Interpretive Quantification

Methodological Explorations for Critical and Constructivist IR

J. Samuel Barkin and Laura Sjoberg, Editors

2017

University of Michigan Press
Ann Arbor

CHAPTER 7

Thinking Counterfactually and with Discipline

Agent-Based Models for Constructing and Deconstructing the Future

IAN S. LUSTICK

The Future: Gaining Traction on a Great Unsolved Problem

The chief purpose of knowledge, according to Niccolò Machiavelli, is to control one's life—not the life one has already lived, but the life to be lived, with all its dangers and opportunities. In other words, the purpose of pursuing knowledge systematically (i.e., the purpose of science) is to solve the problem of knowing as much as possible about the future, a place impossible to visit, filled not only with danger and risk but also perhaps with opportunity. Contemporary social science remains powerfully committed to this ethos. Grant applications promise the prince that science can help effectively defend against potential threats and exploit possible opportunities.

The key difficulty is that the future does not exist and so cannot be directly studied. Neither can it be predicted exactly in all its dimensions. But implicit in the recommendations of Machiavelli, as well as in the studies advanced by policy analysts and in the practices of political scientists who think their work is of value to those who practice politics, is that rational action requires assessments of what will happen that depend on decisions taken in the present; that is, rationality entails assessments of that which cannot be directly examined. Accordingly, a surrogate of some sort must be provided as the target of investigation. These surrogates are stories con-

structed, with more or less self-consciousness, about what the future could look like. To construct these surrogates, Machiavelli used the past and theories about how the world works to construct stories about the future. Machiavelli produced the past he used for this purpose by selective retrieval and narration of various sources about the past, including heavy reliance on Livy and Tacitus. Whether drawing on historians of ancient Rome, cable traffic from relevant embassies, specialized monographs, field research, casual reading of newspapers, or personal prejudices, the traditional and still standard method of both social scientists and policy analysts is to tell stories about how the future could unfold along dimensions selected as relevant or interesting. The technical term for this procedure is *scenario analysis.*

Scenarios are stories, imagined as plausible depictions about what could happen (or would have happened, if attention is directed to retrodiction) and why. Each of these stories is an elaborate counterfactual account. Normally—in the worlds of business and government—a small number (three to five) of such accounts are developed, compared, and analyzed, usually by asking individuals or groups to imagine plausible and interesting futures in a domain in which they are presumed to have some expertise. This process, highly informal but sometimes partially structured, is commonly referred to as BOGSAT (Bunch of Guys Sitting Around Talking). In academia, the sources of these stories can sometimes be explicitly theoretical. For example, a game theorist could describe outcomes by imagining key players interacting according to payoff structures associated with different kinds of games. Aggregate data analysts could accomplish the equivalent task by inferring patterns in the future similar to those measured along relevant dimensions in the past. But since outcomes of interest are commonly too rare to have been observed as categories of aggregate measurement, even aggregate data analysts must use their imagination and judgment to guide the production of their stories about the future.[1] However these scenarios are produced, they constitute, as a set, the basis for analyzing the implications of present policy choices for future outcomes. To talk about how one's actions might affect the future, one must first *construct* an image or a set of images of what the future would be absent one's own contribution. These images, cast dynamically and in emplotted form as "narratives," are thus treated as surrogates for what the analyst (looking forward from the present) hopes that he or she has—a collection of representations of the most plausible futures or kinds of futures that could unfold.

When we direct attention to how these surrogate futures are constructed and deployed as the database for making decisions about probabilities and courses of action, we are engaging in a deconstructive enterprise. In a fundamental sense, this enterprise is the core element of a critical approach. In the context of thinking about the future, constructivism means recognizing that scenarios represent interpretations of the future that are built out of assumptions, theories, purposes, and selected data. A critical approach to the subject matter means identifying and fostering awareness of the limitations of the (interpretive) constructions that result from our engagement with questions of interest about the future. Since claims about events in the present and past involve assessments of the implications of events and patterns prior to them, we cannot logically claim knowledge of the past or present without implying the possibility of knowledge of the future. This does not entail belief in the ability to prophesy what actually will happen in detail, but it does imply the expectation that disciplined analysis will be able to distinguish between credible and less credible forecasts about the ways events may materialize.

In this chapter, I illustrate the use of computer-assisted agent-based modeling as a formal technique for both constructing and critically evaluating our interpretations of the future.[2] Since the assumptions of the model must be explicit and consistent for the computer program to operate, every batch of histories that is produced (i.e., constructed) can legitimately be interpreted as illuminating the implications of the specific assumptions and theories operationalized within the model. Moreover, because every trajectory of the model (i.e., every virtual future) can be subjected to process tracing in exquisite detail, opportunities exist for every future and every batch of futures to be critically assessed—indeed, to be deconstructed into the chains of mechanisms that produced individual outcomes and into the propensities of different theoretical judgments or parameter settings to produce futures with particular features.

To appreciate the advantage of computerized simulation as a generator of accounts of the future, we should first recall that all scenarios—that is, all depictions of the future—are "virtual" in that they exist as or at least are translated from images in the minds of their producers. For even the best analysts, "transition rules" within the domain of their expertise (i.e., the laws of social behavior at all levels of analysis) are, at best, only very partially understood. Thus every story about what the future may hold and every judgment about the relative likelihood of a kind of event under particular conditions must include large doses of arbitrary, even idiosyncratic decisions about what dimensions and variables will dominate, in what combinations, in what sequence, and with what consequences. These are the crutches that human minds require when facing the immense cognitive problem of imagining the implications of what they know or believe about the present for

discerning the future of a complex state of affairs. Contemporary cognitive psychology understands these crutches, these departures from strict Bayesian rationality, as a vast array of heuristics or "psychologics" that trump or simply replace cognitively onerous and informationally demanding processes of evidence-based and probability-based inference.

Which scenarios are produced or endowed with credibility is also affected by political preferences and the structuring of choice, standardly by bracketing a "preferred" scenario with two "extreme" and fairly clearly wrong or distasteful alternatives. Given our cognitive inadequacies and political prejudices, how can moderately credible theory be combined with computerization for systematic understanding of the future? The problem becomes even more daunting when we acknowledge how much of the future is shaped by "accidents" or "random events," that is, by effects of below-the-analytic-horizon events that we can never aspire to know.

We can usefully approach this problem by translating what has been said so far into the language of "counterfactuals."[3] Since the future has not occurred, all claims about it are counterfactuals—conditional statements about something that could or will happen under conditions that have not yet obtained. Since most errors about the future cannot be identified definitively as such until the actual future becomes the past, a critical stance toward claims about the future cannot rest on assertions of truth over falsehood. Such a stance must instead rest on a more sophisticated concept of what is entailed by thinking about the future, that is, on a disciplined approach to the evaluation of counterfactuals. That, in turn, requires being able to think not about the likelihood of individually imagined trajectories within what is assumed to be a normal or near-normal distribution but about differently shaped distributions of trajectories of different types. Needed, in other words, is a map of the space of the future that affords some sense of how uncertainty is distributed within it.

Consider the debate over the catastrophic failure by policy makers to anticipate the financial collapse of 2008. The problem was not that policy makers did not succeed in predicting the exact consequences of the closure of Lehman Brothers but that they failed to understand the shape of the distribution of the state space of the future. From the bursting real estate bubble and the associated collapse of financial markets that inaugurated the Great Recession of 2007–9, we learned that treating an underlying distribution as normal when it is not is immensely dangerous. In their prescient book published four years prior to that event, *The Misbehavior of Markets*, Benoit Mandelbrot and Richard L. Hudson explain that the modern financial theories used by brokers to confidently recommend investment strategies seemed to offer virtually guaranteed returns, which, the authors forecast, would result in economic disaster. The brokers' strategies were based on the expectation that certain extreme fluctuations in market prices were so unlikely that they could not even be considered "rare" and could be relegated to a category of events that were theoretically possible but so improbable as to be disregardable for all practical purposes. Of the (relatively minor) crash of the stock market on October 19, 1987, the authors wrote, "On one day, the Dow plunged 29.2 percent. Something was wrong; The academics said that the fall should not have happened, that it was a once-in-an-eon event."[4]

By looking closely at the "volatility of volatility" in data describing market fluctuations over the previous century, the authors demonstrated that types of events that "should not have happened at all" were better described as implausible, perhaps, and as unpredictable, but certainly not as effectively impossible. The fundamental flaw in modern portfolio theory, the capital asset pricing model, the Black-Scholes equation, and other reigning financial theories was the assumption that the relevant universe of cases has attributes distributed according to a normal, "bell" curve. The underlying distributions driving market behavior are not normal but Cauchian, subject to the regular intrusion of "rare" events of enormous impact that are not common enough to be expected at any particular point in time with any confidence at all but that are plausible enough, given both the shape of the underlying distribution and their magnitude, to warrant taking them into consideration as decisive factors when planning for the future.

The overall point to note for my purpose here is that forecasting in politics means focusing, first at least, on the shape of the distribution of the possible, rather than on whether any particular event or discrete kind of event that is thought to be possible will occur. The strength of the analyses of Mandelbrot and others, such as Nassim Taleb, is that they are able to show emphatically and convincingly that we should expect many more "rare events" to occur across a wide array of domains than our natural inclinations to expect a "bell-shaped" world suggest. The problem is that their conclusions are pitched at such a high level of analysis that their policy prescriptions are extremely limited. Taleb takes an extreme view, that "almost everything in social life is produced by rare but consequential shocks and jumps."[5] But planning is effectively impossible if, as he further claims, "a small number of Black Swans explain almost everything in our world."[6] Indeed, such a view implies not only that planning is impossible but that surprise is inevitable and that protection against maximum loss, rather than organizing to exploit opportunity, is the only rational course of action. But policy makers, including those charged with devising foreign and national

security policies pertaining to a wide array of specific but complex issues, experience considerable variation in the "Cauchiness" of underlying distributions. On occasion, they may even confront domains within which the most important underlying distributions *are* normal. That state of affairs puts a premium on the *a priori* ability of planners to gauge the relative character of the distribution of possible outcomes in their domain of responsibility or with respect to a particular question within that domain.

As noted above, there are good reasons why experts relying on their brains and imagination alone cannot produce a set of scenarios large enough or systematic enough to depict the state space of the future of a complex social system. Yet BOGSAT is still, in one form or another, the standard technique for producing the basis for thinking about the future. To appreciate how agent-based modeling (ABM) can be used to radically improve the way this problem is addressed, let us take a closer look at the distinction between the "actual" future and the space of possible futures.

From any time labeled the "present," we can only imagine or visualize the "future" correctly as a large set of multidimensional trajectories. These trajectories travel through an enormous, though not boundless or totally disorganized, state space of the possible worlds that could evolve from the world that we experience as the present. In other words, looking forward from the present, we can imagine many ways, along many dimensions (some important and some not), in which the actual set of events of our world can vary. Uncertainty about what trajectory will be followed is not only a function of the weakness of our theories. For any complex system, a salient but unanticipatable fraction of the explanation of what actually happens will emanate from causes located below what I have called the analytic horizon of any of our theories. We may treat these factors as random; for most purposes, they are. Still, the deep meaning of chaos theory is that although order may exist in the apparently random intrusions of factors operating outside our theoretical purview, that order is present at a vast remove from our sensors and is therefore irrelevant to an effort to improve the particulars of a forecast.

With these considerations in mind, we can better understand why the process of forecasting political events of interest in a particular setting entails tracing the contours of the space of possible futures along relevant dimensions, rather than searching for the one actual trajectory that our world will follow through that space. "Point predictions" are fool's gold. For complex social problems, systematic analysis cannot produce them, and attempts to do so encourage a fundamentally misleading construction of the problem. What we can achieve are informed judgments about probabilities, associations, and dynamic nonlinearities within distributions of outcomes.

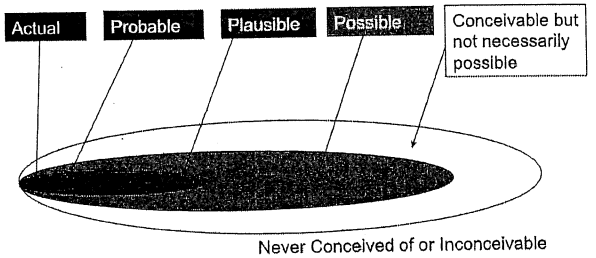

Fig. 7.1. Categories of worlds. The dot indicates the actual future, with all possible outcomes depicted.

Figure 7.1 highlights the irrationality of seeking to predict the actual future, by locating that future as a nearly imperceptible dot in a schematic depiction of intersecting spaces of the conceivable, the possible, the plausible, and the probable.[7] The diagram depicts all possible outcomes as a subset of all those that are conceivable, all plausible outcomes as a subset of those that are possible, all probable (or not improbable) outcomes as a subset of those that are plausible, and the actual outcome as located within the realm of the probable. Except for the dot representing the actual, all other descriptions of the future world in this diagram can be classified as "counterfactuals"—accounts contrary to what the world actually was or what it became.

Although we might prefer to imagine or construct the landscape of the future to be as orderly as is suggested in the depiction in figure 7.1, it is likely to be considerably messier and more complex. Taking a critical stance toward this attractive construction leads us to note, for example, that the ratio of the possible to the conceivable may be substantially greater than we think. Some of the outcomes that we might consider probable may be not only implausible but downright impossible. The "actual" world that emerges may be one of the variants considered not "probable" but only "plausible." Even some of those possibilities we consider relatively probable may be merely possible or actually impossible. For illustrative purposes, figure 7.2 incorporates these potential misconceptions.

Despite the difficulty of producing a reasonably accurate surrogate for the contours of what should be considered "possible," precisely that surrogate is required for any forecasting tool to be consistent with what the "future" actually "is." Whatever technique modelers use to accomplish this task with respect to a domain as complex as international politics, it must be governed by principles or be based on assumptions that enable what is produced to be treated as not only logically but also empirically possible.

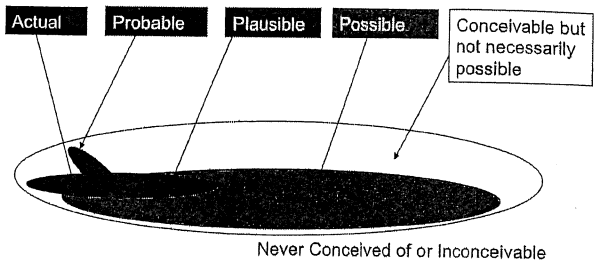

Fig. 7.2. Categories of worlds if imagined incorrectly. The dot indicates the actual future, with possible outcomes depicted with real-world messiness.

The technique must also be capable, at very low marginal cost, of generating large numbers of empirically realizable multidimensional trajectories or chronologies.

ABM and Computer Simulation: Constructing Deconstructible State Spaces of the Future

A large flock of birds traces a trajectory through the sky that is as distinctive as it is unpredictable. One could extravagantly explain this combination of orderliness and unpredictability as the result of a dictatorial leader bird whose whims govern the direction and speed of every other bird in the flock. A better explanation is achieved by using a computer to endow each bird with a simple set of algorithms requiring it to stay near, but not too near, the birds in its immediate vicinity. Numerous computer models of such "boids" show how precisely such an agent-based model can produce flocking behavior, thereby explaining it parsimoniously and in strict conformance with what we know about the information-processing capacities of birds.

Any one run of a boid model, perturbed randomly by the virtual presence of tasty bugs in the line of sight of some boids, traces a flock's trajectory that will be unique among the trajectories the flock could follow. Similarly, each interpretation of the future offered by an agent-based model is a trajectory through an immense space of possible trajectories whose boundaries are established by the interactive implications of the theoretical assumptions of the model, its initial conditions, and the exogenous random stream of tiny perturbations that affect it. Built from and decomposable into identifiable theoretical claims, the model is stylized to conform, at $t = 0$, to a target political system at the present or some stipulated point in the past. By changing

parameters and/or algorithms to reflect adjustments in the theories that we wish to include in the model and by repeating the production of batches of trajectories, we can criticize our own expectations about the future and critically assess the credibility of different theories (once future outcomes can be compared, systematically, to outcome probabilities as registered by model output).

The algorithms that comprise the model's transition rules animate masses of interactions among agents instantiated in ways that condense the limited but high-confidence knowledge available from theoretically and ideographically sophisticated experts and their work. But once the model is animated, the massive interaction effects that arise produce emergent processes of dynamic change across the entire "landscape" of agent behavior, processes that cannot be derived, inferred, or predicted from the algorithms themselves. As individual agents update their state and behavior, the entire array of agents moves forward through time. By collecting data on stochastically perturbed repeated runs of an appropriately assembled model, we can identify outcomes that are typical, plausible, and just possible. Each outcome is consistent with the assumptions, data, and theoretical operationalizations used to build and instantiate the model but is impossible to infer from them. From a constructivist perspective, these algorithms express agentive motivations and have consequences shaped by circumstances, that is, social structure. ABM thus allows a modeling of agent-structure coevolution in ways that both require and exploit contemporary computational power.

Applications of this kind of technology are prevalent in many disciplines, including molecular biology, oncology, archaeology, natural resource management, pharmacology, climatology, immunology, transportation, marketing, and city planning.[8] The approach is now also well established in the social sciences. One of the earliest and still most influential studies animated by this approach was published in 1978 by Thomas C. Schelling, who demonstrated ABM's in-principle fecundity with an agent-based model of segregation.[9] To be sure, Schelling made his key contributions without using a computer and even emphasized the importance of doing necessary calculations manually, but he subsequently developed interest in and spent considerable time investigating how computer programs for exploring his ideas could be designed. Indeed, Schelling-style models run on computers have been used to greatly extend his thinking, by experimenting with different rules that individuals might follow, different patterns of interaction among neighbors, and different tastes for living in integrated or segregated areas.[10]

In political science and other social sciences, computational, bottom-up, or agent-based modeling is a generalization of this method, using computer

simulation to explore the often nonlinear relationship between inputs at the unit level, interaction networks, and outputs at the collective level.[11] ABM has been particularly attractive to researchers in domains where intractability problems make algebraically solvable techniques of formal analysis impossible or when either the complexity of conjectures about macropolitical relationships or the openness of systems involved precludes relying on regression or natural experiments. Among the political science domains where ABM techniques have been successfully deployed are collective mobilization, norm and strategy evolution, constructivist identity theory, secessionism, power sharing, party competition, political communication, national state formation, institutionalization, international treaty making, and the relationship between the structure of the international system and state behavior within it.[12]

Agent-based models are most sensibly deployed to investigate problems that are too complex to be captured algebraically, because of large numbers of relevant dimensions, large numbers of interacting "bodies," or both. Accordingly, there is a strong elective affinity between ABM and computerization, This is because the effects of ABM emerge from the algorithmic behavior and simple interactions of masses of autonomous agents, yielding an otherwise-impossible-to-perform multitude of calculations at each time step—calculations that are straightforward at the agent level but overwhelming if approached as an integrated set. Given the immensity of the possibility space, it is almost certain that building an analytically suitable surrogate for the state space of the future will require computerization. Indeed, the automaticity of computer technology itself helps impose the discipline required for the production of this surrogate, by translating initial conditions and a stable set of theoretical propositions into very large numbers of individually distinctive trajectories.

Each trajectory—that is, each run of a stochastically perturbed ABM—constitutes a distinctive "story," in the sense that every event is fully traceable to the state of the world prior to its occurrence. Since the "laws" governing behavior of types of agents and interactions between types of agents are known, the range of possible outcomes is fully determined (though unknowable). Good computational models, in other words, are capable of combining opportunities for process-tracing, and the distinctiveness and granular, multi-dimensional complexity that traditionally attract constructivist and critical theorists to ideographically oriented research, with the standardization, control, and large *n*'s that enable rigorous testing of nomothetic propositions about a world rightly treated as both orderly and unpredictable.

Mapping the State Space of the Future for Bangladesh

In the late 1990s, I used a rather simple modeling platform known as ABIR (Agent-Based Identity Repertoire) to produce working models of a generic Middle Eastern semiauthoritarian country, to explore the contours of the state space of the future for the Middle East as a whole over a thirty-year period.[13] Substantially refined and elaborated techniques were used in 2002, with a more advanced modeling platform known as PS-I (Political Science–Identity).[14] In response to requests from the US government, this platform was used to produce models and to conduct experiments analyzing prospects for different kinds of stability and instability in Pakistan between 2002 and 2005 and to evaluate repeated cycles of Israeli-Palestinian violence and/or war in the region, considering their implications for stability and the fate of US-friendly regimes in the Middle East.[15] This work was the background for the approach taken by Lockheed Martin's Advanced Technology Laboratories team that participated in a competition sponsored by the Defense Advanced Research Projects Agency (DARPA) and begun in 2005, known as the Integrated Crisis Early Warning System (ICEWS).[16] The competition required modeling teams to reach relatively high levels of accuracy, recall, and precision for forecasts of "events of interest" (e.g., rebellions, insurgencies, domestic political crises, and outbreaks of ethnoreligious violence) in dozens of countries. Forecasts generated automatically from models ingesting data from the late 1990s and early 2000s were tested against what actually happened in those countries in subsequent years.

Most of the modeling effort undertaken by these competing teams used variations of data mining and statistical correlation or pattern-matching techniques. But ABM models produced with PS-I, which passed minimum requirements for forecasting while complementing the ATL team's statistical models, were judged to be especially valuable for the depth and richness of their output and for their potential to assist with option analysis. The result was a separate stand-alone project known as the Virtual Strategic Analysis and Forecasting Tool (V-SAFT), developed under contract from the Department of Defense via the Human Social Cultural Behavioral Modeling Program and the Office of Naval Research. Data presented below are drawn from the array of monthly updated models produced by V-SAFT for a small but growing number of countries.

Bangladesh is one of the nine countries for which monthly updated models were available in 2013. The Bangladesh model's forecast results, which looked forward one year from February 2013, vividly exemplify the

technique of using a theoretically informed, computer-assisted, agent-based model of a real and complex political system to construct an analyzable surrogate for the state space of the future. The presentation of the results here will be followed by a brief account of how the state space that was constructed by running a virtual Bangladesh model was then deconstructed for critically assessing US policy options toward a political crisis threatening the real Bangladesh in late 2013. The published summary forecast of the model looking forward one year in February 2013 reads,

> V-SAFT's Bangladesh model forecasts a somewhat less than 75% probability of a sustained domestic political crisis between March 2013 through February 2014 and only a 10% probability of avoiding any sort of domestic political crisis during this period. Insurgent activity and incidents of rebellion have a roughly 75% probability of appearing in this period, but the likelihood of these activities sustaining themselves for three months or more is very low (below 3%). These probabilities remain more or less the same across the entire forecast period. Nor do probabilities vary widely depending on the complexion of the governing coalition. Under a broad-based nationalist government, a business dominated coalition, or one based on Bengali appeals that is yet not dominated by the Awami League, Bangladesh has the best chance of enjoying a year relatively free from instability or severe instability. But the model shows a wide variety of different coalitions capable of emerging, most of which are associated with some significant instability or severe instability. Probabilities for at least one of these coalitions range from a low of 15% in governments dominated by non-partisan nationalists or business elites, to a high of nearly 65% under governments oriented toward narrow civic and technocratic appeals. Compared to last month's forecasts, a Muslim dominated government is no longer evaluated as plausible, while a business dominated government is to be considered plausible, and not merely possible. V-SAFT's overall assessment for the coming year is that governing elites in Bangladesh will likely face higher than average amounts of illegal and violent political mobilization over the coming year. This represents a substantial trend across the last five months of V-SAFT forecasts for Bangladesh. Keeping in mind what is normal for Bangladesh, this month's forecast registers not just an increasing probability of instability, but an increasing probability of severe instability.[17]

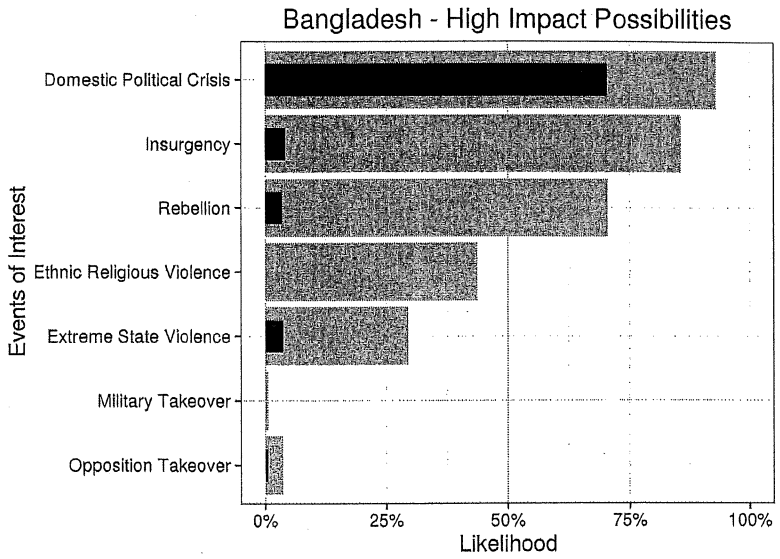

Fig. 7.3. Probabilities for key EOIs of historical importance in Bangladesh, such as the 75 percent chance of a sustained domestic political crisis

Probabilities attached to specific forecasts were calculated directly from proportions of the event space produced by the model exhibiting particular kinds of behavior. Most of the probabilities cited appear in the visuals displayed in this section. For example, the "somewhat less than 75%" likelihood of a sustained domestic political crisis (DPC) over the forecast year and the unlikelihood of Bangladesh avoiding any period of DPC during that time frame are read off the "high impact possibilities" chart (fig. 7.3). The gray "occurs" bar for "domestic political crisis" shows that nearly 90 percent of the weeks comprising all one thousand simulated years of possible Bangladesh futures in the model contain behavior coded as corresponding to the definition of a DPC: organized, mostly nonviolent opposition to the government, significant enough to threaten the integrity of the ruling coalition.[18] The black portion of the bar, extending to roughly 70 percent, shows the probability of a sustained DPC. Precisely, it shows the size of the subset of those weeks featuring DPC and including parts of episodes lasting at least three months. The summary forecast also draws attention to the likelihood of at least some insurgent activity or rebellion. However, in view of the very small size of the black bars in those categories, the forecast notes that sus-

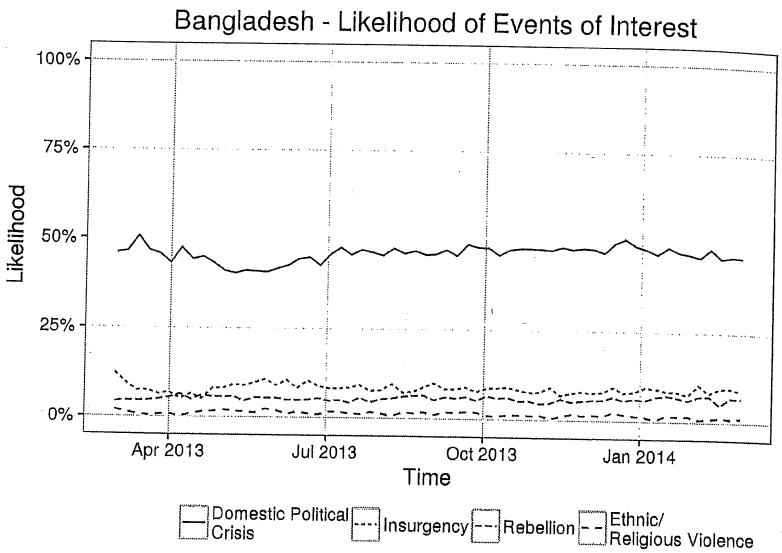

Fig. 7.4. Probabilities for key EOIs in Bangladesh over the forecast year, with DPC relatively stable at 50 percent

tained insurgencies or periods of ethnoreligious violence occurred so rarely in this set of trajectories that they should be considered implausible—in other words, possible but extremely unlikely.

How are probabilities for specific events of interest (EOIs) distributed over the forecast year? Figure 7.4 shows, on the *x*-axis, the months of the coming year and, on the *y*-axis, proportions of futures in the state space, registering EOI activity. As noted in the summary forecast, we see that "these probabilities remain more or less the same across the entire forecast period," with a slight trend toward increasing risk after April 2013.

The summary forecast indicates that instability, including severe instability, is associated with many different constellations of political power that are anticipated as plausible within the forecast year. In figure 7.5, different groups or political appeals that the model identifies as plausibly able to play a dominant and organizing role in Bangladesh governing coalitions over this time period are arrayed on the *x*-axis. The width of the columns associated with political dominance by different groups signifies the proportion of weeks when they are dominant, within all one thousand trajectories; the width of the bars within each column indicates how that amount of the state space is divided into periods (shaded from lightest to darkest) of

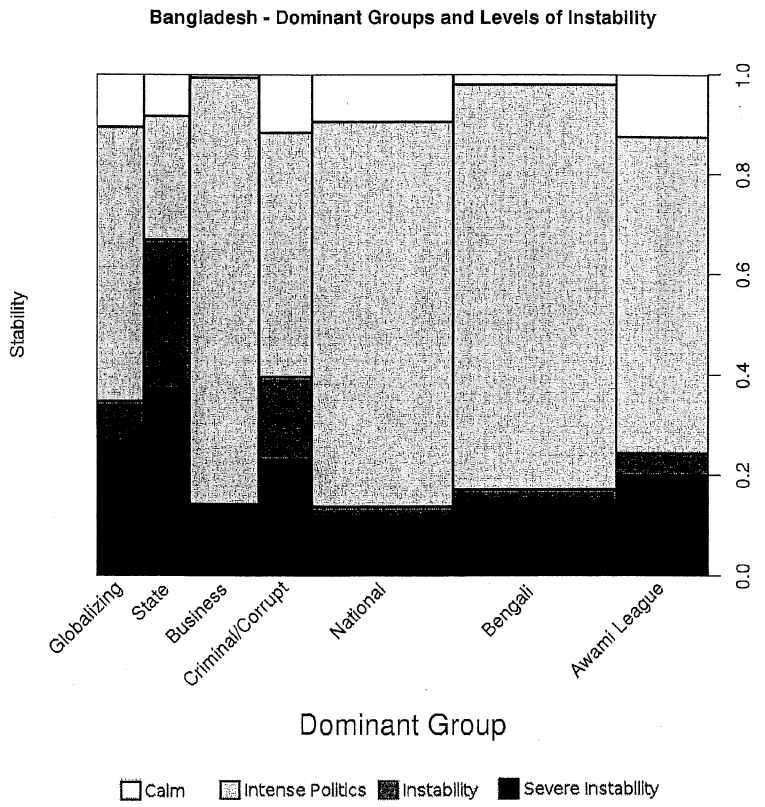

Fig. 7.5. Dominant groups and levels of instability in Bangladesh

calm (white), intense politics (light gray), instability (dark gray), and severe instability (black). Going somewhat beyond the explicit statements in the summary forecast, we can observe that the most likely coalitions—those dominated by nationalist, Bengali, or Awami League ideas and elites—hold out a nearly 80 percent chance of avoiding instability or severe instability, but no government that the model deems plausible has more than a 15 percent likelihood of enjoying calm throughout the year. A government based on alliances led by bureaucratic state elites faces the greatest likelihood of both instability and severe instability.

To probe more deeply, we could see which groups are allies, opponents, or radical opponents of the government when bureaucratic state elites domi-

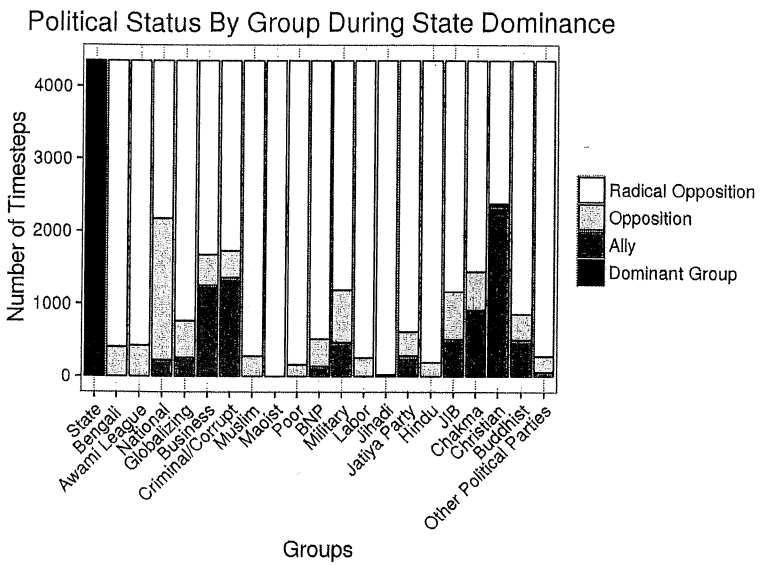

Fig. 7.6. Domestic political hierarchy levels during state dominance

nate the coalition (fig. 7.6) or when "globalizing, Western-oriented" elites dominate (fig. 7.7). We see that these types of governments, when they have significant allies at all, tend to rely on corrupt and criminal elements and on the (highly corrupt) business community.[19]

The *y*-axis of figure 7.6 is the number of model time steps (within the total number) that featured the "state" identity group as dominant. The *y*-axis of figure 7.7 is the number of model time steps (within the total number) that featured the "globalizing" (Western-oriented) identity group as dominant. The shaded bars comprising the columns report the proportion of state-dominant model time (fig. 7.6) and globalizing-dominant model time (fig. 7.7) during which each group was categorized within the DPH as either ally, opponent, or radical opponent.

Again, as noted in the summary forecast, in the relatively unlikely, but still plausible, circumstance that business itself organizes a governing coalition—relying on state bureaucrats, globalizers, and criminal or corrupt networks—prospects for intense politics, if not calm, appear to improve considerably. We can see why this occurs by considering the display in figure 7.8, which shows that when business dominates, the largest groups in Bangladesh are much less likely to adopt a radical oppositionist posture toward the government (the white portion of columns on the left of the display) than when state bureaucrats, globalizers, or criminal or corrupt networks (not shown) dominate.

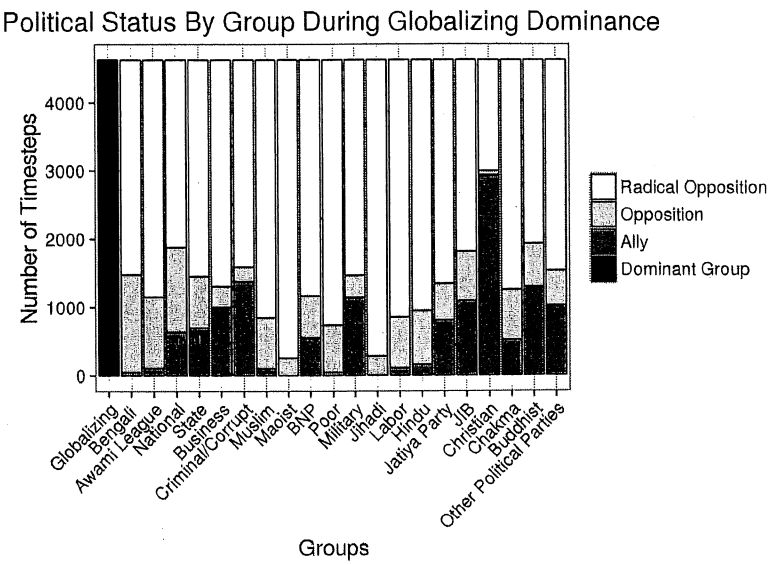

Fig. 7.7. Domestic political hierarchy levels during globalizing dominance

Comparing the state space mapped by running V-SAFT's Bangladesh model for the year beginning March 2013 to mappings produced in earlier months can help identify general trends. For example, in Bangladesh, Islamist-dominated governments appeared prominently enough in the state space to justify treating that outcome as plausible in each of the six previous months' forecasts (meaning that it appears in at least 3 percent of the space of the future). For the year beginning in March 2013, however, the likelihood of Islamist governments dropped below the 3 percent threshold. Accordingly, although such regimes appear in fig. 7.9 in the display on the left, for the year beginning February 2013, they do not appear in the display on the right—the year beginning March 2013. By noting the thin black line at the bottom of the column marked "Muslim" in the "political status by group" display in figure 7.10, we can see that V-SAFT still charted a Muslim-dominated government as possible, just not as "plausible" by the

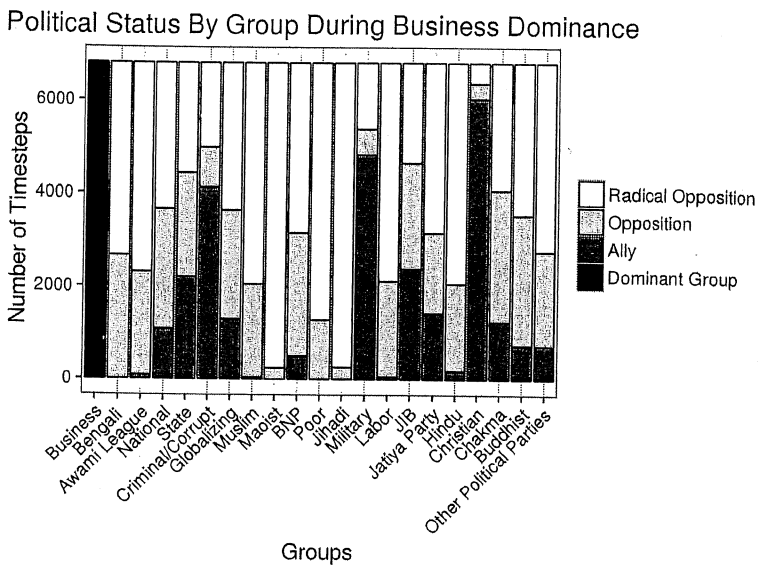

Fig. 7.8. Domestic political hierarchy levels during business dominance

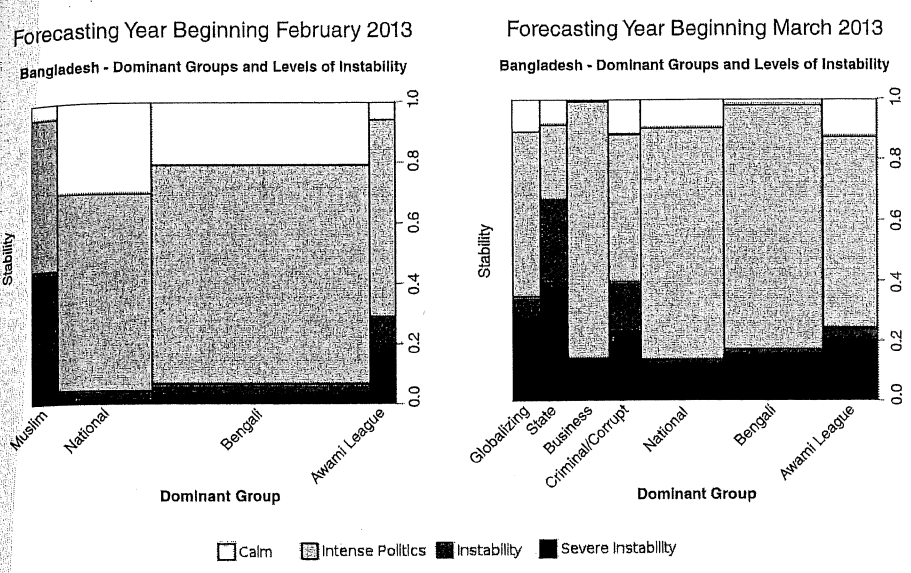

Fig. 7.9. Dominant groups in Bangladesh

3 percent rule. Along the *x*-axis in figure 7.9 are the dominant groups, the first *y*-axis is stability, and the second *y*-axis is the percentage of the space in which the forecast holds.

The final portion of the summary forecast focuses on the trend toward an increasing likelihood of instability, particularly severe instability. This trend is most easily discerned by considering the group of six sequence plots displayed in figure 7.11—beginning with the forecasts made for the October 2012–September 2013 year and ending with forecasts made for the March 2012–February 2014 year. V-SAFT produces these sequence plots by stacking all one thousand trajectories. Each week in each trajectory, traced horizontally by a single line, is shaded to represent either calm (white), intense politics (light gray), instability (dark gray), or severe instability (black). Determining the order of the stacking is the condition present at the beginning of the run and, within that group, the length of time before that condition changes. Different horizontal patterns register different sequences of stabilization or destabilization. By noting the changing prominence of shading horizontally, we can gain a quick sense of time trends toward stabilization (becoming more white or light gray) or toward destabilization (becoming more dark gray or black). By noting the changing prominence of shades

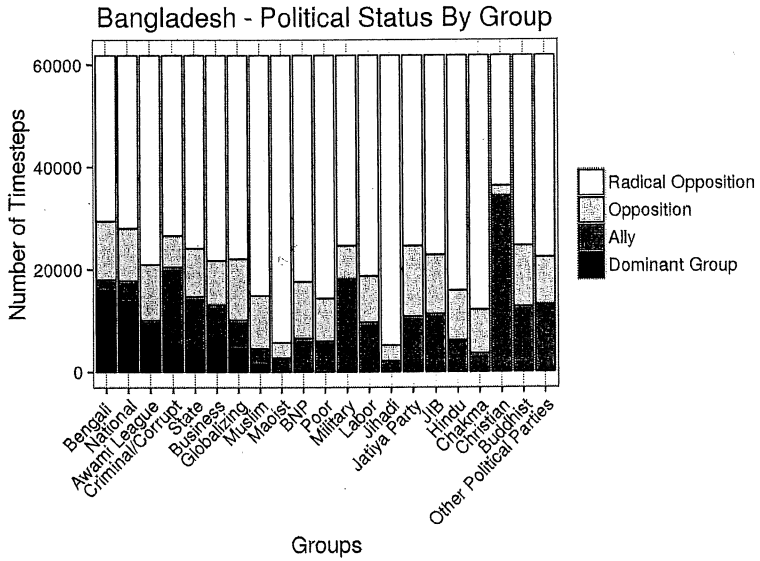

Fig. 7.10. Political status by group in Bangladesh

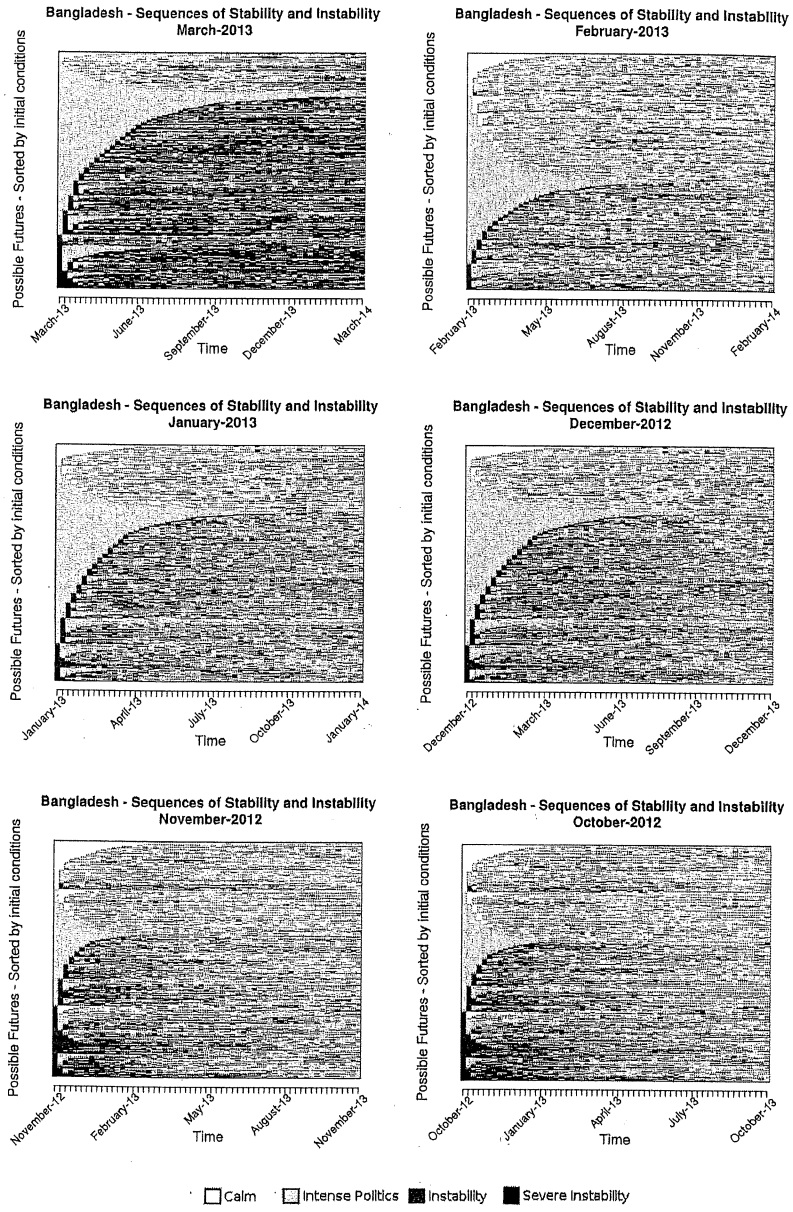

Fig. 7.11. Six sequence plots with forecasts for October 2012 until February 2014. Each of one thousand trajectories is traced horizontally by a single line to represent either calm (black), intense politics (dark gray), instability (light gray), or severe instability (white).

vertically, we can gain a sense of how much of the state space of the future—taking into account all one thousand trajectories—is relatively stable or unstable across the entire square. In the case of Bangladesh forecasts for the year beginning with October 2012 through the year starting in March 2013, we see a trend toward a pronounced shrinkage in the areas of light gray and white and expansion in the areas of dark gray and especially black, suggesting not just a decreasing probability of calm but an increasing probability of either intense political competition or severe instability.

Precisely how productive is this approach to thinking about the future compared to others? How productive might it become? These questions are unanswerable in the context of this chapter, but some evidence for the validity of the results of V-SAFT's depiction of the space of Bangladesh's future can be considered. If we set aside "internal validity" or "verification" questions, considering that—in principle, at least—computer models do offer certain assurances about consistency of propositions and transparency of operations, we can consider the more challenging question of "external validity." How good was the modeling effort as a guide to the future of Bangladesh? This is a daunting question, since once the future is understood as a distribution of possible trajectories, with the actual future understood as but one of them, it is difficult to learn about the model's validity from the failure to forecast correctly in a particular case. But models can be faulted and perhaps improved, if they fail to include outcomes as possible when events show that they actually were. A more systematic and demanding form of validation is to require very large numbers of forecasts from such models—forecasts made with various levels of confidence or with various probabilities attached. The distribution of these forecasts can then be compared to the distribution of actual outcomes for evidence that the contours of the spaces of the future produced by the model resembled the contours of the actual spaces of the future from different points in the past.

Using a technique known as a "separation plot" (developed by Greenbill, Ward, and Sacks),[20] V-SAFT accumulated an archive of forecasts of four EOIs by country and compared them to "ground truth" with respect to whether and when those events occurred. Figure 7.12 displays a separation plot reporting data about EOIs in Bangladesh from October 2012 to November 2013 and V-SAFT forecasts of those EOIs. For each event-month combination listed across the *x*-axis, the shade of each bar represents whether the event occurred and how intense it was according to a review of relevant online sources. The black line represents our forecasted likelihood of each event one month prior (ranging from 0 percent to 100 percent likely). The most accurate forecast would only have shaded bars to the far right, and

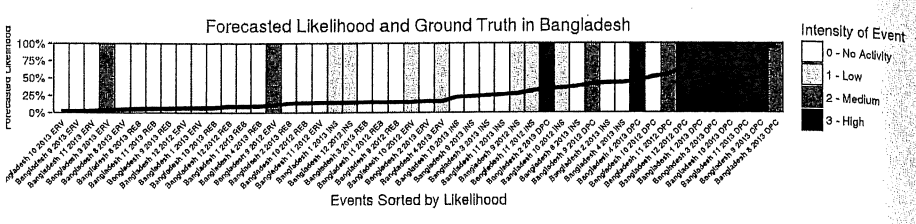

Fig. 7.12. Forecasts arrayed by increasing forecasted probabilities of events plotted against their nonoccurrence, occurrence, and intensity

all of the bars to the far left would be white. Simple visual inspection suggests an encouraging interpretation, since, despite errors, the white columns, where events did not occur, are strongly clustered to the left, where low probabilities were forecasted, while dark columns, where events did occur, are strongly clustered to the right, where high probabilities were forecasted.

V-SAFT models, including virtual Bangladesh, have been used repeatedly to conduct "what-if" experiments, counterfactual experiments that seriously treat questions that scholars have often discussed as whimsy. What, it has often been asked, would have happened if Hitler had been killed in the trenches in World War I, if the asteroid that made the Chicxulub crater had narrowly missed the Earth rather than plunging into the Yucatan, if the battle of Salamis had been fought without the brilliance of Themistocles, and so on? What, as Steven J. Gould put it, would have been the case if the "tape of life" could be "replayed" under slightly different conditions?[21] With the sort of capacity described here, this kind of question is no longer whimsy, and answers provided to it can be evaluated, at least in principle, against criteria much more reliable than rhetorical seductiveness.

In the fall of 2013, Bangladesh was headed toward elections. The opposition party, the Bangladesh Nationalist Party (BNP), fully cognizant of the way it had manipulated past election outcomes when it was in power, demanded that the Awami League dominated government turn over the management of the elections to a neutral caretaker government—one run by the judiciary and the state bureaucracy. A law had been passed requiring this arrangement, but the government argued that progress made to rid Bangladesh of corruption was sufficient to justify the government's direct management of the election, without a caretaker regime. The BNP began a campaign to delegitimize the elections, which threatened to destabilize the country, potentially involving intervention of the military. The United

States, wishing to preserve Bangladeshi democracy while insuring the country's stability, was faced with the problem of deciding whether to push the government to acquiesce in the demands of the Islamist-leaning BNP and the apparent requirements of the law or to maximize prospects for the stability of the country under the rule of leaders Washington favored. By closely analyzing the distribution of futures produced by the virtual Bangladesh model, it was possible to discern a pattern that yielded nonintuitive but compelling insights into the implications of this decision.

We proceeded by identifying a subset of futures produced by the model, in which the election period featured either a caretaker government or a government dominated by the Awami League. Comparison of patterns of instability within those two types of futures (each comprising approximately 17 percent of the entire distribution) suggested that a caretaker government would increase the stability of the election period itself, but at the price of a considerably larger risk of severe instability later. On the other hand, Awami League government management of the election would produce more disturbance during the election period but lower risks of subsequent severe instability. The model also suggested that the risks of subsequent severe instability could be significantly reduced if the composition of the Awami League's government was broadened to include representatives of major sectors and groups not directly associated with the BNP.

Military domination of the country did not emerge within the range of "possible" outcomes that our model forecast, but since this counterfactual was much discussed and much worried about, we conducted "what-if" experiments by rerunning the model to produce two additional batches of trajectories. The script governing the running of each experiment punctuated the trajectories, at a point prior to the scheduled election period, with two types of military interventions, each with a precedent in Bangladesh's political history. The first intervention, high-level coup, featured seizure of the government in the capital by high-ranking officers. The other focused on action by low-level officers on a much more diffuse basis, taking over local governing institutions.

The result can be summarized by noting that a military coup was seen to significantly decrease the likelihood of an apolitical caretaker government afterward. A period of military governance following a coup does become possible and, in the case of a high-level coup, even politically stable. However, following either form of military intervention, political polarization increased, signaled by a marked change in the balance of violent versus nonviolent mobilization.

Conclusion: Why Bother with Another Method?

The principles governing the contributions to this volume are that the worlds we experience and analyze are constructed and that understanding those worlds means critically interpreting them by deconstructing the assumptions and frames of reference they reflect. I have argued that ABM offers a theoretically powerful approach to implementing these principles systematically, supporting my argument by offering the "hard" case" of thinking in a disciplined way about the future. One might even consider this challenge an extreme example of the problem of analysis with a small n and many variables (since, technically, $n = 0$). With computer power as cheap as it is and with islands of good theory ready to be applied to real problems, the approach has the potential, at least, to do much better than any other method available in producing outputs closer to the complexity of the worlds inhabited by politicians and policy makers. For younger scholars, accustomed to finding and using apps of various kinds, the intuition that computers could be used for social science research in exciting new ways is probably more natural than for previous cohorts.

To be sure, portions of the learning curve for using the approach can be steep. But acquisition of the necessary skills is less arduous than may be commonly imagined. For one thing, neither advanced algebra nor calculus is required. Indeed, ABM is designed primarily for problems that cannot, even in theory, be solved mathematically. Twenty years ago, working with ABM would require good programming skills, but modeling platforms are now available that do not assume the ability to read or write in a programming language and that enable substantive theoretical knowledge to be quickly and intuitively encoded.

ABM based in computer simulation is dynamic, visually entrancing, rich in data, and formal. In all these ways, it is nifty. But niftiness per se does not itself justify attention, especially for scholars who are masters of other methods still useful for problems yet unsolved. There are deeper reasons why IR scholars should investigate the technique, other than the increased facility they may thereby enjoy for answering questions they already entertain. Specifically, mastery of this method is likely to significantly expand both the number and kind of questions IR scholars are able to address.

Without a systematic method to produce very large numbers of analytically consistent but substantively distinctive scenarios, both researchers and policy-makers are precluded from thinking about what has not happened yet as distributions of what could happen. Instead their attention is confined to tiny zones in the state space of the possible—zones traced by a few strategically or idiosyncratically crafted stories. As emphasized at the beginning of this chapter, thinking in a disciplined way about the future requires the systematic and controlled construction of very large numbers of scenarios. This requires a method capable of exploiting what is now readily available computing power to conduct large numbers of disciplined thought experiments, thereby offering opportunities to ask more appropriate questions about the future than can be asked with conventional techniques. Methods, in other words, should not be seen as only equipment used by those whose ideas and purposes arise independent of method. Methods, such as ABM, can often prime the pump of insight. Indeed, precisely because this method is based squarely on principles of constructivism and critical practices of deconstruction and because it incorporates randomness, it will act as an intuition pump—changing the way researchers think about the social world, expanding the questions they see it proper and possible to address, suggesting approaches to problems that have seemed intractable, and simultaneously imposing discipline on interpretations of what the world was, is, can be, and why.

But just as a new interpretation of the world may shock or confound those whose expectations it contradicts, a new method may produce incredulity and even horror if it seems to entail combining "high-tech" devices with an "interpretivist" posture that usually celebrates itself as antagonistic to misplaced "scientism." Anyway, why bother with a method that requires investment and training, if one does not feel currently available methods have been exhausted of the contributions they could make? Is the only reason to adopt a new method that it provides better ways to achieve established goals or answer standard questions ? For most researchers in most situations, the answer may well be "yes." Afterall, the very meaning of "normal science," in Thomas Kuhn's terms, is that paradigmatic techniques and methods for solving problems are fully agreed on and seen as effective. However, when researchers are unsatisfied with the questions and answers available in their field, then methodological questions can become nearly indistinguishable from substantive ones. In this circumstance progress is likely to require new methods enabling researchers to pose questions impossible to address with familiar techniques.

Herein lies a serious obstacle to scientific progress, or scholarly work in general. Established researchers have outsized influence on what apprentice researchers judge to be the appropriate methods and training to receive. Successful researchers are successful partly because they have mastered technologies to support investigation of the particular kinds of questions the methods associated with those technologies can answer effectively. The role of senior

researchers as models, supervisors, or judges of new work thus considerably raises the barriers to entry that methods designed for answering different questions must overcome, even if the answers to those questions might be more rewarding—intellectually or socially. Therefore, despite the crucial role that loyalty to established research programs, paradigms, and methods plays in the ability of communities of experts to share and rigorously critique their work, scientific progress cannot rely only on the questions and methods that scholars who are currently successful find "interesting" or useful. Ultimately, disciplines, as research programs, are threatened and destroyed by the degenerative consequences of narrowing substantive foci and the exclusion of new questions deemed unaddressable—and therefore uninteresting.

Whether via the contingencies of Kuhnian paradigm shifts or mechanisms of Lakatosian competition for resources among rival research programs, science does progress. Substantive innovation in theory and in methods available can constitute the proximate impetus for these shifts—most likely a combination of the two. The crucial point is that from the point of view of the advancement of science in general and of the social sciences in particular, it is a fundamental mistake to view methodology as purely a function of the substantive task at hand or of interest only to methodology "specialists." Across all the arts and sciences, methods that are employed both open and close avenues of investigation, creativity, and discovery. Without new methods of seeing or studying the world, many kinds of new questions could never be posed. Consider how artists expanded their horizons as new methods and materials for drawing and painting became available. Or think of poetry, which, as a method, imposes disciplines on the use of language, such as meter and rhyme. These elements of the poetic method restrict language use, thereby forcing a greater role for imagination and creativity. In the hands of excellent practitioners, the result is new ideas, associations, images, and expressions. Without the rules and limits imposed by the method of poetry, these uses of language would otherwise remain undiscovered and unexpressed. In other words, as a method, poetry allows access to terrains of the human imagination that would otherwise remain unexplored.

In the sciences, there could be no Newtonian theory of gravity without calculus, no theory of general relativity without tensor calculus, no microbiology without the microscope, no modern astronomy without the telescope, no experimental particle physics without particle accelerators, and no sophisticated paleontology or serious cross-cultural archaeology without carbon dating. In the social sciences as well, methods help produce the ideas we have and the questions we pose. In the last fifty years, a number of new methods have transformed what we know of as the social sciences and the

practices associated with them. These include Bayesian statistics, natural language processing, mass survey techniques, field experiments, discourse analysis, deconstructionist techniques of textual exegesis, game-theoretic formal modeling, and even word processing. In all fields, new methods that arise in response to unsolved problems and new technological capacities help produce new questions, new ideas, new answers to old questions, and often new sets of questions that themselves become the basis for new subfields and even new disciplines.

Constructivist and critical stances toward domains of interest in the social sciences open the potential for new insights and effective understandings. But only methods capable of exploiting that potential—for systematic elaboration of possibilities and disciplined deconstruction of outcomes—can enable scholars to exploit that potential. I have argued that ABM is such a method, and I have illustrated the plausibility of this claim with a report on the application of one relatively ambitious ABM tool for parsing the space of political futures. V-SAFT's successes with respect to Bangladesh and a range of other countries suggest that deployment of virtualization models of particular political systems can be more effective than standard methods for exploring and evaluating the relationship between variables and past outcomes as well as between policy choices and future outcomes. However, whether V-SAFT is presently developed enough to regularly outperform other techniques should not be the sole or even the primary criterion for giving the kind of investigative tool it represents serious consideration. The real test of a method is not whether it outperforms other methods on traditional tasks but whether it can lead to a level of sophistication and breadth of consideration that lead scholars and scientists to trade traditional tasks for even better ones.

Notes

1. See Gary King and Langche Zeng, "Logistic Regression in Rare Events Data," *Political Analysis* 9 (2001): 137–63.

2. I thank Matthew Reichert for his very helpful and repeated readings of earlier drafts of this essay and Miguel Garces for his critical reading of the text and his expert assistance in the development of data visualizations. Portions of this chapter appeared in "Making Sense of Social Radar: V-SAFT as an Intelligent Machine," in *Sociocultural Behavior Sensemaking: State of the Art in Understanding the Operational Environment*, ed. Jill Egeth, Gary D. Klein, and Dylan Schmorrow (McLean, VA: MITRE, 2014), 317–37.

3. For an extended treatment of the problem of disciplining counterfactual analy-

sis, see Ian S. Lustick, "Tetlock and Counterfactuals: Saving Methodological Ambition from Empirical Findings," *Critical Review* 22, no. 4 (2010): 427–47.

4. Benoit Mandelbrot and Richard L. Hudson, *The (Mis)Behavior of Markets: A Fractal View of Financial Turbulence* (New York: Basic Books, 2004), 77.

5. Nassim Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2010), xxix.

6. Ibid., xxii.

7. This entire approach is related to "modal logic" and "possible worlds" theory in philosophy, as developed to think about a variety of philosophical problems associated with thinking counterfactually. For an extreme but influential view, see David K. Lewis, *On the Plurality of Worlds* (Malden, MA: Blackwell, 1986). For colorized versions of the figures appearing in this chapter, see Lustick, "Making Sense of Social Radar," 317–37, https://www.mitre.org/sites/default/files/publications/SocioculturalSensemaking.pdf.

8. For details and exemplary studies in each of these fields, see http://www.martinos.org/biosystems/projects.php; S. J. Dean, G. J. Gumerman, J. M. Epstein, R. Axtell, A. C. Swedlund, M. T. Parker, and S. Mccarroll, "Understanding Anasazi Culture Change through Agent-Based Modeling," in *Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes*, ed. T. A. Kohler and G. J. Gumerman, Santa Fe Institute Studies in the Sciences of Complexity (New York: Oxford University Press, 2000), 179–205; L. R. Izquierdo, N. M. Gots, and J. G. Polhill, "FEARLUS-W: An Agent-Based Model of River Basin Land Use and Water Management" (paper presented at the conference "Framing Land Use Dynamics," Utrecht, April 2003),; Riccardo Boero and Flaminio Squazzoni, "Does Empirical Embeddedness Matter? Methodological Issues on Agent-Based Models for Analytical Social Science," *Journal of Artificial Societies and Social Simulation* 8, no. 4 (2005): 6, http://jasss.soc.surrey.ac.uk/8/4/6.html sections 4.42–4.47; http://www.talkorigins.org/faqs/genalg/genalg.html; http://biosystems.ucsf.edu/research.html; http://www.msysbio.com/; Joshua M. Epstein, *Agent Zero: Toward Neurocognitive Foundations for Generative Social Science* (Princeton, NJ: Princeton University Press, 2014).

9. Thomas C. Schelling, *Micromotives and Macrobehavior* (New York: Norton, 1978).

10. See Rainer Hegselman, "Thomas C. Schelling and the Computer: Some Notes on Schelling's Essay 'On Letting a Computer Help with the Work,'" *Journal of Artificial Societies and Social Simulation* 15, no. 4 (2012): 9, http://jasss.soc.surrey.ac.uk/15/4/9.html.

11. The most convenient and reliable source for surveying the variety of applications of ABM to social science problems is the *Journal of Artificial Societies and Social Simulation* (http://jasss.soc.surrey.ac.uk/).

12. For a survey of political science applications of ABM, see Ian S. Lustick and Dan Miodownik, "Abstractions, Ensembles, and Virtualizations: Simplicity and Complexity in Agent-Based Modeling," *Comparative Politics* 41, no. 2 (2009): 223–44.

13. See Ian S. Lustick, "Locating the History We Have within the Distribution of Counterfactuals: Laws, Stories, and Computer Simulation" Paper prepared for presentation at Columbia's Center for Historical Social Science, for the workshop on "History, Rationality, and Agency in Social Science," New York, April 16, 2002), http://www.polisci.upenn.edu/ps-i/papers.html.

14. See Ian S. Lustick, "PS-I: A User-Friendly Agent-Based Modeling Platform for Testing Theories of Political Identity and Political Stability," *Journal of Artificial Societies and Social Simulation* 5, no. 3 (2002); Lustick and Miodownik, "Abstractions, Ensembles, and Virtualizations"; Ian S. Lustick, Dan Miodownik, and Roy J. Eidelson, "Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?," *American Political Science Review* 98, no. 2 (2004): 209–29.

15. See Ian S. Lustick, "Simulating the Effects of Israeli-Palestinian Violence, Fundamentalist Mobilization, and Regional Disruption on Regime Stability and USA-Friendly Outcomes in Middle East Polity" (paper presented at the workshop "Origins and Patterns of Political Violence 1: Violence in Civil Wars," Santa Fe, January 16–18, 2004), http://www.polisci.upenn.edu/ps-i/publications/Lusticksantafe.pdf.

16. See Sean O'Brien, "Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research," *International Studies Review* 12, no. 1 (2010): 87–104.

17. Slight disrepancies between the original and this discussion are due to the correction of small errors discovered in the process of reformatting the visuals for this publication.

18. These higher-level concepts require precise operationalizations. For example, DPC is operationally defined as existing "in a time-step when the aggregate influence of protesting agents (protest) multiplied by the protest_threshold (6) is greater than the aggregate influence of the dominant identity (data$dominant_activation)." The routine that implements this category as a rule for data collection is dpc_exists = (protest*protest_threshold>data$dominant_activation).

19. In figures 7.6–7.8 and 7.10, groups and identities listed along the *x*-axis are arrayed from left to right in order of decreasing "aggregate influence" during the periods of dominance by the group named under the column on the extreme left. V-SAFT computes aggregate influence for each group for any time step by summing the individual influence of agents activated on that identity.

20. See Brian Greenhill, Michael D. Ward, and Audrey Sacks, "The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models," *American Journal of Political Science* 55, no. 4 (2011): 991–1002.

21. Concerning the prominence of this meme, implying (incorrectly) that the tape of "history" cannot be rerun, see Lustick, "Tetlock and Counterfactuals," 441–42.