during the past two decades in ways that are instructive. This literature began with deductions based on political theory and limited historical evidence that suggested that democratic states rarely, if ever, go to war against each other (Kant [1795] 1917; Doyle 1986). Vigorous debates about the democratic-peace theory emerged in political science starting in the 1980s, and a progressive research program of many scholars using multiple methods eventually produced a significant, although by no means complete, degree of consensus about how and why democracies differ from nondemocracies in their foreign policy and international-security behavior (Lake 1992; Maoz and Russett 1993; Fearon 1994; Mansfield and Snyder 1995; Elman 1997; Schultz 1998; Russett and Oneal 2001; Reiter and Stam 2002; Gartzke 2007). Our understanding of the effect of democracy on the likelihood and conduct of war has been significantly enhanced by both rich historical and rigorous quantitative scholarship on the democratic peace. We hope that this special issue and this critique can lead scholars to emulate the democratic-peace literature, produce more multimethod research, and contribute further to improved understanding of the causes and consequences of nuclear proliferation.

In the first section, we review the earlier historical case-study and quantitative research on nuclear proliferation and compare its findings with the more recent quantitative literature on proliferation. Here, we note improvements as well as lingering problems with the coding, methodology, and empirical findings. Having established the context for the remainder of the article, we proceed by discussing the accomplishments and remaining problems in the articles in this special journal issue. Here, we analyze all three topics in this issue: (1) the supply and regulation of nuclear technologies, (2) the effects of the spread of nuclear weapons on the likelihood of conflict, and (3) the effects of nuclear weapons on the consequences of crises and conflict. We conclude with a discussion of the implications for future research and policy debates about how to deal with nuclear proliferation.

Previous Studies of Proliferation

Different studies of the causes of nuclear proliferation divide incentives and actors in different ways. In a widely cited analysis of nuclear nonproliferation, Scott Sagan compares three models: security, domestic politics, and norms. The first focuses on military security motivations; the second emphasizes domestic political or bureaucratic coalitions that form to support or oppose nuclear weapons; and the third focuses on norms and prestige considerations that encourage acquisition of nuclear weapons

(Sagan 1996/1997, 1999). He finds strongest support for the security model, although he argues that domestic interests and prestige concerns are “sufficient, but not necessary” conditions for proliferation in a limited number of cases. Scholarly debate has continued about how to weigh these different motivations for proliferation, with some realists arguing for the singular importance of security motivations (Mearsheimer 1990; Frankel 1993) and other scholars focusing on the constraints stemming from domestic economic interests or normative opposition against nuclear weapons (Solingen 1994, 1998, 2007; Katzenstein 1996). Suzette Grillot, Alexander Montgomery, and others have, in contrast, argued that role conceptions and identity play a major role in producing interests in starting and ending nuclear-weapons programs (Chafetz, Abramson, and Grillot 1996; Grillot and Long 2000; Montgomery 2005). Additional studies include cognitive and psychological factors that may contribute to decisions to acquire the bomb (Ogilvie-White 1996, 51-53; Hymans 2006).

Much of this historical case-study work, however, examined only cases in which governments had developed nuclear arsenals. To overcome the limits of research that “selects on the dependent variable,” scholars have also focused on specific cases in which governments chose to get rid of nuclear weapons or forgo the option before they had acquired an arsenal. T. V. Paul, for example, discusses the effects of power and norms and finds that states’ security positions are primarily responsible for their willingness to accept nuclear disarmament (Paul 2000). Some valuable case-study focused accounts include all of these factors without seeking to analytically separate them into individual accounts (Reiss 1988, 1995; Reiss and Litwak 1994; Levite 2002/2003; Campbell, Einhorn, and Reiss 2004).

The first wave of quantitative studies sought to rectify the weaknesses of limited case studies. Charles Kegley’s original quantitative study attempted to measure twenty different national and systemic measures for proliferation in a scattershot fashion, finding significant variation in three categories: threat perception, achievement of international status, and domestic satisfaction. His study is unique among quantitative studies in that it attempts to measure directly a state’s social role-conception in the international system rather than using a proxy for overall status (Kegley 1980). Stephen Meyer divided incentives into three categories—political power and prestige, military security, and domestic politics—and then measured these using fifteen indicators. He found that pariah status, fear of a well-armed rival, and desires for regional pre-eminence were the most important proliferation incentives but that alliances with a nuclear power, membership in the Non-Proliferation Treaty (NPT), and fear of pre-emptive strikes helped to dissuade states (Meyer 1984, 44-74).

In the second, more recent, wave of proliferation studies, Sonali Singh and Christopher Way divide incentives between internal (regime type and liberalization-related incentives) and external (security threats and guarantees) determinants, while Dong-Joon Jo and Erik Gartzke divide their incentives along Sagan’s three models (Singh and Way 2004, 862-65; Jo and Gartzke 2007, 168-71). Singh and Way find strong support for both enduring rivalries and militarized disputes on the external

side and limited support for economic openness and liberalization as domestic political factors. Jo and Gartzke find considerable support for status-driven motives and somewhat less support for domestic political and international security motives.

These studies also include supply-side factors that may limit proliferation, a subject that is given less attention in the qualitative literature. Meyer developed a set of fourteen factors that contribute toward a domestic capability, divided into factors that influence mining and processing of nuclear materials, construction and operation of nuclear facilities, and weapons fabrication; a subset of his indicators were extended through 1992 by Jo and Gartzke. Singh and Way also analyze technological determinants using more general indicators, which are coded into three variables: GDP, a general industrial threshold, and specific industrial capabilities (Meyer 1984, Appendix B; Singh and Way 2004, 867-69; Jo and Gartzke 2006, 5-7). All three models find consistently strong support for supply-side factors as major determinants of proliferation.

The Meaning of Proliferation

An important insight from these studies and their descendants is that more gradations exist regarding the status of a state than its simply being nuclear or nonnuclear. The contribution of the quantitative proliferation literature here is important, but because quantitative studies of nuclear proliferation by Meyer, Singh and Way, and Jo and Gartzke have adopted different methods of coding nuclear activity, they are difficult to compare. Meyer simply coded whether a state had made a "decision" to seek nuclear weapons or not and did not examine whether states succeeded or not (Meyer 1984, Appendix A). Jo and Gartzke (2006, 1-2) operationalized this variable as a measure of whether the "highest decision maker in a given state authorized [or ended] a nuclear weapons program" for official programs or the year in which "nuclear activities increase noticeably" for clandestine programs. They also added acquisition data for the dates on which "each state was ready to quickly assemble nuclear components into nuclear weapons." Singh and Way (2004, 866-67) measure nuclear status along a continuum with four "degrees of nuclearness": no interest, exploration, pursuit, and assembly. The first degree is simply no interest in nuclear weapons. The second degree is exploration, which is "demonstrated by political authorization to explore the option or by linking research to defense agencies that would oversee any potential weapons development." To qualify as pursuing nuclear weapons, states must take steps such as "a political decision by cabinet-level officials, movement toward weaponization, or development of single-use, dedicated technology." Finally, both Jo and Gartzke (2006) and Singh and Way (2004) code states as having fully acquired weapons if they either test or possess a "functional nuclear weapon." The codings from Meyer, Singh and Way, and Jo and Gartzke are compared in Table 1.

Table 1
Nuclear Status Coding Schemes

	Meyer 2004 (1942–1980)	Jo and Gartzke 2007 (1941–2002)		Singh and Way 2004 (1945–2000)		
	Decide	Programs	Possession	Explore	Pursue	Acquire
Acquisition						
United States	1942–	1942–	1945–	*	*	1945–
Russia	1942–	1943–	1949–	*	1945–	1949–
United Kingdom	1947–	1941–	1952–	1945–	1947–	1952–
France	1956–	1954–	1960–	1946–	1954–	1960–
China	1957–	1956–	1964–	1955–	1955–	1964–
Israel	1968–	1955–	1966–	1949–	1958–	1972–
India(1)	1964–1966	1964–1965		1954–	1964–	1974–1974
India(2)	1972–	1972–	1988–	1975–	1980–	1988–
South Africa	1975–	1971–1990	1979–1991	1969–	1974–	1979–1993
Pakistan		1972–	1987–	1972–	1972–	1990–
Pursuit						
Korea, South	1972–1975	1971–1975		1959–	1970–1978	
Libya				1970–	1970–	
Brazil		1978–1990		1953–	1978–1990	
Argentina		1976–1990		1968–	1978–1990	
Korea, North		1982–		1965–	1980–	
Iraq		1973–2002		1976–	1982–	
Iran(1)		1974–1978				
Iran(2)		1984–		1984–	1985–	
Exploration						
Germany		1941–1945		*		
Japan		1943–1945		*		
Switzerland				1946–1969		
Sweden		1946–1969		1954–1969		
Yugoslavia(1)		1953–1963		1954–1965		
Yugoslavia(2)		1982–1987		1974–1988		
Australia				1956–1973		
Taiwan(1)		1967–1976		1967–1977		
Taiwan(2)				1987–1988		
Algeria				1983–		
Romania		1981–1989		1985–1993		

* Before initial observation in data set.

The key idea that nuclear proliferation should not be viewed as a single end-state is valuable. Proliferation is a process by which countries move closer to or away from different thresholds toward developing the bomb. Countries will not necessarily stay solidly in one state of “nuclear latency” or another, as internal and external

conditions that fuel or suppress proliferation may change over time. Governments do not actually need to “decide” to “go nuclear” until the moment that they test a nuclear device; even then, the decision to develop a deliverable weapon or declare nuclear-weapons status can be independent of a nuclear test. The case of India is instructive; the 1974 test device was declared a “peaceful nuclear explosive,” and India did not claim status as a nuclear-weapons state, nor did it weaponize, for a number of years. In 1998, the New Delhi government made two decisions: the first was to test; the second was to declare India’s status as a nuclear-weapons state.

The ambiguity of nuclear proliferation—and consequently, the pitfalls of selecting one coding system or another—becomes apparent when comparing different coding systems for exploration, pursuit, and acquisition of nuclear weapons. Even the dates of the programs for the initial five nuclear states (N5) are debatable; Meyer (1984) has China starting in 1957, while Jo and Gartzke (2007) have it starting in 1956 and Singh and Way (2004) have it jumping directly to pursuit in 1955 without actually exploring nuclear weapons first. After the N5, the data quickly become quite murky. Dates for Israeli acquisition range from 1966 to 1972 (Meyer [1984] does not even code Israel as deciding to seek nuclear weapons until 1968); Libya, Switzerland, Australia, Algeria, and Taiwan’s second attempt at a nuclear-weapons program do not appear in Jo and Gartzke; and Iran’s first attempt in the 1970s is excluded from Singh and Way (2004). Germany and Japan’s World War II programs are not included in the latter data set either.

Complete agreement between Jo and Gartzke and Singh and Way should not be expected, as Jo and Gartzke attempt to code decisions made by officials, whereas Singh and Way look at the capabilities of states. Variation in coding and data sets can be good if the coding rules and decisions are matched up with the appropriate theories. Still, if coding is based on capabilities, as it is for Singh and Way, then Japan and Germany might very well be coded as “exploring” an option, at least for the time before they signed or ratified the NPT, if not today; the histories of these two nuclear-capable states are at least ambiguous, and some historical studies and schemes for coding nuclear status list both states as “hedging” rather than “not interested” (Mackby and Slocombe 2004; Campbell and Sunohara 2004; Levite 2002/2003). Yet, if coding is based on decision making (Jo and Gartzke 2007), Argentina and Brazil probably should not be coded as pursuing nuclear weapons. Both developed uranium enrichment programs, yet the Brazilian Navy’s enrichment program was originally directed at fueling nuclear power plants for submarines (Barletta 1997), while Argentina’s was intended for civilian nuclear power plants (Carasales 1999; Hymans 2001).

As can be seen in Table 1, Meyer’s (1984) data set, coded to end in 1980, missed many programs—including Pakistan, Brazil, Iraq, Iran, Argentina, Sweden, Yugoslavia, Taiwan, and Romania—primarily because of lack of data at the time. While we presumably have somewhat more confidence in some of these estimates now, new information could change these historical or current proliferation data sets

at any point. For example, how should we code Syria, in light of the evidence that a covert nuclear facility was destroyed in 2007 (Albright and Brannan 2008)? Even if we had reliable and complete information on states' nuclear programs, however, the ambiguity about states' programs is not confined to our perceptions and estimates; it is inherent in the phenomenon itself. States do not necessarily make decisions "to go nuclear"; often, the meanings of the programs are constructed by multiple actors within each state as the program progresses. The sociology of nuclear proliferation argues that these programs are better understood by examining the broader political and historical discourses of which nuclear science is a part (Flank 1993/1994; MacKenzie and Spinardi 1995; Abraham 1998); proliferation is a contingent historical phenomenon, not the inevitable result of the workings of the pressures of the international system.

India is often offered as an example of what appears externally as ambiguity but internally might be better characterized as ambivalence. Itty Abraham (2006) argues that India developed a nuclear program as part of a grand modernization scheme with the ultimate end of legitimating its role as an independent, autonomous, self-sufficient postcolonial nation. To this end, India remained "ambivalent" about its nuclear program. It adopted the seemingly contradictory stances of leading the global charge for a robust nonproliferation regime and championing the advances of its nuclear program, to the exclusion of all mentions of the foreign aid it received in its development. "Ambivalence," writes Abraham, "is a permanent feature of the nuclear condition . . . a recognition of the inability to wholly control nuclear events" (2006, 56). As nuclear programs evolve, bureaucracies are established and independent political interests become entrenched. Beyond organizational pressures, however, ambivalence is a result of the awareness that the potential uses for and demands of nuclear science can change over time. The role that the program plays as an expression of state identity, too, can change as the relationship between the state and larger normative-ideational structures changes. Along the road to nuclear weapons, there are many potential paths; the meaning of specific nuclear-technology programs is therefore contingent, not set. Consequently, it is not simply that scholars or the CIA must develop estimates of ambiguous "intent" on the part of potential proliferators; the process of proliferation itself may often be better understood as "ambivalence" inside the governments involved.

Comparing Methods and Results

In addition to the difficulties of defining proliferation, there are also methodological problems with the second wave of quantitative proliferation studies that create further questions about the comparability and validity of their conclusions. Both Singh and Way (2004) and Jo and Gartzke (2007) have multiple stages of proliferation, but they deal with them in very different ways and adopt very different options for analysis. Contained in their methods are implicit assumptions that make little

substantive sense. In Singh and Way's analysis, even states that have no nuclear program are included in their models of nuclear acquisition. Yet, states cannot acquire weapons without first pursuing them,² and it seems unlikely that states will not pursue nuclear weapons without exploring the option first, even if only for a short period of time. Jo and Gartzke effectively count each and every year a state has nuclear weapons as a separate decision to go nuclear, thereby placing excessive weight on states that went nuclear very early. We explore the results of changing these assumptions below. We use these two data sets to demonstrate our first three critiques of the literature: the problems with existing codings of nuclear programs, the lack of proper proxies for important variables, and the need for tighter correspondence between methodologies and empirical questions.

Because of their finely grained coding, Singh and Way look at four different potential transitions. Singh and Way have possible transitions from no nuclear-weapons program to exploration (0 to 1) or pursuit (0 to 2), from exploration to pursuit (1 to 2), and from pursuit to acquisition (2 to 3). States that cancel their nuclear programs revert back to 0, with the exception of India, which moves from acquisition (3) in 1974 to exploration (1) in 1975. Because of left-censoring, the United States never enters any analysis, the USSR only enters into the analysis for acquiring nuclear weapons, the United Kingdom enters only for moving to pursuit, and Japan and Germany's World War II programs never show up. Libya, Pakistan, and China move directly from no nuclear-weapons program to pursuit, skipping exploration. The coding that excludes these programs is potentially problematic, since it removes five of the twenty-six possible cases because of left-censoring and three more from the analysis of exploration that moves to pursuit.

Singh and Way try two different tacks for analysis: first, a hazard analysis of their three different levels, then a multinomial logit. A hazard model for each level (1, 2, 3) with the pool restricted to states at the previous level of development (0, 1, 2) is the most empirically relevant way to analyze these data since it models each level as a prerequisite for the next. We drop all states not already in the first stage (exploring a program) from an analysis of the second stage (pursuit of nuclear weapons) while adding a year of exploration for Libya, Pakistan, and China before they began pursuing; this drops the number of data points from 5,578 to 250.³ We further correct for time by setting the first year of exploration as the first year of being at risk for pursuing nuclear weapons;⁴ once these changes are made, the only variables that remain significant are related to gross domestic product (GDPpc and GDPpc squared, jointly), dispute involvement, and the percentage of democracies. Dropping all states that were not already pursuing from the analysis of nuclear acquisition takes the number of data points from 5,784 to 210, making all variables except industrial threshold statistically insignificant; the data are so sparse that the algorithm does not even converge correctly.⁵ Moving to a more empirically relevant methodology thus undercuts a number of the results. Finally, Singh and Way omit entirely measures of status or any effects that the nonproliferation regime might have; while there are good methodological reasons

for not simply putting in measures of NPT membership or systemic acceptance (see below), eliminating them entirely from the analysis is problematic.

Jo and Gartzke, by contrast, only have two stages: nuclear-weapons programs and nuclear-weapons possession. Unlike Singh and Way, they disregard relatively poorly documented suspected programs, such as those in Algeria, which did not progress very far. While this choice is consistent with their coding rules, they nevertheless do exclude a number of empirically relevant confirmed nuclear-weapons programs such as Libya, Australia, Taiwan's second attempt, and Switzerland. Moreover, their regressions end much earlier than those of Singh and Way (1992 vs. 2000).⁶

Jo and Gartzke's methodology also differs from that of Singh and Way. They perform a probit analysis on the complete data set to determine the effects of the independent variables on starting a nuclear program. However, unlike in Singh and Way, states in Jo and Gartzke do not exit the risk pool for starting a nuclear program once they begin; they remain in the pool the entire time. Consequently, the results are biased toward programs that exist for a longer period of time; the United States, for example, is coded as having a nuclear program for fifty-two years, while North Korea only contributes eleven toward the algorithm.

If Jo and Gartzke's analysis is corrected to eliminate this bias from the risk pool, several very important substantive results change. In Table 2, we report the first differences that result for variables that were significant in any of the three models. We replicate their model for states starting nuclear programs in Model 1, then run the same model while only counting the first year of nuclear acquisition rather than every year in Model 2, then substitute the Singh and Way dates for exploration in Model 3.⁷ *Economic capacity* flips sign and becomes significant in both Models 2 and 3. *Diffusion* becomes negative and insignificant in Models 2 and 3. We suspect that this is due to improper variable specification; their variable is simply the log of the number of years since 1938, not a measurement of diffusion based on direct connections to potential suppliers (as some of the articles in this issue address). *Nuclear threat* becomes insignificant in Models 2 and 3. In Model 3, the effects of being a *Major power* drop, and being a *Regional power* is no longer significant. However, we believe that this change is in part due to mismeasurement, because rather than a measure of prestige, the measures of status used here are based on material power.

Jo and Gartzke usefully grapple with the notions of norms as potential constraints on proliferation. They find that NPT membership (whether a state has ratified the treaty) "decreases the likelihood of having nuclear weapons programs," not a surprising finding; but they also find that their NPT-system variable (the proportion of NPT joiners to the total number of states in the world) does not produce a statistically significant effect on nuclear proliferation and conclude that this "indicates that the NPT has not curbed proliferation incentives since the 1970s" (2007, 179). This last finding, however, is neither robust to alternative measures nor likely to represent a causal relationship. In Model 2 of our tests, the NPT-system variable is positive and significant but becomes negative and insignificant in Model 3. The NPT-system

Table 2
Percentage change in the corrected likelihood of a
Nuclear Program (Jo and Gartzke methods, bold = $p < .10$)

Variable	Model 1: Replication	Model 2: Onset	Model 3: Singh and Way Onset
Latent production capability	142.1%	110.7%	83.4%
Economic capacity	0.7%	-7.2%	-10.8%
Diffusion	61.8%	-5.5%	-3.4%
Conventional threat	246.2%	298.1%	197.4%
Nuclear threat	-81.0%	19.2%	86.5%
NPT membership	-81.1%	-60.3%	59.7%
NPT system	23.2%	45.9%	-36.1%
Major power	1125.8%	1847.5%	847.3%
Regional power	794.9%	378.2%	108.4%
<i>n</i>	4697	4273	4118

effect measured in this manner, however, may just reflect the generally increasing trend for more states' joining the treaty regime through 1992.⁸ Moreover, the Gartzke and Jo NPT-membership coding fails to distinguish between states that join the NPT to cement a bargain with regional rivals (where we would expect to see an effect), cheater states that join the NPT with existing nuclear-weapons programs, such as Taiwan, Iraq, and Libya, and states that join as a part of joining the world order and have no intent of ever developing a civilian nuclear infrastructure, much less nuclear weapons, such as Bhutan, Papua New Guinea, and Fiji. We should not, then, conclude that NPT status and norms are unimportant in curbing nuclear proliferation, but rather, that they have not been properly tested yet.

Finally, the sparsity of the data for nuclear acquisition is particularly troubling. For example, although the official *n* is 250 and 210 for pursuit and acquisition, respectively, in Singh and Way, in effect, there are only 21 unique subjects with 15 failures for moving from *explore* to *pursue* and only 15 subjects and 9 failures for moving from *pursue* to *acquire*. At these levels, perhaps something such as qualitative comparative analysis, which uses fuzzy logic to explore complex interactions between different variables in data sets with a low number of observations, might be more appropriate (Ragin 2000) than traditional statistical analysis.

Progress in Prediction

The articles in this issue further this initial work by focusing on the supply and regulation of nuclear technologies (Kroenig 2009; Fuhrmann 2009) and the effects of the spread of nuclear weapons on the propensity of states to get into conflicts (Gartzke and Jo 2009; Rauchhaus 2009; Beardsley and Asal 2009; and Horowitz

2009) as well as how nuclear weapons affect the dynamics and net outcomes of such conflicts (Beardsley and Asal 2009; Horowitz 2009).

Methodologically, these articles have much to recommend them. They increasingly take into account strategic interaction and endogeneity in various ways, through testing to see if selection into crises affects their conduct and outcomes (Beardsley and Asal 2009; Horowitz 2009) and by using instrumental variables for proliferation (Gartzke and Jo 2009). They carefully include sophisticated notions of the effects of time through splines (Gartzke and Jo 2009; Horowitz 2009), Taylor-series approximations (Beardsley and Asal 2009), autoregression (Fuhrmann 2009; Rauchhaus 2009), or hazard models (Kroenig 2009; Beardsley and Asal 2009). These are important methodological innovations. At the same time, these articles replicate some of the errors of earlier work. We discuss how our first four critiques apply to the articles below, then turn to the curious case of Kargil to elaborate on the fifth.

Proliferation Determinants

Kroenig (2009, this issue) is the only author who directly extends on the proliferation-determinants literature, adding a potentially significant cause of proliferation acquisition: sensitive nuclear assistance. Kroenig's coding of whether states received sensitive assistance or not (and from whom) is a significant improvement over past efforts to determine the supply-side factors influencing proliferation. Kroenig fixes both of the major methodological problems present in previous pieces, counting each nuclear-weapons acquisition only once, and including a censored model with only states that are already pursuing a program. It is worth noting that in the censored model, similar to the results above, very few of the independent variables are significant. In this light, it is particularly important to note that sensitive nuclear assistance is still highly significant; commendably, Kroenig tests his model against both earlier data sets as well as the consensus data set for this issue for robustness, although he does not report the full results.

What Kroenig does not highlight is also important: in robustness checks, he finds that one particular type of civilian nuclear assistance (a state's first reactor) and NPT membership decrease the probability that a state will acquire nuclear weapons. The notion that the spread of civilian technology will inevitably lead to more proliferation has received a lot of attention, primarily because of a few important cases such as India, Iran, and North Korea. Furthermore, Fuhrmann (2009, this issue) argues that defining civilian cooperation more broadly as nuclear cooperation agreements (NCAs) does increase the probability of exploration, pursuit, and acquisition. Kroenig offers an important counterpoint here in the form of evidence that both the NPT and some civilian technologies do decrease the likelihood of weapons proliferation, providing some evidence that the bargain of the NPT is holding, although our caveats regarding NPT membership apply here as well.

Kroenig's main findings are useful but are hardly counterintuitive. He concludes that "states that are better able to produce nuclear weapons, due to either international

assistance or domestic capacity, are more likely to do so.” (p. 162). It is valuable to know that receiving sensitive assistance has been important for new proliferants’ programs, but it is not surprising, to give the most extreme example, that states that received bomb designs from other states were more likely to develop a bomb than states that did not receive such designs.

While Kroenig focuses primarily on the effects of sensitive assistance, Fuhrmann adds to our knowledge about who gives what kinds of nuclear assistance to whom, finding that states make agreements with allies and with other states that share the same enemies. Fuhrmann, uniquely among these articles, breaks down his analysis into several time periods, commendably allowing for greater scrutiny of his results and giving a more accurate picture of the evolution over time of assistance. Fuhrmann’s finding that membership in the NPT actually decreases the probability that a state will receive civilian NCAs is an interesting and potentially troubling finding, given that this evidence seems to run counter to one of the major incentives for states to join the NPT.

This result, however, while statistically sound, appears to be substantively misleading: it is likely to be an artifact of the widespread growth and legitimacy of the NPT, since virtually all states, including many that have no reason to sign such nuclear cooperation agreements, have signed the treaty. Fuhrmann’s NPT result is strictly applicable to the period 1992 to 2000; this outcome, then, likely comes from comparing the small number of states outside the treaty that generally are interested in nuclear technologies of all kinds but did not wish to submit to the demands of the NPT with the vast majority of states (84 percent in 1992) within the NPT, many of which have no interest in nuclear technology of any kind. In short, another interpretation of Fuhrmann’s finding is that NPT-member states that do not wish to acquire nuclear technologies are less likely to get nuclear cooperation agreements than states that wanted nuclear technology before they signed the NPT or the small number of states (India, Pakistan, and Israel) that have remained outside the treaty. This is not a surprising finding. Fuhrmann also finds that “countries exploring nuclear weapons are 107 percent more likely to receive nuclear aid for peaceful purposes” than are countries not exploring a nuclear-weapons option (p. 200). This, too, appears to be a startling finding (albeit one that decreases over time as well). Yet, again, it seems to be caused by comparing states that are exploring nuclear-weapons programs (all of which have civilian programs as well) with states that have never even explored a nuclear-weapons program, a large number of which never had nuclear-power programs at all.

Proliferation Consequences

The second major theme of this issue is the consequences of nuclear proliferation. The previously existing scholarly debate on this subject is both theoretically rich and vigorous but deeply unresolved, in part because it revolves around making predictions about a highly uncertain future. The nuclear optimists (Mearsheimer 1990,

1993; Waltz 1990; Karl 1996; Sagan and Waltz 2003, Ch. 1, 3-4), following realist theory, envision a “perpetual nuclear peace,” maintaining that states and statesmen behave with extreme rationality and will therefore be deterred from using nuclear weapons in crises or wars against other nuclear states. They argue that overcoming preventive war incentives, developing second-strike capabilities, and developing command and control systems are easy. Nuclear pessimists (Feaver 1992; Miller 1993; Sagan 1993, 1994; Feaver, Sagan, and Karl 1997; Sagan and Waltz 2003, Ch. 2-3, 5; Sagan 2004) use organization theory and maintain that nuclear weapons are not controlled by states, but rather, by normal fragile humans in imperfect organizations. Their exploration of nuclear history has uncovered much evidence suggesting that there were close calls to nuclear war during the Cold War, and they predict, therefore, that new proliferators, especially those with poor command and control systems, will be particularly dangerous nuclear-weapons states.

There have been surprisingly few statistical studies, however, designed to determine how common nuclear “near accidents” are or to learn how the acquisition of nuclear weapons influences a states’ propensity to engage in conventional or subconventional war. The exceptions focus on the success rate of extended deterrence and on the effects of nuclear weapons in regional settings. Paul Huth, for example, finds that nuclear weapons increase the likelihood of successful deterrence of attacks on one’s allies (Huth 1988, 1990). S. Paul Kapur combines case-study research and statistical methods in his study of the effects of nuclear weapons in South Asia and finds that nuclear weapons have led to an increase in the likelihood of MIDs between India and Pakistan (Kapur 2007). But we lack broader studies of the consequences of nuclear proliferation on different states’ war and crisis behavior.

The remaining articles in this issue help fill that gap and make significant contributions. They turn the dependent variable in the causes of proliferation studies—a state’s nuclear-weapons acquisition—into the independent variable to assess the consequences of proliferation. Unfortunately, although these articles use the same dates for when individual states acquired their first nuclear weapon, increasing comparability, it is unclear what is the theoretically appropriate threshold in a state’s nuclear-weapons capabilities for nuclear weapons to have an effect on its own or its rivals’ calculations and behavior. Even if the correct date of a state’s first nuclear weapon could be definitively determined (which is not the case today for many new proliferators), should we expect to see deterrent or compellent effects when a state is close to getting a weapon, when it has one weapon, after it has acquired a small arsenal, or only once it has a secure second-strike capability? An opaque proliferating state without a useable nuclear weapon yet might still act more belligerently to encourage other governments to back down in crises; similarly, different adversaries to suspected proliferators may react quite differently to challenges if they have different intelligence estimates of the opponent’s nuclear capabilities.

Take the secret Israeli nuclear program, for example: some scholars date Israel’s nuclear-weapons capability from a covert “decisive test” performed in November

1966, while others claim a crude weapon was first assembled in or just on the eve of the 1967 war; yet, the United States thought that Israel did not have nuclear weapons as late as 1969 (Cohen 2000, 112-14; Montgomery and Mount 2006, 18; Richelson 2006). When should one expect, therefore, that the USSR or Arab states would be influenced by a possible Israeli nuclear weapon? Similar difficulties apply to other states. India tested its "Peaceful Nuclear Explosive" in 1974, which apparently was not easily deliverable; according to some accounts, weaponization was not completed until 1988, and this capability was not definitively demonstrated until the May 1998 nuclear tests. Dates for Pakistan's capabilities range from 1987 to 1990, although again, the May 1998 test date could also apply. North Korea is even more enigmatic; the amount of plutonium that the North Koreans had after reprocessing a limited number of fuel rods by 1992 may not have been enough for even a first weapon; further reprocessing in 2003 definitely gave North Korea a sufficient amount for a few weapons, but this capability was not demonstrated definitively until October 2006. The ambiguity of the status of the arsenal of South Africa in the 1970s and 1980s (Albright 1994, 43; Montgomery and Mount 2006, 10) and the inherited Soviet weapons in Belarus, Kazakhstan, and Ukraine in the 1990s are further cases in point. Even using a first test date is ambiguous; how should the suspected Israeli or South African test in 1979 or the October 2006 North Korean test (which many analysts believe was a failure) be coded?

A second common methodological problem in the consequences of proliferation literature has been how to deal with selection effects and endogeneity. If nuclear states are found to win more disputes than nonnuclear states, is that because they are more powerful or because they cautiously enter only into disputes that they are likely to win? If states that acquire nuclear weapons engage in more conflicts than non-nuclear states, are they really more conflict-prone because they have nuclear weapons, or do they have nuclear weapons because they are conflict prone? The authors contributing to this special journal issue make important progress in dealing with the selection effects and endogeneity problems, but they are less successful in dealing with the difficulties of matching the coding of independent variables to the theory being tested. Some of their findings are also, as was the case with the causes of proliferation articles, quite obvious and unlikely to overturn conventional wisdom.

Gartzke and Jo's article (2009, this issue) is a significant and creative contribution, assessing the effects of the possession of nuclear weapons on militarized dispute (MID) onset and escalation. They find two contradictory effects: while actual nuclear-weapons possession is found to increase the likelihood of initiating a dispute, when substituting an instrument (effectively, the probability of nuclear-weapons possession) instead to correct for possible endogeneity, they find no effect.⁹ Gartzke and Jo suggest that "nuclear weapons have no net effect on dispute propensity" because some nuclear states become more aggressive and others more cautious, and these effects wash each other out. Moreover, they suggest that when facing a nuclear armed state, nonnuclear target states acquiesce before a dispute reaches the

dispute stage, but nuclear states escalate their demands such that new disputes still occur. Again, the net effect of nuclear-weapons acquisition on dispute propensity will be minimal. This is an intriguing finding, and Gartzke and Jo's explanation is one plausible potential cause of the results. We suggest two others. Possession of nuclear weapons should deter disputes with nonnuclear states, but may not if (1) a nuclear taboo exists and both actors believe the dispute is below the threshold of plausible nuclear use (Paul 1995; Tannenwald 2007) or (2) the nonnuclear initiator wishes to demonstrate that it will not be deterred by nuclear weapons, as Stalin did at the beginning of the Cold War (Holloway 1996).

Gartzke and Jo (2009, this issue) have an additional intriguing finding—that “states with nuclear weapons are more likely to receive diplomatic missions from other states than states without nuclear weapons.” They interpret this result as evidence of an increased diplomatic influence that nuclear weapons may bring to a state. This interpretation, however, is problematic for two reasons. First, while having a higher number of connections to other states may result in increased prestige or influence, whether this is the case depends heavily on whom you are connected to and whether the connection is a positive one or seen as a necessary evil (Hafner-Burton and Montgomery 2006; Hafner-Burton, Kahler, and Montgomery 2009). Second, Gartzke and Jo conflate the legal status of official diplomatic recognition with the political influence—arguing that “nations with demands need to be recognized before those demands can be satisfied”—despite the fact that negotiations commonly exist between states that do not have formal diplomatic recognition of each other. In short, diplomatic recognition, political influence, and international esteem are three very different things, and it is difficult to avoid the impression that this measure of “diplomatic influence” was chosen because it was readily available in the COW data set. Gartzke and Jo (2009, this issue) usefully focus our attention on the connection between nuclear weapons and diplomatic influence, but they have not, contrary to their claims, demonstrated that “proliferators appear to prosper by becoming more influential diplomatically.”

Beardsley and Asal's article “Winning with the Bomb” nicely follows on their early work on the effect of nuclear weapons on crisis outcomes, using the ICB data set (Asal and Beardsley 2007). The earlier work demonstrated that as the number of nuclear powers increases, the violence level of a crisis decreases. Beardsley and Asal's new article demonstrates that states that have nuclear weapons have a bargaining advantage as well; the probability of victory (defined as either outright victory or satisfaction with a compromise settlement) in a crisis increases by 14 percent when a nuclear state faces a nonnuclear opponent in any crisis and by 30 percent in “high salience” crises. Moreover, Beardsley and Asal (2009, this issue) find that, controlling for other factors, the crises that nuclear-weapons states have with nonnuclear states are significantly shorter, on average, than are crises between two nonnuclear states. They conclude that “nuclear weapons provide more than prestige, they provide leverage. They are useful in coercive diplomacy, and this must be central to any explanation of why states acquire them” (p. 296).

On further examination, however, the evidence for this conclusion is less impressive than Beardsley and Asal acknowledge. Is a 14 percent increase in the probability of winning an international crisis sufficiently high for statesmen to justify an investment that costs millions of dollars and runs the risk of preventive war and potential opprobrium from the international community? The higher “victory” percentage in high-salience asymmetrical nuclear-weapon cases is more impressive, but more than half of those victories are actually compromise settlements.¹⁰ Moreover, the Beardsley and Asal salience measurement, while a commendable attempt to measure the relevance of nuclear weapons to a crisis, is likely to be endogenous to the conflict itself, since the presence of these weapons affects both the level of violence and the stakes involved. Beardsley and Asal’s salience definition, including both actual military violence and threats of grave damage, strangely excludes cases in which no violence occurred or when only regional threats existed, even if nuclear weapons were highly relevant. Their coding thus leaves out many of the Cold War and post-Cold War crises in which nuclear-weapons use was contemplated or in which threats to attack suspected nuclear-weapons production facilities were made, including the 1954 Dien Bien Phu crisis, the 1961 Berlin crisis, the 1969 Ussuri River crisis, the 1987 Brass Tacks crisis, the 1993–1994 crisis over the North Korean nuclear program, and the 1999 Kargil crisis (see below).

Finally, Beardsley and Asal include a robustness check to ensure that the results are not a function of nuclear states’ selecting themselves into crises, which implies that nuclear states are actually more likely to start crises in the first place. This offers empirical evidence that nuclear states are more likely to enter into crises—even with other nuclear states. Yet, it is not clear that this finding is robust across different codings of nuclear status; indeed, an earlier iteration of the same work with a much broader data set of nuclear-weapons states demonstrated that nuclear states were *less* likely to enter into crises.¹¹ A logical extension of this work would be to study how the outcomes of asymmetrical crises influence the likelihood of further nuclear proliferation; if nuclear states win asymmetric fights, how often do the losers then start nuclear programs?

Rauchhaus’s article—“Evaluating the Nuclear Peace Hypothesis” (2009, this issue)—tests a long-standing proposition in studies of nuclear-deterrence theory and proliferation: the stability–instability paradox. This paradox holds that while symmetrical possession of nuclear weapons may prevent all-out war between two states, it may also increase the likelihood of low-level military conflicts, because one or both of these states might believe that the other government would fear escalation and accept defeat rather than move to higher levels of violence. The logic of the stability–instability paradox led analysts during the Cold War to worry that the USSR might be tempted to use military force against U.S. allies after Moscow had acquired nuclear weapons or a position of nuclear parity. If both governments in a nuclear dyad believed that the other would back down rather than escalate further in a conflict, however, this could lead to a dangerous game of chicken with unpredictable consequences.

The value of Rauchhaus's essay is not in its statistical finding that "when two states have nuclear weapons, the negative coefficient indicates that they are less likely to go to war with one another." We knew that already, and that result was perfectly predictable given Rauchhaus's historical data. Indeed, as he notes, "the data set does not include any instances of two nuclear states' engaging in war with one another. Thus, there is a perfect correlation between the absence of war and dyads in which both states possess nuclear weapons." The value of the article is its sophisticated statistical test of whether two nuclear-armed states are more likely to use threats and military force at lower levels of violence against each other once both have nuclear weapons. Rauchhaus's article provides substantial proof of this proposition: MIDs, the use of force in MIDs, and fatal MIDs are all more likely to occur in symmetrical dyads of nuclear states. Moreover, Rauchhaus carefully controls for potential selection effects and provides a robustness check of his findings by using an alternative coding of the nuclear-weapons acquisition date variable. However, the methodology used here (looking at undirected dyads) does not give as precise a measurement of crisis behavior as directed dyads would, since we cannot tell which state in an asymmetrical dyad started the crisis.

Horowitz's (2009, this issue) article, unlike most of the literature on proliferation, takes into account the possibility that states may change their behavior over time because of learning and carefully assesses which state initiates and reciprocates each dispute. His valuable article finds that "new" nuclear states are slightly more likely to have their disputes reciprocated (and are much more likely to reciprocate others' initial moves) than are nonnuclear states, but that over time, this diminishes to the point at which experienced nuclear states are less than half as likely as nonnuclear states to have their disputes reciprocated. "Mature" nuclear states are also slightly less likely to reciprocate others' disputes than are nonnuclear states.

Horowitz robustly tests different codings for nuclear acquisition, finding that including nuclear inheritors (Belarus, Kazakhstan, and the Ukraine) does not affect the results. Vicarious learning, a particularly difficult-to-measure but potentially important concept, is partially addressed in a robustness test. Presumably new nuclear states enter a world in which all states have learned something about nuclear disputes, so their learning curve might be less steep. Future research could attempt to model vicarious learning as a diffusion process that occurs proportional to the strength of network ties between states (e.g., the United Kingdom is likely to learn lessons from the United States but not from the Soviet Union; Iran might learn more effectively from Pakistani experience with nuclear weapons than from the Israeli or U.S. experience).

These are valuable articles presenting intriguing findings. Further research will be needed, however, to assess whether the causal mechanisms suggested in these studies are actually driving the results. More statistical work will also be needed, as noted above, to provide more robust confidence in the statistical findings. Future work, however, must be more careful than past research in taking into account disconfirming evidence as noted below.

The Curious Case of Kargil

Statistical studies are a useful, indeed essential, method of testing social-science theories about how the world works, and the articles in this issue have clearly helped generate new insights about the sources of nuclear proliferation and its consequences. The use of quantitative research methods, however, too often leads scholars to ignore important disconfirming evidence or to treat it as “measurement error.” Such disconfirming evidence should instead lead to questioning about the accuracy of the initial test of a theory or research about why there may be exceptions to the rule. Consider, for example, the curious case of Kargil.

The Kargil war of 1999 is curious in two ways: first, because it offers apparently damning evidence against both the democratic-peace theory and the nuclear-peace theory, and second, because quantitative scholars on both subjects have largely ignored the implications of the Kargil war. The war occurred in the spring of 1999, when Indian armed forces discovered that Pakistani Northern Light Infantry soldiers, disguised as *Mujahideen* guerilla fighters, had crossed the Line of Control in Indian-held Kashmir and had taken up fortified positions in the mountains above the town of Kargil. India and Pakistan were clearly nuclear-weapons states after their 1998 nuclear tests. The 1999 conflict should be coded as an interstate war, since the most reliable estimates of the fatalities in Kargil, from the Kargil Review Committee (2000), set the number of Indians dead at 474 (p. 23) and the “lowest estimate of regular Pakistani Army casualties is 700 killed” (p. 98), for a conservative minimum of 1,174. This is well above the 1,000 battle deaths for war criteria used in the COW and other data sets. Kargil is an exception to the democratic-peace theory as well because India and Pakistan score a +9 and +7, respectively, in the Polity IV data set for 1999 (Marshall and Jagers 2003). Still, all the articles in this journal issue have ignored this fact and its implications for the robustness of the tests about the consequences of proliferation. A leading work in the democratic-peace literature does at least address the Kargil war: Russett and Oneal recognize Kargil as an apparent exception to the theory, but they incorrectly suggest that it may not count as a war because many of the deaths were “Islamic guerrillas, not regular Pakistani troops” (Russett and Oneal 2001, 48).

It would be better to analyze apparent exceptions to the rule to develop alternative theories and hypotheses for future research. In the case of Kargil, the Pakistani decision to send troops into Indian-held Kashmir was apparently made by military leaders, with minimal involvement of the democratically elected prime minister (Bennett-Jones 2002, 102-3; Sagan 2009). This leads to a testable hypothesis that military-run governments or civilian governments with inadequate operational control over the military may behave differently regarding nuclear weapons and war initiation. The fact that the Pakistani military also opposed the withdrawal from Indian-held Kashmir and the cease-fire settlement in 1999 also suggests that such governments may behave differently regarding nuclear weapons and escalation in

crises and conflicts (Sagan and Waltz 2003, 97; Musharraf 2006, 96). Future qualitative and quantitative research will be needed to test these hypotheses.

Implications

This revival of the quantitative study of proliferation is a welcome advancement in nuclear-proliferation studies. Some of the questions and debates in this subfield can only be fully tested when using statistical methods. Others will require mixed methods, combining historical case studies, deductive reasoning, and quantitative research. These articles constitute significant progress in our understanding of the causes of nuclear proliferation, the propensity of states to initiate disputes, and the outcomes of those disputes. As with much good research, they provoke as many questions as they answer. Here, we have noted some of the puzzling and potentially contradictory results, offer some tentative explanations, and propose an agenda for further research.

The findings of the articles on the causes of proliferation usefully focus attention on the technology assistance to potential proliferators. The next wave of research will need to disaggregate the kinds of assistance much more than Fuhrmann and Kroenig do in their articles. While it is not surprising that designs for nuclear weapons make proliferation more likely, it would be helpful to know, for example, whether uranium enrichment or plutonium reprocessing were more commonly abused by covert proliferators, including states that started weapons programs and did not complete them. It would also be helpful to differentiate between NCAs signed and those that were actually implemented, to avoid giving inappropriate weight to cases, such as the German NCAs with Brazil, that were not in the end honored and thus could not contribute to nuclear-weapons programs or acquisition (Squassoni and Fite 2005). Disaggregation of different types of NCAs would also be useful; nuclear assistance with medical devices and light-water power reactors are unlikely to aid nuclear-weapons programs, while other types of civilian assistance are much more contentious. The relationship between the NPT, covert nuclear-weapons programs, and civilian nuclear assistance also needs to be examined more thoroughly before jumping to negative conclusions about the treaty. Further historical and statistical study is required to unravel some of these puzzles: Why have different states joined the treaty, and how has the treaty influenced their subsequent behavior (Sasikumar and Way 2007)? How do different technologies affect proliferation? How much does sensitive assistance decrease the time to acquisition of nuclear weapons? With the potential for a dramatic increase in the use of nuclear power expected in the future, these questions have immediate policy relevance.

The articles on the consequences of proliferation also point to the need for further research. One important research agenda would focus more on initiation of conflicts to learn if possession of nuclear weapons makes states more likely to initiate or be

the target of disputes. In their raw results, all the articles on proliferation consequences suggest that possessors are more likely to be in crises, although only two have specific results for initiation (Gartzke and Jo 2009, this issue; Horowitz 2009, this issue). Gartzke and Jo posit that increased initiation is not because of nuclear possession itself, but Horowitz's results also suggest that possessors become more likely to initiate disputes with experience. This is a very bleak conclusion for nuclear optimists: while nuclear-weapons states may have their challenges reciprocated less often over time, they may issue more challenges. Future research could usefully focus on whether all states fit into that pattern, since some types of regimes or leaders are more likely than others to be aggressive if they get nuclear weapons. While Beardsley and Asal (2009, this issue) and Horowitz (2009, this issue) have taken important first steps in understanding the net effect of nuclear weapons on conflict and bargaining outcomes, additional research is clearly also required here. If initiation is actually not affected (or increased) by nuclear weapons, this should have an effect on the bargains that these states get, with or without crises or threats of force. Future work in this important area of research could focus less on the overall trends and more on the outliers: for example, are there certain states or certain kinds of regimes that do not become more prudent over time after getting nuclear weapons? Finally, a useful follow-on study regarding the stability–instability paradox would be to combine Rauchhaus's statistical methods and historical case studies to examine directly the dynamics of escalation from the threat to the use of force and from uses of force to fatal MIDs to determine whether the fear of escalation by one side or both determined the governments' behavior or whether other factors were more important. That research could provide insights into how confident we should be about the likelihood that the “nuclear peace” will continue to exist in the future if more states acquire nuclear weapons.

The exclusion of important theories of proliferation (or their inclusion with inadequate measures) also tempers some of the conclusions reached here and requires additional research. Scholars have grappled for a long time with the notions of prestige, norms, social roles, and psychology as motivations for the causes of proliferation. Yet, since Kegley (1980), none of the quantitative articles on the causes of proliferation have successfully incorporated these theories into their analysis. This affects articles both on the causes and likelihood of proliferation and on its consequences, as the reasons for proliferation could potentially have significant effects on how nuclear weapons are used once they are acquired. Difficult-to-measure concepts such as civil–military relations and bureaucratic power and autonomy have also rarely been incorporated into quantitative analyses,¹² or when incorporated, have been measured using existing measures of regime type, which were developed for other purposes. Yet, civilian control of the military and central government control over nuclear scientists could have important effects on both the likelihood of proliferation and its consequences.

Finally, the results that quantitative studies produce do not necessarily make for sound policy advice; what may be true across an entire population does not necessarily make good sense for dealing with individual cases. The finding that civilian nuclear assistance may often lead to proliferation, for example, should not lead scholars to oppose the Clinton and Bush administrations' agreements to give light water-moderated nuclear-power reactors to North Korea, for that policy option is conditional on North Korea's dismantling its more proliferation-prone graphite-moderated reactor and returning its weapons-grade plutonium stockpile to the International Atomic Energy Agency (IAEA). The observation that nuclear-armed dyads are less likely to go to war against each other similarly should not necessarily lead to the prescription that more nuclear states would be better, since the possible destruction of a full-scale war between nuclear powers could outweigh the benefits of decreased likelihood of conventional conflict.

The curious case of Kargil underlines this final lesson for both the democratic peace and the nuclear peace: rare does not mean never. This is less damaging for the democratic peace; as long as democracy decreases war, there is still a net benefit that results from an increase in the number of democratic states. But if nuclear weapons do not eliminate the possibility of nuclear war entirely, their proliferation could lead to a disastrous outcome over time. While nuclear optimists may therefore take some comfort in the results of some of these articles, nuclear pessimists will remind us that unless the probability of nuclear war is zero, the potential consequences of nuclear proliferation may still outweigh the benefits.

Notes

1. MID 3.10 does include Kargil as a dispute that had a maximum hostility level of war but has it as the endpoint of a dispute starting in 1993; see Kapur (2007, 19). COW 3.0 only has wars through 1997, while ICB 8.0 lists Kargil as "serious clashes," not "full-out war."

2. The cases of Belarus, Kazakhstan, and Ukraine are excluded from both data sets, since the temporary location of nuclear weapons on their soil resulted from the breakup of the Soviet Union rather than from a domestic program.

3. Singh and Way footnote this variant in their analysis (2004, 875, note 32); we believe that this is the more empirically relevant question.

4. This corrects for entry and exit times of the risk pool. In Singh and Way's (2004) analysis, a state is at risk of moving to the third stage (pursuing) nuclear weapons from its first date of entry into the system. In this corrected analysis, states enter the risk pool for pursuit of nuclear weapons only in their first year of exploring nuclear weapons.

5. We omit a full analysis here because of space constraints; STATA data sets and do-files for replication are available on the *JCR* Web site.

6. Like the Singh and Way left-censoring, the right-censoring of these data at 1992 is potentially important but beyond the scope of this analysis.

7. We replicate their Table 1, Model 1 (Jo and Gartzke 2007, p. 178). We used Clarify 2.1 (King, Tomz and Wittenberg 2000, Tomz, Wittenberg and King 2003) to analyze these differences; all variables were set to their medians for the baseline except *Latent production capability*, which was set to 5, and increased to their 75th percentile value except for dummy variables and variables whose 75th percentile is still zero (i.e., *Conventional threat*), which were set to 1.

8. The NPT-system variable is correlated 0.9131 with the year; the fraction of states in the system increases rapidly from .47 in 1970 to .84 in 1992.

9. While we critique Jo and Gartzke (2007) for including multiple observations for nuclear possession, in the Gartzke and Jo (2009, this issue) article, this model makes more sense in this context, since it is attempting to produce an instrument that correlates with the presence of nuclear weapons, not just with the initial acquisition. Nonetheless, we find the instrument to be potentially problematic, since probit and logit functional forms can lead to misspecification (see Angrist and Krueger 2001, 80), and important variables such as status are excluded from the instrument.

10. Similarly, in high-salience crises with symmetrical nuclear dyads, Beardsley and Asal record a 21 percent increase in victory compared to a dyad with two nonnuclear states; however, this is entirely because of an increase in compromise settlements, as absolute victory actually decreases.

11. The American Political Science Association (APSA) version of the article (Beardsley and Asal 2007) had a negative coefficient on Nuclear A and a small positive coefficient on Nuclear B. This data set had earlier dates for India (1974), Pakistan (1987), and North Korea (1992) and included Ukraine, Kazakhstan, and Belarus from 1991 until 1996, 1995, and 1996, respectively.

12. One notable exception is Sechser (2004).

References

- Abraham, Itty. 1998. *The making of the Indian atomic bomb: Science, secrecy and the postcolonial state*. London, UK: Zed.
- . 2006. The ambivalence of nuclear histories. *Osiris* 21 (1): 49-65.
- Albright, David. 1994. South Africa and the affordable bomb. *Bulletin of the Atomic Scientists* 50 (4): 37-47.
- Albright, David, and Paul Brannan. 2008. *The Al Kibar reactor: Extraordinary camouflage, troubling implications*. Washington, DC: Institute for Science and International Security.
- Angrist, Joshua D., and Alan B. Krueger. 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 15 (4): 69-85.
- Asal, Victor, and Kyle Beardsley. 2007. Proliferation and international crisis behavior. *Journal of Peace Research* 44 (2): 139-55.
- Barletta, Michael. 1997. The military nuclear program in Brazil. Working paper, Center for International Security and Cooperation, Stanford University.
- Beardsley, Kyle, and Victor Asal. 2007. Winning with the bomb. Paper read at 103rd annual meeting of the American Political Science Association, Chicago, IL, August 30 to September 2.
- . 2009. Winning with the bomb. *Journal of Conflict Resolution* 53:278-301.
- Bennett-Jones, Owen. 2002. *Pakistan: Eye of the storm*. New Haven, CT: Yale University Press.
- Campbell, Kurt M., Robert J. Einhorn, and Mitchell Reiss, eds. 2004. *The nuclear tipping point: Why states reconsider their nuclear choices*. Washington, DC: Brookings Institution.
- Campbell, Kurt M., and Tsuyoshi Sunohara. 2004. Japan: Thinking the unthinkable. In *The nuclear tipping point: Why states reconsider their nuclear choices*, ed. K. M. Campbell, R. J. Einhorn, and M. Reiss, 218-53. Washington, DC: Brookings Institution.
- Carasales, Julio C. 1999. The so-called proliferator that wasn't: The story of Argentina's nuclear policy. *Nonproliferation Review* 6 (4): 51-64.
- Chafetz, Glenn, Hillel Abramson, and Suzette Grillot. 1996. Role theory and foreign policy: Belarussian and Ukrainian compliance with the nuclear nonproliferation regime. *Political Psychology* 17 (4): 727-57.
- Cohen, Avner. 2000. Nuclear arms in crisis under secrecy: Israel and the lessons of the 1967 and 1973 wars. In *Planning the unthinkable*, ed. P. R. Lavoy, S. D. Sagan, and J. J. Wirtz, 104-24. Ithaca, NY: Cornell University Press.
- Doyle, Michael W. 1986. Liberalism and world politics. *American Political Science Review* 80 (4): 1151-69.

- Elman, Miriam Fendius. 1997. *Paths to Peace: Is Democracy the Answer?* Cambridge, MA: MIT Press.
- Fearon, James D. 1994. Domestic political audiences and the escalation of international disputes. *American Political Science Review* 88 (3): 577-92.
- Feaver, Peter D. 1992. Command and control in emerging nuclear nations. *International Security* 17 (3): 160-87.
- Feaver, Peter D., Scott D. Sagan, and David J. Karl. 1997. Proliferation pessimism and emerging nuclear powers. *International Security* 22 (2): 185-207.
- Flank, Steven. 1993/1994. Exploding the black box: The historical sociology of nuclear proliferation. *Security Studies* 3 (2): 259-94.
- Frankel, Benjamin. 1993. The brooding shadow: Systemic incentives and nuclear weapons proliferation. In *The proliferation puzzle: Why nuclear weapons spread and what results*, ed. Z. S. Davis and B. Frankel, 37-78. Portland, OR: Frank Cass.
- Fuhrmann, Matthew. 2008. Exporting mass destruction? The determinants of dual-use trade. *Journal of Peace Research* 45 (5): 633-52.
- . 2009. Taking a walk on the supply side: The determinants of civilian nuclear cooperation. *Journal of Conflict Resolution*, 53:181-208.
- Gartzke, Erik. 2007. The capitalist peace. *American Journal of Political Science* 51 (1): 166-91.
- Gartzke, Erik, and Dong-Joon Jo. 2009. Bargaining, nuclear proliferation, and interstate disputes. *Journal of Conflict Resolution*, 53:209-233.
- Gartzke, Erik, and Matthew Kroenig. 2009. A strategic approach to nuclear proliferation. *Journal of Conflict Resolution*, 53:151-160.
- Grillot, Suzette R., and William J. Long. 2000. Ideas, beliefs, and nuclear policies: The cases of South Africa and Ukraine. *Nonproliferation Review* 7 (1): 24-40.
- Hafner-Burton, Emilie M., Miles Kahler, and Alexander H. Montgomery. 2009. Network analysis for international relations. *International Organization*, forthcoming.
- Hafner-Burton, Emilie M., and Alexander H. Montgomery. 2006. Power positions: International organizations, social networks, and conflict. *Journal of Conflict Resolution* 50 (1): 3-27.
- Holloway, David. 1996. *Stalin and the bomb*. New Haven, CT: Yale University Press.
- Horowitz, Michael. 2009. The spread of nuclear weapons and international conflict: Does experience matter? *Journal of Conflict Resolution*, 53:234-257.
- Huth, Paul K. 1988. Extended deterrence and the outbreak of war. *American Political Science Review* 82 (2): 423-43.
- . 1990. The extended deterrent value of nuclear weapons. *Journal of Conflict Resolution* 34 (2): 270-90.
- Hymans, Jacques E. C. 2001. Of gauchos and gringos: Why Argentina never wanted the bomb, and why the United States thought it did. *Security Studies* 10 (3): 153-85.
- . 2006. *The psychology of nuclear proliferation: Identity, emotions, and foreign policy*. Cambridge, UK: Cambridge University Press.
- Jo, Dong-Joon, and Erik Gartzke. 2006. Codebook and data notes for "Determinants of nuclear weapons proliferation: A quantitative model" in *Journal of Conflict Resolution*.
- . 2007. Determinants of nuclear weapons proliferation: A quantitative model. *Journal of Conflict Resolution* 51 (1): 167-94.
- Kant, Immanuel. [1795] 1917. *Perpetual peace: A philosophical essay*, ed. M. C. Smith. London: Allen and Unwin.
- Kapur, S. Paul. 2007. *Dangerous deterrent: Nuclear weapons proliferation and conflict in South Asia*. Stanford, CA: Stanford University Press.
- Kargil Review Committee. 2000. *From surprise to reckoning: The Kargil Review Committee Report*. New Delhi: Sage.
- Karl, David J. 1996. Proliferation pessimism and emerging nuclear powers. *International Security* 21 (3): 87-119.

- Katzenstein, Peter J. 1996. *Cultural norms and national security: Police and military in postwar Japan, Cornell studies in political economy*. Ithaca, NY: Cornell University Press.
- Kegley, Charles W. 1980. International and domestic correlates of nuclear proliferation: A comparative analysis. *Korea and World Affairs* 4:5-37.
- King, Gary, Michael Tomz, and Jason Wittenberg. 2000. Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science* 44 (2): 347-61.
- Kroenig, Matthew. 2009. Importing the bomb: Sensitive nuclear assistance and nuclear proliferation. *Journal of Conflict Resolution*, 53:161-180.
- Lake, David A. 1992. Powerful pacifists: Democratic states and war. *American Political Science Review* 86 (1): 24-37.
- Levite, Ariel E. 2002/2003. Never say never again: Nuclear reversal revisited. *International Security* 27 (3): 59-88.
- Mackby, Jenifer, and Walter Slocombe. 2004. Germany: The model case, a historical imperative. In *The nuclear tipping point: Why states reconsider their nuclear choices*, ed. K. M. Campbell, R. J. Einhorn, and M. Reiss, 175-217. Washington, DC: Brookings Institution.
- MacKenzie, Donald, and Graham Spinardi. 1995. Tacit knowledge, weapons design, and the uninvention of nuclear weapons. *American Journal of Sociology* 101 (1): 44-99.
- Mansfield, Edward D., and Jack Snyder. 1995. Democratization and the danger of war. *International Security* 20 (1): 5-38.
- Maoz, Zeev, and Bruce Russett. 1993. Normative and structural causes of democratic peace, 1946-1986. *American Political Science Review* 87 (3): 624-38.
- Marshall, Monty G., and Keith Jagers. 2003. *Polity IV project: Political regime characteristics and transitions, 1800-2002*. College Park, MD: Integrated Network for Societal Conflict Research.
- Mearsheimer, John J. 1990. Back to the future: Instability in Europe after the Cold War. *International Security* 15 (1): 5-56.
- . 1993. The case for a Ukrainian nuclear deterrent. *Foreign Affairs* 72 (3): 50-66.
- Meyer, Stephen M. 1984. *The dynamics of nuclear proliferation*. Chicago, IL: University of Chicago Press.
- Miller, Steven E. 1993. The case against a Ukrainian nuclear deterrent. *Foreign Affairs* 72 (3): 67-80.
- Montgomery, Alexander H. 2005. Ringing in proliferation: How to dismantle an atomic bomb network. *International Security* 30 (2): 153-87.
- Montgomery, Alexander H., and Adam J. Mount. 2006. Misunderestimation: Explaining US failures to predict nuclear weapons programs. Paper read at 102nd annual meeting of the American Political Science Association, Philadelphia, PA, August 31 to September 3.
- Musharraf, Pervez. 2006. *In the line of fire: A memoir*. New York: Free Press.
- Ogilvie-White, Tanya. 1996. Is there a theory of nuclear proliferation? An analysis of the contemporary debate. *Nonproliferation Review* 4 (1): 43-60.
- Paul, T. V. 1995. Nuclear taboo and war initiation in regional conflicts. *Journal of Conflict Resolution* 39 (4): 696-717.
- . 2000. *Power versus prudence: Why nations forgo nuclear weapons*. Vol. 3, *Foreign policy, security and strategic studies*. Montreal, Quebec: McGill-Queen's University Press.
- Ragin, Charles C. 2000. *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Rauchhaus, Robert. 2009. Evaluating the nuclear peace hypothesis: A quantitative approach. *Journal of Conflict Resolution*, 53:258-277.
- Reiss, Mitchell. 1988. *Without the bomb: The politics of nuclear nonproliferation*. New York: Columbia University Press.
- . 1995. *Bridled ambition: Why countries constrain their nuclear capabilities*. Woodrow Wilson Center special studies. Washington, DC: Woodrow Wilson Center Press.
- Reiss, Mitchell, and Robert S. Litwak, eds. 1994. *Nuclear proliferation after the Cold War*. Woodrow Wilson Center special studies. Washington, DC: Woodrow Wilson Center Press.

- Reiter, Dan, and Allan C. Stam. 2002. *Democracies at war*. Princeton, NJ: Princeton University Press.
- Richelson, Jeffrey. 2006. *Spying on the bomb: American nuclear intelligence, from Nazi Germany to Iran and North Korea*. New York: Norton.
- Russett, Bruce M., and John R. Oneal. 2001. *Triangulating peace: Democracy, interdependence, and international organizations*. New York: Norton.
- Sagan, Scott D. 1993. *The limits of safety: Organizations, accidents, and nuclear weapons*. Princeton, NJ: Princeton University Press.
- . 1994. The perils of proliferation: Organization theory, deterrence theory, and the spread of nuclear weapons. *International Security* 18 (4): 66-107.
- . 1996/1997. Why do states build nuclear weapons? Three models in search of a bomb. *International Security* 21 (3): 54-86.
- . 1999. Rethinking the causes of nuclear proliferation: Three models in search of a bomb. In *The coming crisis: Nuclear proliferation, U.S. interests, and world order*, ed. V. A. Utgoff. Cambridge, MA: MIT Press.
- . 2004. The problem of redundancy problem: Why more nuclear security forces may produce less nuclear security. *Risk Analysis* 24 (4): 935-46.
- , ed. 2009. *Inside nuclear South Asia*. Stanford, CA: Stanford University Press.
- Sagan, Scott D., and Kenneth N. Waltz. 2003. *The spread of nuclear weapons: A debate renewed*. 2nd ed. New York: Norton.
- Sasikumar, Karthika, and Christopher R. Way. 2007. Paper tiger or barrier to proliferation? What accessions reveal about NPT effectiveness. Paper read at the 103rd annual meeting of the American Political Science Association, Chicago, IL, August 30 to September 2.
- Schultz, Kenneth A. 1998. Domestic opposition and signaling in international crises. *American Political Science Review* 92 (4): 829-44.
- Sechser, Todd S. 2004. Are soldiers less war-prone than statesmen? *Journal of Conflict Resolution* 48 (5): 746-74.
- Singh, Sonali, and Christopher R. Way. 2004. The correlates of nuclear proliferation: A quantitative test. *Journal of Conflict Resolution* 48 (6): 859-85.
- Solingen, Etel. 1994. The political economy of nuclear restraint. *International Security* 19 (2): 126-69.
- . 1998. *Regional orders at century's dawn: Global and domestic influences on grand strategy*. Princeton, NJ: Princeton University Press.
- . 2007. *Nuclear logics: Contrasting paths in East Asia and the Middle East*. Princeton, NJ: Princeton University Press.
- Squassoni, Sharon A., and David Fite. 2005. Brazil as litmus test: Resende and restrictions on uranium enrichment. *Arms Control Today*, October.
- Tannenwald, Nina. 2007. *The nuclear taboo: The United States and the non-use of nuclear weapons since 1945*. Cambridge, UK: Cambridge University Press.
- Tomz, Michael, Jason Wittenberg, and Gary King. 2003. *Clarify: Software for interpreting and presenting statistical results* (2.1).
- Waltz, Kenneth N. 1990. Nuclear myths and political realities. *American Political Science Review* 84 (3): 731-45.