International Studies Quarterly (2022) **66**, sqac036

One without the Other? Prediction and Policy in International Studies

Kristian Skrede Gleditsch
University of Essex, UK and
Peace Research Institute Oslo, Norway

Salient events such as the COVID-19 pandemic and the Ukrainian crisis and Russian invasion lead to interest and debate on how research can inform policy responses. Research can provide important evidence on the potential consequences of different actions with regard to specific objectives. However, research conclusions also remain uncertain and typically depend on many assumptions. I argue that since policy prescriptions entail claims about future consequences, they must be seen as predictions. Although prediction is difficult, especially about the future, we can have better and more informed discussions about policy consequences if we pay attention to what we have learned about predictions and predicting better. Moreover, beyond whether predictions are ultimately correct or not, it is useful to consider what we can learn from them. In some cases, it is more helpful to understand how specific inputs influence predictions than to focus only on approaches that maximize overall fit. Predictions may not be intended to influence policy, but research is more likely to be useful to non-academics if presented in a clear and accessible manner. Finally, since the outcomes we try to predict depend on policy choices, we can potentially improve predictions by thinking about how decisions are made.

Acontecimientos destacados, como la pandemia de la COVID 19, así como la crisis en Ucrania y la invasión rusa, suscitan el interés y el debate sobre el modo en que la investigación puede suministrar información para las respuestas en materia de políticas. La investigación puede proporcionar pruebas importantes sobre las posibles consecuencias de diferentes acciones con respecto a objetivos específicos, pero, sin embargo, las conclusiones de la investigación siguen siendo inciertas y suelen depender de muchos supuestos. Sostenemos que, dado que las recomendaciones en materia de política implican aseveraciones sobre las consecuencias futuras, deben considerarse como predicciones. Aunque hacer predicciones sea difícil, especialmente sobre el futuro, podemos llevar a cabo debates, mejores y más informados, sobre las consecuencias en materia de políticas si prestamos atención a lo que hemos aprendido sobre las predicciones y a hacer mejores predicciones. Además, más allá de si las predicciones resultan ser correctas o no, es útil considerar lo que podemos aprender de ellas. En algunos casos es más útil entender como contribuciones concretas influyen en las predicciones que centrarse solo en enfoques que maximicen el grado de adecuación general. Puede que las predicciones no se destinen a influir en las políticas, pero es más probable que la investigación sea más útil para no académicos si se presenta de una manera clara y accesible. Por último, puesto que los resultados que intentamos predecir dependen de las decisiones en materia de políticas, tenemos el potencial para mejorar las predicciones si pensamos en cómo se toman las decisiones.

Les événements d'importance majeure tels que la pandémie de COVID-19 ou l'invasion de l'Ukraine par la Russie incitent à se pencher sur la manière dont la recherche peut informer les réactions politiques. En effet, la recherche peut fournir des preuves cruciales sur les conséquences potentielles de diverses actions, au regard d'objectifs spécifiques. Toutefois, ses conclusions demeurent également incertaines, et s'appuient généralement sur de nombreuses hypothèses. J'avance que, dans la mesure où les mesures politiques impliquent des affirmations relatives aux conséquences futures, elles doivent être considérées comme des prédictions. Or, bien qu'il soit difficile de prévoir l'avenir, il devient possible d'avoir des discussions plus efficaces et mieux informées en matière de conséquences politiques dès lors que nous prêtons attention à ce que nous savons des prédictions, et des moyens de les améliorer. Par ailleurs, au-delà du caractère avéré ou non des prédictions, il est important de réfléchir à ce que nous pouvons apprendre d'elles. Dans certains cas, il est plus utile de comprendre comment des données spécifiques influencent des prédictions plutôt que de se focaliser sur des approches consistant à optimiser ces dernières. Si les prédictions ne doivent pas être destinées à influencer la politique, la recherche a davantage de chances d'être utile à un public non universitaire si elle est présentée d'une manière à la fois claire et accessible. Enfin, dans la mesure où les conséquences que nous essayons de prévoir dépendent de choix politiques, nous pouvons potentiellement améliorer les prédictions en réfléchissant à la manière dont les décisions sont prises.

Introduction

Like many others, I have spent much time since March 2020 reviewing and reflecting on research on COVID-19 and the policy responses to the pandemic. This has in turn inspired me to reflect on research and policy in my own field. The crisis over Russian demands on Ukraine in early 2022 and the subsequent Russian invasion of Ukraine is at the time of

Kristian Skrede Gleditsch is a Regius Professor of Political Science at University of Essex and a research associate of Peace Research Institute Oslo. His research interests include conflict and cooperation, democratization, and prediction. He was the president of the International Studies Association 2021–2011.

Author's note: This article is based on my presidential address to the 2022 annual meeting of the International Studies Association in Nashville, TN, USA, on March 31. I am grateful for helpful comments from and discussions with Barış Arı, Kyle Beardsley, Mia Bloom, Ursula Daxecker, Han Dorussen, Erik Gartzke, Jack Goldstone, Faten Ghosn, Theodora-Ismene Gizelis, Nils Petter Gleditsch, Håvard Hegre, Cullen Hendrix, Nils Lid Hjort, Carl Henrik Knutsen, Steve Pickering, Andrea Ruggeri, Håvard Strand, and Jonas Vestby as well as other participants at presentations for the May 4 2022 University of Essex Regius Lecture in Political Science and a workshop organized by the "Stability and Change" program at the Centre for Advanced Studies, Norwegian Academy of Science and Letters in Oslo, on May 11–12. I would like to dedicate this article to my late supervisor, close collaborator, and personal friend Michael D. Ward. We had many useful discussions on forecasting, and his work on conflict prediction has had a major influence on the field. I hope he would have enjoyed this presidential address.

Gleditsch, Kristian Skrede (2022) One without the Other? Prediction and Policy in International Studies. *International Studies Quarterly*, https://doi.org/10.1093/isq/sqac036

writing stimulating a great deal of interest in what research and evidence can tell us about possible responses and their consequences.

My main concern here is on a specific aspect of the broader research and policy link, namely the relationship between prediction and policy. At first, some might find this to be an odd combination of topics, since prediction and policy are usually seen as unrelated. If one is generous, one might perhaps concede that both topics tend to generate controversy, although they do so for very different reasons.

The idea that social science should aim to predict is at best controversial. Philosophers of science such as Carl Hempel (1966) have told us that explanation and prediction are two sides of the same coin. However, many social scientists reject out of hand that we should evaluate theories based on their ability to generate successful predictions.[1] It is often argued that since prediction about the future is so difficult, we should not waste time on it. I do not agree, but more on this later.

In contrast, many people seem unable to get enough of policy or policy implications of research. We may have controversy over specific policies or prescriptions. However, the idea that research should have policy relevance is widely embraced, and funders and policymakers often call for research to be "more relevant" to policy.[2]

Prediction and policy are clearly not inherently related. Scholars very often do one without worrying at all about the other. My argument is that policy and prediction *ought* to be closely related—like love and marriage. Perhaps we can do one without the other,[3] but I will try to convince you that doing one well inevitably requires us to pay attention to the other at some point.

A presidential address should not be a mystery novel, so let me anticipate my three key points:

1) Since policy prescriptions entail predictions about future consequences, policy analysis cannot avoid prediction. However, we can have better and more informed discussions about policy consequences if we pay attention to what we have learned about predictions and predicting better.
2) Moreover, beyond the question of whether predictions are ultimately correct or not, it is useful to think about what we can learn from them. In some cases, it will be more helpful to understand how specific inputs influence the derived predictions rather than whether some modeling approach maximizes overall fit over an alternative. Predictions may not be intended to influence policy, but research is more likely to be useful to non-academics if presented in a clear and accessible manner.
3) Finally, since the outcomes we try to predict depend on policy choices, we can potentially improve predictions by thinking about how decisions are made.

Prediction, Research, and the Future

Social scientists in general, and international studies scholars in particular, tend to have a rather ambiguous relationship to prediction and forecasting. On the one hand, researchers are often forward-looking, and many are eager to argue that their work is "useful for understanding the future." However, on the other hand, there is often also a deep reluctance to try to evaluate theory on their predictions, using formal scoring, and some see a focus on prediction as clashing with a more important goal of "explanation."

In high school, I was often told that studying history was important since learning about the past would help not repeat past mistakes and learn lessons for the future.[4] Similar arguments are often made on the value of theory for international studies.[5] Most students are drawn to our field because they want to understand what is likely to happen in the future and what may influence this. For example, what will world politics look like with a more powerful China?[6] How will COVID-19 change our world?[7] What do theories of civil war tell us about the risk of political violence in the United States?[8]

Many researchers offer informal predictions about future events, more or less explicitly based on some theoretical framework. A famous example in our discipline is Huntington's (1993) prediction that we would have a "clash of civilizations" after the Cold War, once the superpower rivalry had abated. In many cases, the precise meaning of more informal predictions remains unclear. This is one important reason why one can still find debates about whether Huntington was right or not, almost twenty years after the original publication.[9] For the record, I personally think Huntington's prediction was wrong. I have written elsewhere (with Erik Gartzke) about how we see much more conflict within than between civilizations, even when we take into account other factors affecting the risk of conflict such as distance (see Gartzke and Gleditsch 2006).[10]

[1] See, for example, the 2012 New York Times op'ed by Jacqueline Stevens, asserting that "political scientists are lousy forecasters" (https://www.nytimes.com/2012/06/24/opinion/sunday/political-scientists-are-lousy-forecasters.html, accessed May 15, 2022) or the Hechter et al. (1995) symposium on prediction in the social sciences.

[2] See, for example, the discussion in Walt (2005). For a dissenting view, see Kalyvas and Strauss (2020).

[3] In his 1955 song "love and marriage", Frank Sinatra claims that "you can't have one without the other."

[4] The quote by George Santayana (1905) that "[t]hose who fail to learn from history are condemned to repeat it" has been popularized through the claim that Winston Churchill invoked it in a speech to the House of Commons in 1948 (see, e.g., https://winstonchurchill.org/resources/in-the-media/churchill-in-the-news/folger-library-churchills-shakespeare/, Accessed May 15, 2022). However, there appears to be no evidence on record that Churchill actually said this (https://api.parliament.uk/historic-hansard/people/mr-winston-churchill/1948, Accessed May 15, 2022).

[5] See, for example, Mearsheimer and Walt (2013, 437).

[6] See, for example, Allison (2017) and https://impakter.com/can-chinas-rise-to-power-peacefully-here-is-what-political-theories-tell-us/, Accessed May 15, 2022.

[7] https://theconversation.com/i-spoke-to-99-big-thinkers-about-what-our-world-after-coronavirus-might-look-like-this-is-what-i-learned-146986, Accessed May 15, 2022.

[8] See, for example, https://www.newyorker.com/news/daily-comment/is-a-civil-war-ahead, Accessed May 15, 2022, and Walter (2022).

[9] See, e.g., the symposium at https://www.e-ir.info/publication/the-clash-of-civilizations-25-years-on-a-multidisciplinary-appraisal/, Accessed May 15, 2022. The lack of clarity about the specific implications of the argument is also illustrated in the controversies over proposed empirical tests, as evidenced by Huntington's (2000) dismissal of Russett, Oneal, and Cox's (2000) effort to evaluate his argument.

[10] Huntington (1993, 38) provides an interesting discussion of the risk of conflict between Russia and Ukraine, worth quoting in full: "In 1991 and 1992 many people were alarmed by the possibility of violent conflict between Russia and Ukraine over territory, particularly Crimea, the Black Sea fleet, nuclear weapons and economic issues. If civilization is what counts, however, the likelihood of violence between Ukrainians and Russians should be low. They are two Slavic, primarily Orthodox peoples who have had close relationships with each other for centuries. As of early 1993, despite all the reasons for conflict, the leaders of the two countries were effectively negotiating and defusing the issues between the two countries. While there has been serious fighting between Muslims and Christians elsewhere in the former Soviet Union and much tension and some fighting between Western and Orthodox Christians in the Baltic states, there has been virtually no violence between Russians and Ukrainians." One might argue that Huntington's prediction was correct at the time, but that a prior common identity has drifted apart over time, or that tension has been exacerbated by competing pressures such as proposed Ukrainian North Atlantic Treaty Organization (NATO)

However, although I believe Huntington was wrong and exaggerated conflict across civilizations relative to the extent of conflict within civilizations, he is in other respects admirable in trying to make an explicit prediction, offered in advance of the events the theory is held to predict. In many cases, people confidently claim to have predicted things, but they only identify alleged successful predictions after events have already occurred. For example, we now see many statements claiming that the fall of the Afghan government was "predictable" after the Taliban seized control.[11] The arguments offered may well be valid and important. But unless they are actually made in advance, arguments that something "was predictable" are really post-dictions, made in retrospect. It is much easier to predict something after the fact, with the benefit of hindsight. It is also hard to resist the temptation to tailor our explanations around the known facts. Well-intended people often convince themselves that they believed something all along once they have observed what they wish to explain. This kind of "self-scoring" risks exaggerating people's view of how much we can actually explain and anticipate about the future.[12]

Alternatively, many statements about the future are so general that they are consistent with almost any outcome. Claims that certain events or outcomes are "possible" or "cannot be ruled out" ultimately tell us little beyond that the probability is not exactly 0 or 1. They do not rule out anything between these extremes or give us meaningful indication of how likely or unlikely the outcome may be. Philip Tetlock reminds us that "vague verbiage" make many predictions unclear or at best imprecise (Tetlock, Mellers, and Scoblic 2017).[13]

Predicting More Clearly

We can overcome the problems that plague informal prediction by trying to be more precise. For example, we can specify exactly what the predicted event or outcome entails (e.g., what is meant by conflict and what would be evidence for a trend or shift). We can try to be more precise on the timeline or interval for the prediction (e.g., state a period from date1 to date2 rather than more ambiguous statements such as "after the Cold War"). We can try to provide more specific likelihood estimates for events, or identify explicit odds or bets. Moreover, we can try to state distributions for outcomes under different scenarios. As Bertrand Russell (1923) alerts us to, being precise helps us realize and identify what is vague.[14]

Instead of just focusing on how persuasive explanations seem once facts are known, we can evaluate predictions on genuinely new information. There are famous examples in the natural sciences where theories are tested on new events. Eddington, for example, used a solar eclipse to test an implication of Einstein's general theory of relativity, based on the gravitational deflection of an object passing by the sun.[15] We may have few similar examples in our discipline, and many claims are probabilistic and not amenable to decisive tests. Yet, it is becoming more common to explicitly evaluate implications separately from theory development.[16] Researchers sometimes look at how models fare when applied to new data and try to clearly separate between training and testing data.[17] Others preregister an experiment or study in advance in order to tie their hands and resist any temptation to tinker with the original plan once more details of the results are known (e.g., Miguel et al. 2014).

In addition to the ability to predict individual events, it is also useful to focus on variation and our ability to identify when something does not occur. That someone consistently predicts economic doom, for example, does not mean that they anticipated the financial crisis in a meaningful way. Paul Samuelson (1966) reminded us that the predictive ability of stock markets depended not just on the recessions correctly anticipated but also how many of these anticipated recessions did not occur.[18]

Prediction Is Difficult—Especially about the Future—Yet Useful

Social scientists sometimes latch on to a lack of predictive ability as reflecting problems with research or approaches they dislike. John Lewis Gaddis (1992/93), for example, argued that the inability to anticipate the end of the Cold War demonstrated the limited usefulness of much international relations theory. But very often, the relevance of prediction in the social science itself is dismissed outright, as in Jaqueline Steven's assertion that "research aimed at political prediction is doomed to fail."[19]

membership that Huntington could not have foreseen at the time. Tetlock (1998) notes that experts on world politics often embrace counterfactuals that change history when these protect their forecast ("I was wrong, but I was almost right"). At a minimum, this raises question over the alleged enduring nature of civilizations claimed by Huntington.

[11] See, for example, https://www.newstatesman.com/politics/2021/08/afghanistan-diary-fall-kabul-was-predictable-if-you-were-there, Accessed May 15, 2022.

[12] See Tetlock (2005) for a more systematic overview on the predictive accuracy of political experts.

[13] The ironic term "definite maybe" illustrates the absurd combination of both absolute confidence and uncertainty. I was first introduced to "definite maybes" by The Kinks 1983 song, but there are earlier claims attributing this to the movie producer Samuel Goldwyn (although it is unclear if he actually said this, see https://quoteinvestigator.com/2015/08/01/definite-maybe/, Accessed May 15, 2022). My personal favorite example is the 1948 Norwegian children's song *Blåbærturen* [The bilberry picking trip] by Alf Prøysen. Here, a group of children picking billberries are frightened by rustling of the leaves, leading one to pronounce that "it might perhaps be a bear, it could well be an ox, and at least it is certain that it could possibly be a cow." In the end, it turns out to be their own dog, but of course the prediction is worded so that it could never be wrong, irrespective of any outcome (the song itself can be found at https://www.dailymotion.com/video/x1tfqma, Accessed May 15, 2022).

[14] Each of these is important, but attention to being explicit may not always suffice. In 2008, Bryan Caplan and Raphael Frank agreed to a bet on whether the number of deaths from terrorism and riots in France between 2008 and 2018 would exceed 500. In 2018, however, they failed to agree on the actual number of deaths over the period. Caplan cited 256 deaths from terrorism from the Global Terrorism Database and the apparent absence of deaths from rioting as evidence that he was correct in predicting that the number would be lower than 500. In contrast, Frank claimed that the number exceeded 500, presenting a list including deaths from Germanwings Flight 9525 (a flight from Spain to Germany that was crashed by the co-pilot on French soil) as well as many people killed by the French police as deaths from riots. See https://www.econlib.org/archives/2018/05/i_win_my_french.html, Accessed May 15, 2022.

[15] See the detailed description on https://en.wikipedia.org/wiki/Eddington_experiment, Accessed May 15, 2022 and the 2008 BBC dramatized account https://www.bbc.co.uk/programmes/b00ft62c, Accessed May 15, 2022.

[16] Popper (1959 [1934]) highlighted the difference between the context of discovery (i.e., "the act of conceiving or inventing a theory") and the context of verification, which could only be satisfactorily carried out on genuinely new data.

[17] For example, Bishop (2006) and Ward, Greenhill, and Bakke (2010).

[18] Samuelson (1966, 92) noted that it was an understatement that "market downturns predicted four out of the last five recessions … [since] Wall Street indexes predicted nine out of the last five recessions! … And its mistakes were beauties!" See also McArdle's discussion of predictions of the 2008 financial crisis (https://www.theatlantic.com/business/archive/2008/10/future-shock/4235/, Accessed May 15, 2022).

[19] https://www.nytimes.com/2012/06/24/opinion/sunday/political-scientists-are-lousy-forecasters.html, Accessed May 15, 2022).

There are a number of reasons why people argue that the social sciences cannot aspire to prediction. I cannot enumerate all these arguments in full here, but some of the more prominent involve the role of human agency, overwhelming complexity, and Karl Popper's (1972) idea that human systems are more like amorphous and shifting clouds than deterministic mechanical clocks. However, it is notable that different arguments often converge on the implication that predicting social phenomena will be difficult. According to critics, prediction is unrealistic since it is difficult, and we should settle on trying to "explain" observed facts as an alternative.[20] If anything, the current trends in the social science explicitly favor post-diction—can we find exogenous variation in a past event that can account for contemporary outcomes, and then claim to have better causal inference devoid of endogeneity problems? This may be a good strategy for publishing in a prestigious journal, even if our findings are essentially similar to previous observational studies. However, waiting for exogenous shocks is unlikely to be helpful for forecasting. Grant applications that focus on predictions about the future also tend to have a particularly hard time.

It is trivially true that it is difficult to predict, especially about the future.[21] It is also true that it is difficult to do research with the future as evidence. As such, it is good advice to steer students toward projects that can be examined with empirical material that we have observed already. However, inherent difficulty is by itself not a good argument for dismissing prediction, and it does not make a low ability to predict any less relevant. If our theoretical arguments fail to provide a good or simple guide to predicting future events, then why should we be so confident that they are correct?

Rather than sweeping limited predictive ability under the carpet, we should try to understand why we face problems in making predictions. Some things may be inherently unpredictable, but we can learn something if we try to establish if this actually is the case and, if so, why. In some cases, we may have theoretical reasons to expect indeterminacy. [22] Erik Gartzke (1999), for example, has argued that war onset in crises is inherently unpredictable, since the relevant features that determine if we see war will not be observable to actors at time. If war results when offers made are "too low," or below the opponent's reservation point for war as an outside option, and actors simply guess about reservation points with a random distribution, then overestimates should be as likely as underestimates (see also Chen et al. n.d.).

We can learn useful things by examining more closely what may be unpredictable and what is not.[23] Gartzke's bargaining framework highlights indeterminacy in whether an agreement will be reached in a situation where two actors are already in a dispute. However, we may be able to identify a great deal of regular patterns in terms of which sets of dyads are more likely to have contentious issues flare-up and a higher risk of crises that may see use of force (e.g., Gleditsch and Ward 2013). This can in turn be helpful for analyses of the impact of potential conflict management efforts. For example, although conflict over contentious issues such as territory often gives rise to recurrent disputes, these appear to be amenable to be solved in a stable manner through formal agreements between the antagonists (e.g., Mitchell and Hensel 2007). It is much more difficult to reach effective agreements when repeated interstate conflict arises from domestic conflict, since the stability of any agreement signed will depend on non-state actors partly outside the control of the states (Schultz 2010). This type of analysis can also help us establish when history can be our guide and when the past is less useful as a guide for the future (e.g., King and Zeng 2007). Much has been made of the potential problems of "black swans," or events that cannot be predicted based on previous observed events (e.g., Taleb 2007). However, if "white swans" that can be predicted from past events are more numerous, then focusing on them may be more useful in practice for many purposes (see Gleditsch 2017; Guo, Gleditsch, and Wilson 2018). Former International Studies Association (ISA) President Bruce Bueno de Mesquita (2011) has done extensive work documenting on how combining game-theoretic models with detailed information on actors, preferences, and power can predict to outcomes of negotiations. Finally, if we find that a model predicts works well in one setting but less well in another, then we may learn something about the specific scope conditions. For example, Drew Bowlsby et al. (2020) report that many conflict prediction models that have performed well in the past appear to do less well at the present, and this helps set a useful research agenda on what may have changed in conflict over time.

In short, predictive scoring allows us to do better research; it guards against overconfidence in our conclusions, and it helps us recognize limits to prediction and change over time.

Predicting Better

Even if prediction is difficult, are there ways to make it relatively better? Over the last decades, there have been major advances in research on prediction. Despite complexity and Popper citing clouds as less predictable systems, weather forecasting is really a success story. One hundred years ago, the father of weather forecasting Lewis Fry Richardson (2007 [1922]) devised equations to represent atmospheric flow. Since the 1950 advances computing power has made it possible to apply these to data to forecast weather ahead.[24]

In the social sciences, election forecasts have gained prominence, and we now know much about what approaches work relatively better and have more realistic measures of our uncertainty.[25] In international studies, we see considerable interest in conflict prediction.[26] An important impetus here has been the State Failure Task Force—now Political Instability Task Force (PITF)—which since

[20] See, for example, Dowding and Miller (2019) and Schrodt (2014).

[21] This is another famous quote of unclear origin, attributed to many including Niels Bohr and Yogi Bera.

[22] A classical case for indeterminacy in economics is the so-called efficient markets hypothesis, or the argument that a fully efficient market would incorporate all the relevant information in pricing stocks (Fama 1970).

[23] Analyses of the evidence for efficient markets are somewhat mixed, and this has led to useful discussions about the level of aggregation at which markets can be said to approximate efficiency as well as understanding how bubbles can arise and apparently irrational individual behavior (e.g., Shiller 2000). Evidence on behavioral tendencies that defy normal definitions of rationality has influenced research on international relations. Johnson (2004) suggests that human tendencies to overconfidence imply a higher risk of the types of mistakes that could lead to a failure to reach agreement. See also Lake (2010/11) on behavioral economics and possible explanations of the Iraq war.

[24] Richardson also did important work on modeling violent conflict, noting that the frequency–severity distribution of many violent events followed a regular pattern (i.e., a power law) where the severity of events is inversely proportional to their frequency (Richardson 1948), and that the distribution of war outbreaks over time was consistent with a random Poisson distribution (Richardson 1944).

[25] See, for example, Silver (2012) for a popular overview or Gelman et al. (2020) for a more recent assessment of challenges.

[26] Anticipating future conflict has often been held up as a goal for conflict research (e.g., Singer and Wallace 1979), but outside informal predictions about individual events we have had few attempts of comprehensive or global predictions of conflict risks until the early 1990s. Metternich, Gleditsch, and Dworschak (2021) provide a more detailed survey on forecasting in international relations.

its inception in the 1990s has tried to look at what social science theories can contribute to help inform risk assessment on conflict and instability for the US Federal Government (see Esty et al. 1998). My predecessor as ISA President, the late Ted Gurr, played an important role in developing this project. One prominent example is the work of Jack Goldstone et al. (2010), highlighting how distinguishing between different types of political institutions can help identify the risk of political instability.[27]

Predictive conflict research is a vibrant area, and we now have a range of different approaches, often allowing for direct comparisons on the same data and events and dialogue between contributions (see, e.g., Ward et al. 2013; Hegre et al. 2017; Vesco et al. 2022). In my own modest contributions to conflict prediction, I have looked at issues such as the ability to improve on baseline models of civil conflict by including measures of plausible grievances and the potential contribution from real-time information from event data coded from news media (see, e.g., Buhaug, Cederman, and Gleditsch 2014; Chiba and Gleditsch 2017). In Buhaug, Cederman, and Gleditsch (2014), for example, we contrasted a model looking at group-based horizontal inequalities (HI) as an influence on grievances that could lead to civil war with standard models of purely state-based characteristics or vertical inequality (VI) measures. Using annual data on countries for 1960–1999 for model training, we then used the estimated probabilities from the models for 1999 to predict to civil war onset out-of-sample for the 2000–2009 period. While the VI model predicted four out of the twenty-six onsets over the period, the HI model identified eight. This is obviously not perfect foresight, but it is still twice as many as the alternative and not a negligible difference.

We also have improvements in general methodology and an understanding of the traits and types of reason that allow some "superforecasters" to predict better than others. In particular, Tetlock and collaborators argue that forecasting is improved when we break up problems into smaller parts and reason separately about these, think about future events in terms of scenarios instead of single outcomes, and use Bayesian updating to adjust initial predictions as we learn more information (e.g., Tetlock and Gardner 2015).

In sum, predicting political events remains difficult, but it is clearly possible to do better.

Prediction and Policy

At this point, I would like to go back to policy. Grant applications often ask researchers to detail broader implications, and many are very eager to emphasize that research is important if it is "policy relevant."[28] My usual response to calls to be more policy relevant is that research can only be relevant to policy objectives that are clearly specified in advance. Research may be able to tell you whether one proposed policy A is more likely to be helpful for an objective than another policy B. However, in many cases, we have multiple objectives. In some cases, policies A and B may also entail costs of a different kind than the potential benefits. In many cases, researchers are eager to talk about policy objectives, but research alone cannot tell you what your priorities ought to be or how to weigh different objectives and concerns against one another. To use a pandemic example, restrictive public health measures can reduce mortality and disease, but if they also reduce economic activity, undermine education, and increase mental health problems, what metric can we use to weigh lives saved against economic losses and loss of education? Or in Afghanistan, what should be the relative weight in US policy of minimizing active conflict, protecting women's rights, and curtailing the cost of ongoing military operations?

In short, policy priorities must inevitably be political decisions, and researchers have no special competence here over the general public. Many people implicitly assume that everyone would be likely to accept their own views as inherently reasonable if they had the same information. There may be areas where we all would converge if we had the same information. However, knowledge problems and political priorities are distinct, and I for one have many views and preferences that I know are not widely shared. That policy must be political is as trivially true as the claim that prediction is difficult. To dismiss prediction because it is difficult is a cop-out. But so is dismissing interest in policy as something that lies outside research.

A more useful and interesting response is to think about how we can have better debates and analysis of policy. What does this mean in practice? In short, I think it is useful to separate debate about

1) policy objectives, or what we wish to achieve;
2) policy alternatives, or what one might think that we could do;
3) policy consequences, or what we think different policies may lead to; and
4) cost–benefit analysis—how do we weigh consequences of one type against each other, and different objectives against each other if there may be tension between them.

All of these points are important, but my main interest here is in (3)—namely that policy consequences are future outcomes. A claim that policy A will have consequences X is thus a prediction.

For example, in early 2022, we can find many statements on the consequences of imposing sanctions on Russia for perceived aggression against Ukraine before the invasion, or retaliation after Russia invaded, either in terms of the costs imposed or effects on Russian behavior.[29] But if it is difficult to predict about the future, then it is also difficult to predict policy consequences. Still, it is common to make claims about the consequences of policies with no regard for uncertainty, which Charles Manski (2013, 2020) calls "policy prescriptions with unrealistic certitude."

The analogy between prediction and claims about policy consequences is helpful not just in stressing uncertainty, but

[27] This is sometimes referred to as the PITF model, although the PITF actually considers a range of different models. Walter (2022) discusses the relevance of this research on political institutions on the risk of civil war and political violence in the United States, as some observers argue that democracy has "declined precipitously in the United States." See https://www.washingtonpost.com/news/monkey-cage/wp/2018/07/03/one-third-of-the-worlds-population-lives-in-a-declining-democracy-that-includes-americans/, Accessed May 15, 2022.

[28] The UK Economic and Social Research council, for example, requires research applications to submit plans for "economic and societal impact, which is the demonstrable contribution that excellent social and economic research has on society and the economy, and its benefits to individuals, organisations or nations," including "instrumental impact - influencing the development of policy, practice or services, shaping legislation and changing behaviour" and "conceptual impact – contributing to the understanding of policy issues and reframing debates." See https://www.ukri.org/councils/esrc/impact-toolkit-for-economic-and-social-sciences/defining-impact/, Accessed May 15, 2022.

[29] For statements on the likely effectiveness or ineffectiveness of sanctions, see, for example, https://www.politico.com/news/magazine/2022/01/21/russia-sanctions-stop-putin-energy-markets-us-invasion-527524, Accessed May 15, 2022 and https://www.independent.co.uk/voices/biden-putin-conference-ukraine-sanctions-b1971627.html, Accessed May 15, 2022.

also in allowing us to bring to the table some of the advances in research on explicit predictions to evaluate claims about policy. If you think severe sanctions would make Russia less assertive in demands on Ukraine and the West, what are the premises and assumptions that underpin the conclusion? To evaluate predictions, they should be transparent and replicable. We need to "show our work," in order to allow others to understand how we how we get to our conclusions.[30] Only by spelling out assumptions and our degree of confidence in them can we get a clear sense of the premises for a policy as well as uncertainty and potential pitfalls.

In many cases we have competing policy advice or prescriptions—or in Manski's terminology, "dueling certitudes." For example, someone might assert that even severe sanctions are unlikely to be effective against Russia, if the costs are not suffered directly by leaders (e.g., Kaempfer, Lowenberg, and Mertens 2004). In such cases, looking at claims as predictions and the details in how they are produced can often help us understand the sources for the divergence. For example, do we have different conclusions because people make opposite assumptions about different things, such as the bite that sanctions would have on decisionmakers? Or do we have different objectives in the first place, such as stressing the punitive effect sanctions as an end itself versus likely effects on changing behavior? Disagreement can also arise over other features that may influence the final outcome, such as the likelihood that high energy prices and demand for natural gas will undermine willingness to engage in the most costly sanctions.[31]

Research and prediction alone cannot directly tell you what decisions you should make—we need to specify objectives first. However, it can help make the basis for decisions more explicit and allow for a more informed debate about policy and what people actually disagree on. People may share the same objectives and priorities, yet disagree on the consequences of proposed actions. Or people may argue for different policies and actions precisely because they disagree on the objectives themselves, or people may have similar objectives but weight relative priorities differently. If they have different beliefs about the state of the world, can we specify scenarios and assign prior likelihoods to them? What would constitute more evidence for one perspective over another as time goes by?

Some people might argue that working through policy as predictions and with transparency is very cumbersome and would make things much more complicated. To this I can only say "yes," but this is an inherent feature and not a bug. The fact that topics are important is not a good justification for thinking or acting fast; indeed, transparent thinking and reasoning about uncertainty seem all the more important, not less. In Daniel Kahneman's (2011) terminology, we need to rely more on system II (i.e., slower, analytical reasoning) rather than system I (i.e., fast, instinctive/intuitive perceptions).

It is sometimes suggested that people prefer simple answers and that researchers can better support policy if they offer simplification and clear decisions as policy audiences do not tolerate uncertainty and ambiguity (Manski 2020). The idea that policy prescriptions should exaggerate certainty to be persuasive is sometimes associated with Harry Truman, who is supposed to have called for a one-handed economist, deploring their tendency to bring up competing concerns rather than simple decisions.[32] It is questionable if decisionmakers really are averse to uncertainty or that what they want from researchers is to tell them what their priorities ought to be or what to decide. But even if it were, we should be hesitant to claim political authority from our role as academics or trade off influence from recognizing the appropriate role of politics and uncertainty.

If we wish to be relevant to society, then we should also be candid about the limits to our knowledge. For example, in the 2016 US presidential elections, some forecasts such as the Princeton Election Consortium led by neuroscientist Sam Wang claimed that there was a 99 percent probability that Hillary Clinton would win. However, Nate Silver's 538 forecast was much more explicit on the uncertainty and clearly identified about a one-third chance of a Donald Trump victory, given the many potential chances for an electoral college majority even with a considerably lower popular vote share.[33] Uncertainty must be large for most issues relevant to international studies. If we wish our research to be useful for policy, then we should acknowledge and try to reflect uncertainty.

Adding Policy to Prediction

The relevance of prediction to policy, I hope, should be obvious at this point. However, the claim that we should not have one without the other also implies that taking policy more seriously must yield some benefits for prediction. Why would this be the case?

A first point is that predictions are most useful when we can learn something directly from them. In many cases, it is more helpful to show how a single feature can provide additional predictive impact, or what has "oomph" in the words of Deirdre McCloskey and Stephen Ziliak (1996). The most useful predictions are not always those that maximize overall predictive fit, especially if we do not fully understand what accounts for the improvement. For example, Netflix held a competition to determine the best algorithms for suggestions for content to subscribers, yet declined to use the actual winning algorithm as it was deemed to be too complicated.[34] Understanding the importance of single inputs to an algorithm might be more useful for content develop-

[30] During the COVID-19 pandemic, for example, there was an important debate over whether one should follow a "first-doses-first" (FDF) approach to vaccination or stick closely to the dosing interval that had been used in the clinical trials. Arguments in favor of FDF highlighted the opportunities for maximizing coverage and the likely benefits to immunity from a longer interval between doses, based on data from other vaccines. Arguments against typically emphasized the lack of direct data for anything outside what had been done in the vaccine trials (which had clear incentives to choose a short interval and larger doses to demonstrate effectiveness and ensure faster approval rather than examine optimum intervals or dosing). Cowen noted in January 2021 that one could only evaluate the strength of the counterarguments based on explicit expected value calculations, and how these had not been specified. His call for analysts to "show your work" has broader relevance. See https://marginalrevolution.com/marginalrevolution/2021/01/first-doses-first-show-your-work.html, Accessed May 15, 2022.

[31] As of the time of writing, Russia is expected to have increased earnings from energy sales despite sanctions, given the increase in prices following the invasion of Ukraine. See https://www.reuters.com/business/energy/russia-expects-earn-96-bln-more-april-due-high-oil-prices-2022-04-05/, Accessed May 15, 2022.

[32] There appears to be limited recorded evidence that Truman ever said this. Interestingly, the first printed record of a similar statement—appearing in a serialized editorial published in multiple newspapers in 1936—does not mention economists, but criticizes the Republican party for endorsing two inconsistent objectives on soil conservation, calling for a "one-armed platform writer." See https://quoteinvestigator.com/2019/04/10/one-handed/, Accessed May 15, 2022. Manski (2011) discusses a similar likely apocryphal statement attributed to Lyndon B. Johnson, who is reported to have said "ranges are for cattle - give me a number" when presented with a forecast with a range of values.

[33] See Gelman and Azari (2017) for a more extended discussion.

[34] See https://www.wired.com/2012/04/netflix-prize-costs/, Accessed May 15, 2022. In the case of conflict research, Kapoor and Narayanan (2021) argue that many applications suggesting major improvements from machine learning al-

ment than an algorithm that works well but in ways that are harder to interpret (see also Rudin 2019).

A second related point is that predictions sometimes, if not always, are intended to speak to policy audiences. This is a very different audience than academic researchers, without formal training in statistics or data science. This does not mean that people are inherently unable to deal with uncertainty and technical details. However, researchers would do well in trying to understand how to communicate research better. The 2021 ISA book of year by Christoph Meyer, Chiara de Franco, and Florian Otto (2019), for example, provides an extended analysis of what makes early warnings relatively easier to understand or more persuasive to nonacademic target audiences.

A third point stems from the argument that prediction is more difficult in situations that involve human decision-making or human agency. Instead of seeing this as an argument against the possibility of prediction, we could consider opportunities for incorporating policy decisions and agents in prediction scenarios. People act on what they believe or know at the time, not what we learn after the fact.[35] Thus, if we believe that individuals are more likely to believe A than B, or weight concerns C more than D, are they more likely to make decisions X, Y, or Z? Can we identify what other actors are likely to infer from decisions? Can we specify a range of possible scenarios that might happen and how the likelihood of decisions shift with specific events or new information? Thinking systematically about predicting decision-making in this way can help us identify factors that influence the final outcomes we wish to predict.

Policy and Prediction: Doing One Better with the Other

I have argued that virtually any effort to talk about policy will entail claims about prediction, and we can do one better by doing more of the other. Second, I have argued that predictions can also benefit from attention to learning, better communicating their potential value to others, and thinking about how policy and decisions can make influence outcomes and make specific predictions more or less likely to come through. Working through policy as predictions will not inevitably make us more likely to converge on decisions—doing so would require first reaching agreements on policy objectives and priorities. And if we do not, let us recognize that disagreement on objectives can be perfectly acceptable. Democracy entails embracing pluralism and the fact that others may not have the same objectives as I do or weight things differently. But at least we can have a more informed debate about what people disagree on and why. Former ISA president David Lake (2011) has drawn attention to how the focus on "-isms" and "rhetorical competition" draws too much attention to selective evidence primarily aimed at demonstrating superiority and takes us away from the more important task of evaluating the logic of theories and their ability to explain (see also Wagner 2007). Forcing us to predict to common specified problems and the same data can help us have more dialogues and fewer monologues. And greater transparency and debates might ultimately lead to better decisions, as research on groupthink and cognitive biases suggests.[36]

I started this talk with a love and marriage analogy. If you are very perceptive, you will latch on to the fact that marriage is really akin to a contract (e.g., Grossbard-Shechtman and Lemennicier 1999). If love really truly was everlasting and fully certain, then lovers would have little need to get married. However, in real life, marriage can be a useful contract, since we do not know if the other party may have changing views or be tempted to fall in love with someone else. Marriage makes it harder to walk away from a relationship, and gives each party some rights if the other should not hold up their part. Well-structured marriage can allow love to flourish in good and not so good days.

Policy relevance in research may be as tempting as love at first sight. However, if researchers want to be relevant to society and policy in a transparent manner, then they would do well commit itself to a social contract by offering clearer and explicit predictions over policy.

Funder Information

Funding for this article was provided by Economic and Social Research Council(ES/S007156/1).

References

ALLISON, GRAHAM T. 2017. *Destined for War: Can America and China Escape Thucydides's Trap?* Boston, MA: Houghton Mifflin.

BISHOP, CHRISTOPHER M. 2006. *Pattern Recognition and Machine Learning*. New York: Springer.

BUENO DE MESQUITA, BRUCE M. 2011. "A New Model for Predicting Policy Choices: Preliminary Tests." *Conflict Management and Peace Science* 28 (1): 65–85.

BUHAUG, HALVARD, LARS-ERIK CEDERMAN, AND KRISTIAN SKREDE GLEDITSCH. 2014. "Square Pegs in Round Holes: Inequalities, Grievances, and Civil War." *International Studies Quarterly* 58 (2): 418–31.

BOWLSBY, DREW, ERICA CHENOWETH, CULLEN HENDRIX, AND JONATHAN D. MOYER. 2020. "The Future is a Moving Target: Predicting Political Instability." *British Journal of Political Science* 50 (4): 1405–17.

CHEN, CHONG, JORDAN ROBERTS, SHIKSHYA ADHIKARI, VICTOR ASAL, KYLE BEARDSLEY, EDWARD GONZALES, AND NAKISSA JAHANBANI et al. n.d. "Tipping Points: Challenges in Analyzing International Crisis Escalation." *International Studies Review* in press.

CHIBA, DAINA, AND KRISTIAN SKREDE GLEDITSCH. 2017. "The Shape of Things to Come? Extending the Horizontal Inequality Model to Prediction with Event Data." *Journal of Peace Research* 54 (2): 275–97.

CLARK, CORY J., AND PHILLIP E. TETLOCK. n.d. "Adversarial Collaboration: The Next Science Reform." In *Political Bias in Psychology: Nature, Scope, and Solutions*, edited by C.L. Frisby, R.E. Redding, W.T. O'Donohue and S.O. Lilienfeld. New York: Springer. https://www.researchgate.net/publication/352101647_Adversarial_Collaboration_The_Next_Science_Reform

DOWDING, KEITH, AND CHARLES MILLER. 2019. "On Prediction in Political Science." *European Journal of Political Research* 58 (3): 1001–18.

ESTY, DANIEL C., JACK GOLDSTONE, TED ROBERT GURR, BARBARA HARFF, PAMELA T. SURKO, ALAN N. UNGER, AND ROBERT CHEN. 1998. "The State Failure Project: Early Warning Research for US Foreign Policy Planning" In *Preventive Measures: Building Risk Assessment and Crisis Early Warning Systems*, edited by John L. Davies and Ted Robert Gurr, 27–38. Boulder, CO: Rowman and Littlefield.

FAMA, EUGENE. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (2): 383–417.

GADDIS, JOHN LEWIS. 1992/93. "International Relations Theory and the End of the Cold War." *International Security* 17 (1): 5–58.

GARTZKE, ERIK A. 1999. "War Is in the Error Term." *International Organization* 53 (3): 567–87.

gorithms over simpler linear model alternatives arise due to "data leakage," where information from the testing data has seeped into the training data, for example, through common imputation on the full data set or variables that proxy for the outcomes such as conflict types.

[35] See Møller and Skaaning (2021) on the Ulysses principles and ways to overcome the temptation to rely on ex post information not available at the time in explaining historical events.

[36] Groupthink was originally discussed by Janis (1972). Recent research highlights the potential advantages of "adversarial collaboration" (Clark and Tetlock n.d.), and the potential value of taking into account contrarian minorities in improving forecasts (see Powell et al. 2022).

Gartzke, Erik A., and Kristian Skrede Gleditsch. 2006. "Identity and Conflict: Ties that Bind and Differences that Divide." *European Journal of International Relations* 12 (1): 53–87.

Gelman, Andrew, Jessica Hullman, Christopher Wlezien, and George E. Morris. 2020. "Information, Incentives, and Goals in Election Forecasts." *Judgment and Decision Making* 15 (5): 863–80.

Gelman, Andrew, and Julia Azari. 2017. "19 Things We Learned from the 2016 Election." *Statistics and Public Policy* 4 (1): 1–10.

Gleditsch, Kristian Skrede. 2017. "Ornithology and Varieties of Conflict: A Personal Retrospective on Conflict Forecasting." *Peace Economics, Peace Science and Public Policy* 23 (4): 1–4.

Gleditsch, Kristian Skrede, and Michael D. Ward. 2013. "Forecasting is Difficult, Especially about the Future: Using Contentious Issues to Forecast Interstate Disputes." *Journal of Peace Research* 50 (1): 17–31.

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, and Jay Ulfelder et al. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science* 54 (1): 190–208.

Grossbard-Shechtman, Shoshana, and Bertrand Lemennicier. 1999. "Marriage Contracts and the Law-and-Economics of Marriage: An Austrian Perspective." *Journal of Socio-Economics* 28 (6): 665–90.

Guo, Weisi, Kristian Skrede Gleditsch, and Alan Wilson. 2018. "Retool AI to Forecast and Limit Wars." *Nature* 562 (7727): 331–33.

Hechter, Michael, Timur Kuran, Charles Tilly, Randal Collins, and Edgar Kieser. 1995. "Symposium on Prediction in the Social Sciences." *American Journal of Sociology* 100 (6): 1520–1615.

Hegre, Håvard, Nils W. Metternich, Håvard Mokleiv Nygård, and Julian Wucherpfennig. 2017. "Introduction: Forecasting in Peace Research." *Journal of Peace Research* 54 (2): 113–24.

Hempel, Carl. 1966. *Philosophy of Natural Science*. Upper Saddle River, NJ: Prentice Hall.

Huntington, Samuel P. 1993. "The Clash of Civilizations?" *Foreign Affairs* 72 (3): 22–49.

———. 2000. "Try Again: A Reply to Russett, Oneal & Cox." *Journal of Peace Research* 37 (5): 609–10.

Janis, Irving L. 1972. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*. Boston, MA: Houghton Mifflin.

Johnson, Dominic D.P. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press.

Kaempfer, William H., Anton D. Lowenberg, and William Mertens. 2004. "International Economic Sanctions against a Dictator." *Economics & Politics* 16 (1): 29–51.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Penguin.

Kalyvas, Stathis, and Scott Straus. 2020. "Stathis Kalyvas on 20 Years of Studying Political Violence." *Violence: An International Journal* 1 (2): 389–407.

Kapoor, Sayash, and Arvind Narayanan. 2021. "(Ir)Reproducible Machine Learning: A Case Study." Accessed May 15, 2022. https://reproducible.cs.princeton.edu/irreproducibility-paper.pdf.

King, Gary, and Langche Zeng. 2007. "When Can History Be Our Guide? The Pitfalls of Counterfactual Inference." *International Studies Quarterly* 51 (1): 183–210.

Lake, David A. 2010/11. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War," *International Security* 35 (3): 7–52.

———. 2011. "Why 'Isms' Are Evil: Theory, Epistemology, and Academic Sects as Impediments to Understanding and Progress." *International Studies Quarterly* 55 (2): 465–80.

Manski, Charles F. 2011. "Policy Analysis with Incredible Certitude." *The Economic Journal* 121 (554): F261–89.

———. 2013. *Public Policy in an Uncertain World: Analysis and Decisions*. Cambridge, MA: Harvard University Press.

———. 2020. "The Lure of Incredible Certitude." *Economics & Philosophy* 36 (2): 216–45.

McCloskey, Deirdre N., and Stephen T. Ziliak. 1996. "The Standard Error of Regressions." *Journal of Economic Literature* 34 (1): 97–114.

Mearsheimer, John J., and Stephen M. Walt. 2013. "Leaving Theory Behind: Why Simplistic Hypothesis Testing is Bad for International Relations." *European Journal of International Relations* 19 (3): 427–57.

Meyer, Christoph, Chiara de Franco, and Florian Otto. 2019. *Warning about War: Conflict, Persuasion and Foreign Policy*. Cambridge: Cambridge University Press.

Metternich, Nils W., Gleditsch, Kristian Skrede, and Christoph Dworschak. 2021. "Forecasting in International Relations." In *Oxford Bibliographies in International Relations*, edited by Patrick James. New York: Oxford University Press.

Miguel, Edward, Colin Camerer, Katherine Casey, Joshua Cohen, Kevin M. Esterling, Alan Gerber, and Rachel Glennerster et al. 2014. "Promoting Transparency in Social Science." *Science* 343 (6166): 30–31.

Mitchell, Sara McLaughlin, and Paul R. Hensel. 2007. "International Institutions and Compliance with Agreements." *American Journal of Political Science* 51 (4): 721–37.

Møller, Jørgen, and Svend-Erik Skaaning. 2021. "The Ulysses Principle: A Criterial Framework for Reducing Bias When Enlisting the Work of Historians." *Sociological Methods & Research* 50 (1): 103–34.

Popper, Karl. 1959 [1934]. *The Logic of Scientific Discovery*. Camberley: Hutchinson.

Popper, Karl R., ed. 1972. "Of Clouds and Clocks: An Approach to the Problem of Rationality and the Freedom of Man." In *Objective Knowledge: An Evolutionary Approach*, 206–55. Oxford: Clarendon Press.

Powell, Ben, Ville Satopää, Niall J. MacKay, and Philip Tetlock. 2022. Skew-Adjusted Extremized-Mean: A Simple Method for Identifying and Learning from Contrarian Minorities in Groups of Forecasters." Accessed May 15, 2022. https://ssrn.com/abstract=4004029 or http://dx.doi.org/10.2139/ssrn.4004029.

Richardson, Lewis Fry. 1944. "The Distribution of Wars in Time." *Journal of the Royal Statistical Society* 57 (3–4): 242–50.

———. 1948. "Variation of the Frequency of Fatal Quarrels with Magnitude." *Journal of the American Statistical Association* 43 (244): 523–46.

Richardson, Lewis Fry. 2007 [1922]. *Weather Prediction by Numerical Process*. Cambridge: Cambridge University Press.

Rudin, Cynthia. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1: 206–15.

Russell, Bertrand. 1923. "Vagueness," *Australasian Journal of Psychology and Philosophy* 1 (2): 84–92.

Russett, Bruce M., John R. Oneal, and Michaelene Cox. 2000. "Clash of Civilizations, or Realism and Liberalism Déjà Vu? Some Evidence." *Journal of Peace Research* 37 (5): 583–608.

Samuelson, Paul. 1966. "Science and Stocks." *Newsweek*, September 19. 92.

Santayana, George. 1905. *Reason in Common Sense*. London: Constable.

Schrodt, Philip A. 2014. "Seven Deadly Sins of Contemporary Quantitative Political Analysis." *Journal of Peace Research* 51 (2): 287–300.

Schultz, Kenneth A. 2010. "The Enforcement Problem in Coercive Bargaining: Interstate Conflict over Rebel Support in Civil Wars," *International Organization* 64 (2): 281–312.

Shiller, Robert J. 2000. *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

Silver, Nate. 2012. *The Signal and the Noise: Why Most Predictions Fail – but Some Don't*. New York: Penguin.

Singer, J. David, and Michael D. Wallace. 1979. *To Augur Well: Early Warning Indicators in World Politics*. Beverly Hills, CA: SAGE.

Taleb, Nassim N. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Tetlock, Philip E. 1998. "Close-Call Counterfactuals and Belief-System Defenses: I was Not Almost Wrong but I was Almost Right." *Journal of Personality and Social Psychology* 75 (3): 639–52.

———. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

Tetlock, Philip E., Barbara A. Mellers, and J. Peter Scoblic. 2017. "Bringing Probability Judgments into Policy Debates via Forecasting Tournaments." *Science* 355 (6324): 481–83.

Tetlock, Philip E., and Dan Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. New York: Random House.

Vesco, Paola, Håvard Hegre, Michael Colaresi, Remco Bastiaan Jansen, Adeline Lo, Gregor Reisch, and Nils Weidmann. 2022. "United They Stand: Findings from an Escalation Prediction Competition," *International Interactions* 48 (1). Accessed May 15, 2022. https://doi.org/10.1080/03050629.2022.207074.

Wagner, R. Harrison. 2007. *War and the State: The Theory of International Politics*. Ann Arbor, MI: University of Michigan Press.

Walt, Stephen M. 2005. "The Relationship between Theory and Policy in International Relations." *Annual Review of Political Science* 8: 23–48.

Walter, Barbara F. 2022. *How Civil Wars Start: And How to Stop Them*. New York: Random House.

Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "Perils of Policy by P-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47 (4): 363–75.

Ward, Michael D., Nils W. Metternich, Cassy L. Dorff, Max Gallop, Florian M. Hollenbach, Anna Schultz, and Simon Weschle. 2013. "Learning from the Past and Stepping into the Future: Toward a New Generation of Conflict Prediction." *International Studies Review* 15 (4): 473–90.