

Chapter 3

The Simple, the Trivial, and the Insightful: Field Dispatches from a Formal Theorist

Branislav L. Slantchev¹

This advice begins with a disclaimer. I have a view of formal modeling that – while being shared by many fellow modelers and philosophers of science – is at variance with what seems to be the prevalent one in the discipline. All my thoughts on the use of formal models in research are bound up with that view, and are probably not useful if one does not share it.

Contrary to popular opinion, the biggest hurdle to effective modeling is not the absence of advanced mathematical skills. Instead, the problem lies with a hazy conception – shared by both proponents and critics – of what models are supposed to accomplish. The received view in the discipline seems to be that the primary purpose of (formal) models is to make predictions that are then ‘tested’ empirically.² On this account, models are hypothesis-generating machines – insert assumptions, crank through solutions, spit out predictions – and the benefit of formalism is to make the process more rigorous, and thus more scientific.³ Since ‘research practice in political science currently revolves around theory testing’, the sole value of the model is nearly always taken to lie in its ability to withstand empirical scrutiny.⁴

¹ I thank David Wiens for many useful discussions and William Clark for his comments on this chapter. Prepared for the SAGE Handbook for Political Methodology.

² Jim Granato and Frank Scioli, ‘Puzzles, proverbs, and omega matrices: the scientific and social significance of Empirical Implications of Theoretical Models (EITM)’, *Perspectives on Politics*, 2, 2 (2004), pp. 313–23, p. 315; Rebecca B. Morton, *Methods & Models: A Guide to the Empirical Analysis of Formal Models in Political Science* (Cambridge University Press, Cambridge, 1999); Donald P. Green and Ian Shapiro, *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science* (Yale University Press, New Haven, 1994). See James Johnson, ‘Models among the political theorists’, *American Journal of Political Science*, 58, 3 (2014), pp. 547–60 on what he calls ‘the standard rationale’ for models in political science.

³ Nancy Cartwright, *The Dappled World: A Study of the Boundaries of Science* (Cambridge University Press, Cambridge, 1999), p. 184.

⁴ Kevin A. Clarke and David M. Primo, *A Model Discipline: Political Science and the Logic of Representations* (Oxford University Press, Oxford, 2012), p. 11; Johnson, ‘Models among the political theorists’, p. 80.

This view is incoherent in logic and unimplementable in practice.⁵ It offers an impoverished interpretation of the role of models in research by denying outright their *raison d'être*: efficient and effective communication. Overcoming the fundamental hurdle to modeling requires one to recognize that models are merely specific arguments.

Models are Arguments

A model is not evaluated if its predictions are not analyzed, regardless of how true the assumptions of the model are believed to be.⁶

Models are closed deductive systems, which simply means that their conclusions follow from their premises. Given these premises, the conclusions are true irrespective of empirical referents. Models cannot be incorrect when their inferential rules are followed. No amount of ‘testing’, no matter how carefully designed, can alter this basic fact. The notion that hypothesis-testing somehow confers validity on a model or that it constitutes the core element in scientific practice has been roundly debunked.⁷

Moreover, all premises in these deductive systems are almost invariably false, in the sense that they do not correspond exactly to anything in the real world. This is true of any argument that purports to explain any social or natural phenomenon, not just models, and certainly not just the formal ones. For example, one might criticize a formal model for having the ‘unrealistic’ assumption that uncertainty over a parameter is represented by, say, the uniform distribution. But the non-formal argument that relies on the concept of uncertainty is not ‘more realistic’ because it does not make that assumption. It is exactly the opposite: without being specific about the concept, it might be impossible to evaluate as an argument. The virtue of the modeling exercise is that it can establish that the claim holds for the uniform distribution, which in turn could be used to establish whether it holds for a class of distributions or even arbitrary ones. Being non-specific and vague does not make an argument more general or ‘realistic’. If one is going to critique a model because of ‘unrealistic’ assumptions, one might as well give up any attempt to explain anything.

The proof of the model is not in its empirical consummation.⁸

If one thinks of models as arguments, then it quickly becomes clear what their role in research must be: give specific expression to a line of thinking, communicate it effectively and persuade the audience of its usefulness. It is the ability to perform these tasks well that defines a good formal model, so let us unpack them a bit.

First, formal models force one’s argument to be *specific*. That is, in order to represent abstract concepts with the mathematical formalism of a model, one is invariably forced to define

⁵ James Johnson, ‘Models-as-Fables: An Alternative to the Standard Rationale for Using Formal Models in Political Science’ (Department of Political Science, University of Rochester, NY, 2017).

⁶ Morton, *Methods and Models*, p. 102.

⁷ Clarke and Primo, *A Model Discipline*; Nancy Cartwright, ‘Models: parables v. fables’, in R. Frigg and M.C. Hunter (eds), *Beyond Mimesis and Convention: Representation in Art and Science* (Springer, Dordrecht, 2010), pp. 19–32; J. Johnson, ‘How not to criticize rational choice theory: pathologies of “common sense”’, *Philosophy of the Social Sciences*, 26 (1996), pp. 77–91.

⁸ One might wish to keep in mind the distinction between models (which is what I am discussing here) and a theory (which I will not discuss). Think of models as the building blocks of a theory: useful (or not) specifications of concepts and ideas that comprise a theoretical explanation. Theories can, and should be, subjected to empirical testing in a way that models cannot, and should not, be.

more or less precisely a specific representation of that abstract concept. For instance, we often use the rather abstract concept of ‘power’ in our research. In international relations, the term is ubiquitous, and yet it is never quite clear what it means. There are hundreds of articles about what ‘power’ could mean, and how it is to be understood in different contexts, and it is nearly impossible to evaluate research that deploys this concept (theoretically or empirically) because its meaning is protean. And yet we feel quite confident that something called ‘power’ is important in understanding international relations, not the least because policy makers seem so concerned about it.

When one writes down a formal model that wishes to use ‘power’, the abstract must become concrete. In the standard models of crisis bargaining, for example, ‘power’ is related to an actor’s expected payoff from war. It could remain undifferentiated – the higher the payoff, the more powerful an actor is – or become even more specific, relating to particular other concepts. For instance, it could be modeled as relating to the *costs* of war (the lower one’s costs, the more powerful one is), to the *probability of victory* (the more likely one is to win the war, the more powerful one is) or to the *valuation of the stakes* (the higher the stakes, the more powerful one is). Each of these specific formalizations of ‘power’ is subtly different from the others even if they all affect the expected payoff from war. For example, the *costs* only determine an actor’s own war payoff and nothing else, whereas the *probability of victory* determines the war payoffs of both actors (if one becomes more likely to prevail, the other one must be less likely to do so). *Valuation*, on the other hand, still relates to one of the actors, but determines both its war and peace payoffs.

At first blush, it appears to matter little how power is formalized: some of the most basic results in the crisis bargaining literature have been derived with each specification. However, research soon showed that the specific formalization does, in fact, matter for many important conclusions. Uncertainty about the opponent’s ‘power’ is a very common ingredient in theories of war, but once we have specified a precise formal definition of power, we are forced to adopt a precise formal definition of uncertainty as well. Thus, one could distinguish among uncertainty about the costs (independent private value), the probability of victory (interdependent value and uncorrelated types) and the undifferentiated value of war (correlated types). Fey and Ramsay have shown in a very general framework that the different sources of uncertainty have very different implications for the probability of war and the possibility of war-avoiding mutually acceptable settlements in crisis bargaining.⁹ In other words, it matters very much what specific conceptualization of the abstract notion of ‘power’ one uses for one’s argument.

Thus, contrary to the oft-repeated allegation that models are ‘too abstract’, models are in fact quite specific. They force us to give a particular expression to our line of thinking. The resulting clarity reduces the definitional burden of arguments and enables sharper communication.

Second, formalization lays bare the *structure* of the argument, which ensures its internal validity and simplifies communication. Consider the canonical crisis bargaining model used by Fearon to identify the risk–return trade-off as an important cause of war.¹⁰ Amid its myriad of simplifying premises, the model assumes that one of the actors can make a take-it-or-leave-it

⁹ Mark Fey and Kristopher W. Ramsay, ‘Uncertainty and incentives in crisis bargaining: game-free analysis of international conflict’, *American Journal of Political Science*, 55, 1 (2010), pp. 149–69.

¹⁰ James Fearon, ‘Rationalist explanations for war’, *International Organization*, 49, Summer (1995), pp. 379–414. The trade-off arises because actors are unsure how much they must concede to each other to make peace mutually preferable to war. Conceding a lot makes the other actor more likely to accept a settlement but also forces one to forego some of the benefits of peace. Conceding too little, on the other hand, risks pushing the opponent into a fight. Generally, an actor that is uncertain about what the other one’s minimal acceptable terms are would make an offer that carries some risk of rejection but that yields somewhat better terms if accepted.

demand, whose acceptance ends the game peacefully but whose rejection leads to war. Ultimatum games are quite popular in both economics and political science not because they are particularly representative of real-world situations but because modeling a bargaining situation opens a Pandora's box of additional assumptions about time horizons, discounting, inside and outside options, sequencing of moves, timing of offers and responses and many others. (These usually remain rather buried in non-formal negotiation models where they are often tacitly assumed.)

One problem is that strategic bargaining models (ones where the analyst specifies the structure of the interaction; e.g., only one actor makes demands or the two alternate) that incorporate uncertainty tend to produce infinitely many solutions, as opposed to the ultimatum game, which usually yields sharp unique results. Because of that, the results of Fearon's analysis had to be tentative because it was not known just how dependent they were on that assumption. Powell extended the model to an alternating-offers bargaining framework and showed that, unlike the common models of that type, his also yielded a unique result that mirrored exactly the risk–return trade-off identified by Fearon's ultimatum game.¹¹

This appeared to give us a stronger warrant to accept the argument until Leventoglu and Tarar showed that both the result's uniqueness and its risk–return trade-off aspect were dependent on another structural assumption: a proposer whose offer gets declined is not permitted to attack (only an actor responding to an offer could do so).¹² This premise is inconsistent with the standard assumption in international relations – that any actor can take military action whenever it chooses to do so – and relaxing it wipes out the canonical result. Their model shows that crisis bargaining can involve significant delays without escalating to war, and so the risk–return trade-off mechanism is less robust than previously thought. By varying only specific premises, scholars have been able to gain a much better understanding of how a particular argument works.

Third, formalization forces one to confront one's own demons of *unstated assumptions*. I have often heard the breezy dismissal of formal models with 'you can always concoct a model that yields any result you want', an assertion with the Schrödinger quality of being apparently both true and false at the same time.

It is true that anything can be described by some model when there are no restrictions to the premises one is permitted to make. It is also true that nothing in the modeling technology itself restricts the premises except perhaps to ensure that they are not mutually exclusive. The supposed (but famed) 'rigor' of models does not extend to the definition of the system itself or to the interpretation of its parameters and results.¹³ Models can readily produce conclusions that are absurd (these are usually easy to spot) or trivial (because the premises assume them). This happens most often when people try to reverse-engineer an argument by working from the desired conclusion to the premises needed to generate it. So, the charge is correct, at least when it describes sloppy modeling practices.

The charge, however, is almost comically wrong when leveled against models designed from first principles and with careful attention to detail, as any formal theorist worth their salt would tell you. Constructing a valid argument that does not beg the question can be surprisingly difficult; the analysis can often be startling, and the entire process quite edifying. There is much to be learned

¹¹ Robert Powell, 'Bargaining in the shadow of power', *Games and Economic Behavior*, 15 (1996), pp. 255–89.

¹² Bahar Leventoglu and Ahmer Tarar, 'Does private information lead to delay or war in crisis bargaining?' *International Studies Quarterly*, 52 (2008), pp. 533–53.

¹³ Branislav Slantchev, 'On the proper use of game-theoretic models in conflict studies', *Peace Economics, Peace Science and Public Policy*, 23, 4 (2017), pp. 1–14.

from attempts to formalize one's intuition. There is a lot of trial and error involved in getting the argument right (meaning, to have a model that is both non-trivial and solvable), and often the conclusions are not the same as the ones the analyst expected to obtain. Models are great disciplinarians, and they can teach us when our intuition has gone astray or when our 'straightforward' argument turns out to require a small army of auxiliary premises to sustain.

My favorite example of the process in print (in addition to my own trials and tribulations) comes from Schelling.¹⁴ He was dealing with the problem of surprise attack and reasoning from an intuition about an armed burglar surprised by an armed homeowner whose house he has broken into. Presumably, both prefer that the burglar just leave quietly, but the problem is that neither is sure whether the other might shoot. If he thinks that the homeowner might shoot, the burglar becomes more likely to shoot first. But the homeowner knows that and now becomes even more fearful about getting shot, so he becomes more likely to shoot first. But the burglar also knows that his own fears are compounding the homeowner's, which makes him even more trigger-happy. This escalating spiral of mutual fears causes one of them to pull the trigger, ending the interaction with an outcome both would have preferred to avoid.

Schelling's intuition is compelling, but he decided to formalize it to see how such a 'multiplier effect' could arise. He then proceeds through a series of different formalizations, none of which yields the desired result. The failure in each iteration teaches him something about the problem that he had tacitly assumed without knowing how crucial it was for the inference he was making.

Even if his resolution leaves something to be desired from a modern standpoint, the exercise was clearly useful.¹⁵ Not only did the author learn a lot about the subtleties of his intuition, but the exposition of the unsuccessful attempts to formalize it has great value for scholars who wish to build upon the insight. Knowing where the blind alleys are is crucial to progress.¹⁶

It is for these three main reasons that I almost invariably ask graduate students to formalize their arguments. Even the practice of constructing a model without solving it can be enormously beneficial, as it forces one to at least identify the basic premises and overall structure of its logic. Sometimes this is enough to expose a fatal weakness, as even rudimentary models can identify problematic lapses in reasoning.

Because models are correct by definition, specifying the premises and working systematically through their implications toward the conclusion has many benefits. We can tell whether a specific conceptual definition is consistent with a set of inferences, whether a collection of premises really does yield a claimed consequence and whether seemingly disparate arguments share a common core. We can easily agree on what a definition means, on what the precise assumptions are and on how the results are obtained. This means we can build on each other's work more effectively and transmit that knowledge with a smaller chance of miscommunication.

None of these benefits are unique to formal models. People can make, and have made, arguments that are specific, well-constructed and non-trivial without formalizing them. What I am saying is that formalization facilitates the process for the analyst and democratizes it by making the argument instantly accessible to anyone with a modicum of training. To give an analogy, it was

¹⁴ Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, Cambridge, MA, 1980), ch. 9.

¹⁵ For a modern treatment of this problem, see Sandeep Baliga and Thomas Sjöström, 'Arms races and negotiations', *Review of Economic Studies*, 71, 2 (2004), pp. 351–69.

¹⁶ In this respect, the scholarly practice of publishing only what 'works' is quite a detriment to knowledge accumulation. Much can be learned from failed attempts to formalize an argument, and I wish people would be more upfront about the fact that the published model is probably the fifth iteration attempted.

not formal musical notation that made Beethoven's Ninth possible. But it was this notation that allowed him to transmit what was in his mind efficiently, that has permitted generations of composers to build on his approach and that has enabled audiences to enjoy the results.¹⁷

Think of your model as an argument, and of yourself as a persuader.¹⁸

When to Model, or How to Get Inspired

Although every argument can be modeled, formalization can be especially appropriate for certain kinds of arguments. Since published work almost never describes the inspiration behind the models and the evolution of the arguments before they ended up in print, in this section I will give examples from my own experiences. In general, I resort to formal modeling when I have some intuition about a phenomenon of interest but find myself asking, 'How does this argument work?' Here are several instances of that question that produced formal models.

The Puzzling Case

Logic, especially when human beings are involved, is often no more than a way to go wrong with confidence.¹⁹

I teach US foreign policy to undergraduates. One of the lectures is about the Korean War, and I never felt I understood why the US and China ended up in a war over North Korea. Political scientists who have tackled this question usually frame it as 'Why did the US miss clear warning signals from China and extend the war beyond the 38th parallel?' My reading of the history, however, indicated that the signals were anything but 'clear' – the Chinese chose an Indian intermediary dismissed by its own government as biased, they failed to intervene when it would have made most sense militarily (right after Inchon) and, perhaps most critically, they appeared to have failed to prepare for war.

The last of these was crucial: nobody will take you seriously if you threaten war but do nothing to get ready for it. Assured that the Chinese were bluffing, the Americans forged ahead, only to run into a vast mass of Chinese troops who swept them back south. Unbeknown to anyone, China had entered North Korea with enough strength to shatter the UN advance and eventually stalemate the war despite reinforcements sent by the United States. If the Chinese were serious about their threats, why not show that they were preparing to fight? This would have been a clear signal, but they chose not to make it, which in turn misled the United States into invading the north. Neither side wanted to fight the other, and yet war is what they got. Why?

This case was specially puzzling considering existing crisis bargaining theory. According to all our models, an actor who is willing to fight should make that preference known by some sort of costly diplomatic or military move. This is the only way to convey resolve and hopefully convince the opponent to make war-avoiding concessions. The Chinese behavior ran contrary to all these models, as they had deliberately concealed the one move that could have conceivably had a deterrent effect on the Americans. Why would they do this?

¹⁷ Ironically, Beethoven's markings with respect to dynamics and tempo were often misunderstood, and to this day there is controversy about how his work must be performed. This is nothing that more formalism could not have fixed.

¹⁸ Deirdre N. McCloskey, *The Rhetoric of Economics* (2nd ed.) (Madison: The University of Wisconsin Press, 1998).

¹⁹ David Weber, *At All Costs* (Baen, Riverdale, 2005), p. 664.

I was reading a science fiction novel by David Weber at the time, and there was a scene in it where a very powerful military vessel on patrol detected a pirate ship in the distance. The captain's problem was that if she chased the pirates openly they would probably have the time to make their escape. So, she disguised the engine emission signature to mimic a merchant ship and lured the pirates into coming close enough for her to attack them. This was a fictional scenario, but one could think of any number of historical episodes involving such an ambush, so I wondered if the Chinese behavior had such an element to it.

The key was the possible self-denying aspect of revealing military strength: what if doing so gave the opponent an opportunity to do something that eroded your advantage? In the fictional story, the pirates would run away, thwarting the purpose of the patrol. In the Chinese case, the United States could use the information to target the troops with devastating effect because of its superior firepower and control of the skies (as it happened, Mao was still wrangling with Stalin over Soviet planes, since the Chinese air force was non-existent). Thus, if one believed that the encounter was likely to end in war, it could make perfect sense to conceal one's strength and gain the advantage of surprise.

This now presents a dilemma that the fictional captain did not have, but an actor engaged in crisis bargaining would: failure to reveal strength makes it more likely that the opponent would not offer significant concessions, which in turn would make war more likely. So, while feigning weakness could potentially be useful in war, it might not necessarily be optimal if it also made war more likely. This is the kind of problem – with incentives pointing in opposite directions – that is especially well suited for game-theoretic modeling.

I started with the canonical ultimatum crisis bargaining model and merely modified it so that after the initial demand was rejected, both actors chose how much to mobilize for the war, with their mobilizations being costly but also improving their chances of victory. Since a strong actor (one with low costs of arming) would mobilize more aggressively, it could cause the other to respond with similarly aggressive mobilization. This, in turn, would weaken the incentive to reveal one's strength through the initial demand. The analysis revealed that this was indeed the case: there were circumstances in which a strong actor would not make a risky but revealing high demand but would instead pretend to be weak by mimicking the demand of the weak type. The cost of that was a lower peace payoff but the benefit was improved war payoff. The dilemma could produce behavior contrary to the canonical models but consistent with what the Chinese had done.

The analysis also showed that the assumption of two-sided incomplete information was unnecessary: it complicated the math without changing the basic insight, which could be obtained under asymmetric information about the actor making the demand. When I sent the revised version to the journal, a referee pointed out that the fundamental result might be had with an even simpler setup. He/she was right, and the published version centers on a much simpler model where only the informed responding actor can mobilize additional resources at a fixed cost.²⁰

The model also yielded several surprises. First, there was the unexpected application: Jeff Ritter pointed out to me after a presentation that the mechanism could explain the puzzle of secret defensive alliances – they enhance one's fighting potential but deny one any deterrent advantages, the precise dilemma that could end in a feint. Second, there was the implication about the venerable explanation for war as being caused by mutual optimism. If an actor holds optimistic beliefs about the expected payoff from war, then it would be loath to make concessions, which in turn would make war more likely because it would weaken the opponent's incentive for peace. The crisis signaling literature had suggested that this optimism could be reduced by costly signals of resolve.

²⁰ The article preserves some of the original analysis as an extension to analyze endogenous tactical incentives.

One's apparent will to fight should lead the opponent to revise their estimate about war, and if they are still willing to fight, then this in turn should lead the original actor to lower their estimate as well. The feigning weakness argument showed that a strong actor could deliberately foster false optimism in the opponent, which short-circuits the inferences: this actor could not use the opponent's willingness to fight as evidence to lower their own estimates about war since that willingness is based on the wrong inference induced by the feint. This would strengthen mutual optimism and make war more likely than the signaling literature would suggest.

Thus, my inability to offer a coherent explanation of a historical case to my undergraduates, coupled with the coincidental reading of a science fiction novel, led to novel insights about crisis bargaining.

I find history both fascinating and often puzzling. It offers a tremendous menu of opportunities to not understand something and is a fertile source of inspiration for research. I teach an entire graduate course on the history of international relations that is designed to puzzle students and generate ideas for study. It is important to realize that historical episodes (and empirical patterns) can only be puzzling considering existing explanations: one must know enough to understand that one does not understand something.

Being puzzled productively requires quite a bit of preparation, which is why it is often difficult for scholars tackling unfamiliar areas of research. Sometimes one's puzzlement is due to ignorance of existing work and is easily resolved by a literature review. This is why I am not a big fan of advice that tells scholars to avoid the literature review until after they have clarified their ideas. In fact, searching the literature with a particular idea in mind can be especially productive and efficient.

The inconsistent assumption

Varian contends that academic journals 'really aren't a very good source of original ideas'.²¹ I think what he means by that is that it is hard to get inspired by academic work to create something 'original'. I am using scare quotes because I never quite understood the emphasis on originality in the profession – sometimes there is a very good reason an idea did not appear in print before you had it, and it is not because you are a genius. Be that as it may, here is an example motivated by published research.

The study of crisis bargaining and escalation has relied on stylized models neatly summarized and analyzed by Fearon.²² He distinguishes between behaviors that involve sunk costs (paid irrespective of the outcome) and those that involve audience costs (paid only if an actor backs down after making a threat). As an example of the former, Fearon gives 'building arms or mobilizing troops'. He is very careful to note that these actions 'may affect the state's expected value for fighting' (p. 70) and that it would be 'more realistic to have the probability that the defender wins in a conflict depend on m [the level of mobilization]' (p. 72). He opts not to do this in order to keep sunk costs and tying hands analytically distinct.

I wondered whether this distinction was distorting. It was very difficult to conceive of a military move that did not alter the distribution of power between the actors. This meant that any such move would alter not only the incentives of the actor making it but also those of its opponent:

²¹ Hal R. Varian, 'How to build an economic model in your spare time', *American Economist*, 41, 2 (1997), pp. 3–10, p. 3.

²² James D. Fearon, 'Signaling foreign policy interests: tying hands versus sinking costs', *Journal of Conflict Resolution*, 41, 1 (1997), pp. 68–90.

tying one's hands might simultaneously untie the opponent's, affecting the overall probability of war in the crisis in ways not anticipated by the original models.

To check this intuition, I modified that standard escalation game so that both actors could choose their mobilization levels during the crisis, with the probability of victory dependent on these allocations.²³ Sure enough, there were important differences in the inferences. For example, in contrast to Fearon's results, the military threat model (MTM) shows that bluffing is possible. It also undermined the long-established result in the crisis bargaining literature that militarily stronger actors are likely to obtain better deals but must run a higher risk of war in order to do so.²⁴ In the MTM, these actors still get better deals, but their risk of war might be *lower* because they can compel the other side by more aggressive mobilization (the untying hands effect). The MTM also showcased the importance of considering the costs of maintaining peace through mutual deterrence with high military allocations, which led me to another paper.

The tacit assumption

An essential, but often overlooked, assumption in the canonical crisis bargaining model is that peace is costless.²⁵ The high mobilizations without war in the MTM alerted me that this assumption might be problematic. I should have known this from Powell's earlier contribution that showed how the possibility of an armed peace depended on the long-term costs of deterrence, but since the point had been peripheral to the goals of his article, I had missed both it and its implications for crisis bargaining.²⁶

Clearly, if peace were to be costlier than war, then the bargaining puzzle would fall apart: actors would fight because war was preferable to any negotiated outcome for at least one of them. This would shift the focus from the now trivial problem of war under these circumstances to the unexamined problem of how actors would create these circumstances in the first place. In other words, if they knew that making peace too costly would lock them into war, would they pursue strategies that do so anyway?²⁷

I took the direct approach: since arming is costly irrespective of the outcome, I decided to look at how actors paid for military power. Since I had been reading a lot about the emergence of centralized government in early modern Europe, my head was full of examples of kings who could not tax very effectively but who borrowed a lot and sometimes failed to pay back their debts. War finance through borrowing instead of taxation also seemed appropriate because funds would be instantly accessible (unlike tax proceeds), and this appeared important considering the fact that the vast majority of interstate wars only last for a few months.²⁸

²³ Branislav L. Slantchev, 'Military coercion in interstate crises', *American Political Science Review*, 99, 4 (2005), pp. 533–74.

²⁴ Jeffrey S. Banks, 'Equilibrium behavior in crisis bargaining games', *American Journal of Political Science*, 34, 3 (1990), pp. 599–614.

²⁵ Strictly speaking, all that is necessary for the canonical results is for the status quo to be less costly than war. This is usually modeled as the negotiation outcome being costless.

²⁶ Robert Powell, 'Guns, butter, and anarchy', *American Political Science Review*, 87, 1 (1993), pp. 115–32.

²⁷ This line of reasoning was independently pursued by Andrew Coe, 'Costly Peace: A New Rationalist Explanation for War', Manuscript, Department of Political Science, Vanderbilt University, 2011, who identified three other causes of peace being potentially costlier than war.

²⁸ Branislav Slantchev, 'How initiators end their wars: the duration of warfare and the terms of peace', *American Journal of Political Science*, 48, 4 (2004), pp. 813–29.

As usual, I started with the canonical model, allowed both actors to determine the distribution of power through their military mobilizations and merely allowed one of the actors to borrow to increase the resource base, thereby permitting larger military allocations. I assumed that the actor was committed to repaying the debt if the bargaining ended peacefully or if the war ended in victory, but that the debt was repudiated if the war ended in defeat.²⁹ The analysis then revealed conditions under which the actor would incur a debt so high that the other would not be willing to concede enough to enable its peaceful repayment, and as a result the interaction would end in war under complete information. Further analysis revealed the importance of the actor's efficiency in converting financial resources into military capability, a topic never explored formally before.

The review process beefed up the article substantially, since the referees wanted me to allow both sides to borrow and wanted me to account for interest on the debt that was somehow related to the risk of default. These analyses took several months to complete and showed that the fundamental insight was not dependent on these simplifying assumptions in the original. And so the article conveyed a new, and different, argument about the causes of war.³⁰

The unsatisfying argument

Among the most fertile sources of inspiration are attempts to explain your arguments to others. Teaching students, discussing with colleagues and sometimes even just chatting with friends and family have all, at one point or another, stopped me dead in my tracks in the sudden realization that my argument does not quite work, either because there seem to be missing steps or because it is making potentially distorting assumptions that themselves need to be explained. Being puzzled on one's own by reading is much harder than being stumped by someone's question. There is probably no limit to the inanity of ideas I can come up if I work in isolation, and the healthy skepticism of others is a crucial corrective. That is why I advise students to talk about their ideas as much as they can, to anyone who will listen. Instead of becoming defensive about criticism, look at it as an opportunity to develop a better argument.

My first example is a model that came about from an offhand comment during a lunch break while I was still in graduate school. I was working through the literature on audience costs for a course assignment and was chatting with a faculty member (I cannot recall whom) in the lounge while waiting for the microwave to warm up my lunch. He had asked me what I was working on, and I was explaining the idea behind audience costs – that leaders who escalate a crisis are punished if they back down – when he interrupted me by asking: 'why would they do that?' As I was giving the usual 'national honor and prestige' answer, I began to realize that it involved an uncomfortable amount of hand-waving and that a good explanation would require these costs to arise endogenously in the model. In other words, there should be a reason for the audiences to be willing to impose costs on leaders for backing down. At this point, the microwave pinged, and the conversation shifted to something else. But the question bugged me.

The problem was that it was not actually at all clear to me why someone would punish a leader for avoiding war, especially if bluffing was an optimal strategy. I did not do much with this because I had to finish my dissertation, but a few years later I was discussing the importance of leaders with Hein Goemans and suddenly recalled the puzzle. We had a back-and-forth about this, and I searched the (very small) formal literature on the topic only to find a couple of scholars

²⁹ Technically, all that is needed for the results is that default is more likely after defeat. This makes debt service less costly in expectation if a war is fought than if peace obtains, which is enough to trigger war under some conditions.

³⁰ Branislav L. Slantchev, 'Borrowed power: debt finance and the resort to arms', *American Political Science Review*, 106, 4 (2012), pp. 787–809.

asking the same question. I was not satisfied with the answers, so set out to model the problem myself.

My model strips away all detail – like the presence of a foreign actor – that did not seem pertinent to the analysis of the interaction between a leader, a policy being implemented and the domestic audience.³¹ I found a model developed by Dur to deal with the persistence of bad policies, and adapted it for my purpose, reasoning that a legitimate reason for punishing a leader (and thus imposing audience costs) would be the audience realizing that the policy implemented is bad and so preferring a more competent leader.³² The nuance of the argument was that the imposition of costs had to happen with positive probability during the interaction, rather than being assumed as a hypothetical threat with the leader then taking action to avoid it. (If the imposition of costs remained a zero-probability event in the model, then the purported explanation of audience costs would amount to an assumption.)

As I developed my intuition, I realized that since the argument hinged on the asymmetry of information about the policy quality between the leader and the audience, the model might be useful in exploring other potential sources of relevant information, such as a political opposition and a possibly biased media.³³ Consequently, the model expanded to include these actors (along with the entailing auxiliary structural premises), and generated some surprising insights. Among them was the result that in the absence of a robust and unbiased free press, democracies were no more likely to generate audience costs for their leaders than autocracies. This was in contradiction to Fearon's working hypothesis in the original paper that claimed the opposite. Moreover, the argument helped explain why mixed regimes could be especially sensitive to policy failures.

My second example is a model that came about from a discussion with Christina Schneider, who had been researching decision-making in international organizations (IOs). We were dog-walking and sharing what we knew about the literature on the topic when she mentioned the problem of compliance with IO decisions. When would actors abide by collective decisions without an exogenous system of enforcement? This quickly led to another question: how would actors agree on such collective decisions? Since most organizations involve voting, the answer seemed easy. But then we realized that the insights about voting come from models that assume that the outcome is enforceable and voters who disagree do not get to work to overturn decisions. Moreover, almost all such models assume that voting is sincere, which is not an issue in the setting they were developed in (secret votes) but that could be quite problematic where votes are public, as they are in many IOs. Indeed, many empirical studies implicitly rely on public votes being sincere when they use them to measure preference similarity of member states.

Thus, we ended up with a question: what makes voting in IOs meaningful in the sense that actors are likely to reveal their preferences with their vote and abide by the collective decision even if they disagree with it? We began formalizing the problem that very evening, and several

³¹ Branislav L. Slantchev, 'Politicians, the media, and domestic audience costs', *International Studies Quarterly*, 50, 2 (2006), pp. 445–77.

³² Robert A.J. Dur, 'Why do policy makers stick to inefficient decisions?' *Public Choice*, 107, 3/4 (2001), pp. 221–34.

³³ Kenneth A. Schultz, 'Domestic opposition and signaling in international crises', *American Political Science Review*, 92, 4 (1998), pp. 829–44; Matthew Baum, 'Going private: public opinion, presidential rhetoric, and the domestic politics of audience costs in U.S. foreign policy crises', *Journal of Conflict Resolution*, 48 (2004), pp. 603–31.

months later had an answer.³⁴ The modeling exercise here also led us to some more fundamental issues such as conceptualizing of international cooperation not merely in terms of the free-rider problem, as the widespread reliance on repeated games with Prisoner's Dilemma preferences does, but also in terms of a conflict of interest between groups of like-minded actors with resources to pursue divergent policies. The introduction of this competitive element in the cooperation problem brought the original puzzle into sharper focus but also opened up a host of related issues for research.

In both examples, the impetus behind the model arose from the feeling that the existing arguments were not quite right because they relied on a premise that was itself in need of explanation. This premise might have appeared for modeling convenience or because of the adoption of a model developed for an apparently related but in fact quite different context. In neither case was the empirical validity of the assumption relevant.

Some bad ideas

As I indicated above, thinking about published research could give one insight that begs to be formalized. For this to work, however, one needs to know enough about the substantive phenomenon being studied to understand which premises in the existing argument might be distorting the conclusion, and thus warrant change. Without this 'inductive' step, modifying existing models for the sake of 'relaxing assumptions' and 'making them more general' could turn out to be of interest only to a handful of modelers, or could prove a pointless exercise altogether. Do not formalize for the sake of formalization.

One unfortunate consequence of the misguided 'scientism' in the social sciences is the insistence that hypotheses be derived from a formal model. Somehow, this is supposed to imbue them with rigor and validity, never mind the fact that almost all such models are concocted after the fact and are absurdly trivial. I have argued at length about the supposed 'rigor' of formal models elsewhere,³⁵ so here I will limit myself to the following injunction: do not formalize just to give a formal gloss of your hypothesis (or to pretend that you are getting some precise point estimates).

As I will argue below, a good first step in building a model is to use an existing one and modify it as little as possible to adapt it to the problem under consideration. Thus, someone interested in crisis bargaining might start by looking at models of strikes or pre-trial negotiations: in all these settings the actors possess a power to hurt each other in the absence of an agreement, and in some there is also uncertainty about who is going to prevail in a costly 'fight' if negotiations fail.

This sort of importation must be done very carefully, though, because there might still be very important differences between the contexts, and if one borrows the idea without accounting for them, the result could be worse than useless: it might in fact be harmful to subsequent research. Political scientists are especially prone to borrowing models from economists without due consideration of the different context or the fact that the economists themselves very often neglect crucial political considerations in their own models. It is very easy to end up with a very ornamental model and pages of equations that amount to no insight whatsoever. Sometimes all you need is a small change in a premise to adapt the model, but sometimes you might as well build your model from scratch. Do not attempt to shoehorn your intuition into an existing framework.

³⁴ Christina J. Schneider and Branislav L. Slantchev, 'Abiding by the vote: between-groups conflict in international collective action', *International Organization*, 67, 4 (2013), pp. 759–96.

³⁵ Slantchev, 'On the proper use of game-theoretic models in conflict studies', *Peace Economics, Peace Science and Public Policy*, 23, 4 (2017), pp. 1–14.

Have Puzzle, Will Model

Most of my research time is spent thinking about puzzles and trying to make persuasive arguments that resolve them. When I decide to formalize something, I also tend to spend considerable time on what the model should look like. Understand that this is almost never the unidirectional process suggested by how the idea appears in print: idea → model → solution → interpretation. Instead, one should expect to make several attempts to formalize the idea, sometimes along with full or partial solutions to the models, and sometimes one might even have to modify the original puzzle in light of what the analysis uncovers. The process looks like this: simplify → model → solve → realize either model or question were not quite right → change appropriate premise → iterate. This goes on until one is satisfied about the match between the question and the model.³⁶ It is worth keeping a record of the failed attempts to avoid having to repeat going down blind alleys and to assist with the write up and responses to referees (who often suggest things one has tried already).

When building the model, be aware that there is no guide that can tell you whether the model is going to be useful. Do not strive for “realism”. Keep it simple, and realize that all models, even the most ‘realistic’, are false. Utterly and irredeemably so. Strive instead for minimalism and elegance. Remember, you are going to be making an argument, so it pays to be clear, precise and concise, and to have as few moving parts as possible to convey your intuition. Resist the temptation to show off modeling chops with complex math. Any real mathematician is going to laugh at such a folly anyway. Unnecessary bells and whistles do not make the model ‘more realistic’; instead, they make it harder to solve and even harder to follow. Proliferating parameters and premises decreases the warrant to believe the robustness of the result: how can one be sure that it is not entirely dependent on one of these myriad assumptions or some unlikely combination of them?

Scholars are often tempted to make things unnecessarily complex (which is why impenetrable jargon is always *de rigueur* in seminars), mistaking incomprehensibility for profundity. The same is true for formal models: one can easily build a formidable-looking array of Greek letters and numbers without realizing that no flower of an idea survives in the forest of equations. Sometimes the authors themselves do not understand how these models work – having plowed to a solution in a mechanical fashion – let alone the audience, who might be duly impressed by the math fireworks but leave scratching their heads and quickly forgetting about the idea. One might hope that diligent scholars would pore over one’s brilliant but obfuscated work to tease out the intellectual gems concealed amid the baroque ornamentation, but most of us are neither Plato nor Aristotle, so what are the chances of that? Most likely, the work will perish in the scholarly wilderness for lack of citation sustenance. If you want your ideas to make a difference, you need the audience to understand them, which means it is your responsibility to make them as clear as possible. Models should be as complex as they have to be, but no more.³⁷

How can this be done? For starters, relating the model to something known can be helpful. This is why I talked about beginning with an existing model, especially one that the audience is

³⁶ This reality also makes a mockery of the idealized notion that one constructs a deductive model in some pristine analytical void and then ‘tests’ its conclusions. The process I am describing is a mix of induction and deduction, all informed by the researcher’s knowledge of empirical patterns, other models, and cases.

³⁷ For a discussion of what model complexity is and isn’t, see Catherine C. Langlois, ‘Are complex game models empirically relevant?’ *Conflict Management and Peace Science*, 35, 4 (2017), pp. 1–15. She also critiques Allan and Dupont’s assertion of a ‘tendency toward theoretical elegance to the detriment of empirical correctness’ that supposedly plagues formal models. I leave it to the reader to decide what to make of an assertion that uses ‘empirical correctness’ as a criterion to evaluate models.

likely to be familiar with. Building on previous work is not merely good scholarly practice (acknowledging the contributions of others should be the *sine qua non* of research), but also a valuable aid in fleshing out the argument – how similar is it to others, and what makes it different. It also helps communicate the ideas efficiently and effectively – the audience is more likely to grasp something that is not too far from what it already knows, and since it can then evaluate your argument better, it is more likely to be persuaded of its merits.

If an existing model is not readily adaptable, then you must build one from scratch. Begin by specifying who the actors are, what they want, what they think they know, what their constraints are and what they can do. There will be many tough choices to make here, and it is not at all obvious initially which are better. Should you limit yourself to two actors? If the interaction is dynamic, should it be just two periods or some sort of infinite-horizon game? If there is incomplete information, should you use two types or a continuum? If the latter, should you use an arbitrary distribution or something convenient analytically, like the uniform or normal distributions? Is the action space discrete or continuous? How many opportunities to act will the actors have, and in what order will they move?

The problem here is that when you are building a model from scratch, there is always the temptation to make the argument more general and the model more ‘realistic’. You might not know what assumptions are going to make it intractable and what assumptions might make it trivial. So, there will be trial and error here. Build a model and try solving it. This will give you some intuition about its moving parts and how they interact. It will also give you some ideas about how to improve the model.

I cannot stress enough how important the process of building a satisfactory model is, how messy and iterative it can be, and how long it might take. Published papers make it look like the model sprang from the mind of the author fully formed and perfectly adapted to the task, like some sort of mathematical Athena from the head of Zeus. In my experience, nothing could be further from the truth. Coming up with a model that is an adequate representation of your argument is hard work, full of trial and error, over a very uncertain timeline. It has on occasion taken me months of trying different specifications and sets of premises to arrive at a model that is a reasonable approximation of my intuition while simultaneously being solvable.

This is where minimalism and elegance become crucial. By minimalism I mean trying to come up with the simplest non-trivial model you can think of that formalizes your intuition. Your goal is to expose the structure of your argument as cleanly as possible, so any unnecessary parameter or detail should be mercilessly pruned.

I have taken to writing a paragraph in my papers, right before the model specification section, where I enumerate the essential features the model should have in order to represent the puzzle the argument is going to address. This paragraph comes very naturally after the literature review, which has situated the puzzle in the relevant literature and shown why current research has not answered the question being posed. This review identifies gaps in existing arguments and points to premises that need to be incorporated into or omitted from the new argument. This paragraph also ‘sells’ the model to the audience by justifying its premises explicitly. Anything that cannot be justified in this paragraph should be removed from the model.

Elegance is an elusive concept. It is something that one recognizes when one sees it, but that cannot be defined very precisely. Minimalism certainly helps, but there is more to it. Does the model seem ‘natural’ for the question being posed (this is where that paragraph also helps)? Or does it incorporate some odd premises that artificially constrain the actors in their choices? All assumptions are false, but some are beyond silly and are likely to make your argument

unacceptable despite its deductive rigor.³⁸ Always remember that the strength of your argument will rest not on the idea that your agents are optimizing but on what you have them optimizing over: their preferences and constraints.³⁹ Do the payoffs reflect reasonable preferences or do they appear contrived and complicated? Anything with more than a few parameters or with very specific functional forms begins to look suspicious to me. Is the notation intuitive or clunky? There is no standard notation in game theory, so here it is best to adopt the notation used in well-cited articles or textbooks such as *Game Theory* by Fudenberg and Tirole.⁴⁰ Here it is best to emulate the specification of prominent and well-cited models that you admire.

The build \rightarrow solve \rightarrow build again \rightarrow solve again \rightarrow build again \rightarrow etc sequence suggests that one could benefit from being smart about the analysis. Instead of trying the most general case, go through a few simpler variants first. Use numerical examples to get a handle on what might be possible in the model, and some intuition about the parameter space where it can happen. Simulations and graphs are an excellent way of exploring the model before you begin solving it analytically. Plot the payoffs and vary the strategies of the other players to see what form the best responses might take. You might notice a pattern. For example, some relationship between payoffs from two choices seems to persist no matter what values you assign to some parameter. Try to prove analytically that the optimal choice is independent of that parameter. You might see abrupt changes in the optimal solution. Try to prove that it changes form at some threshold value. Once you derive the best responses, program them and then explore the comparative statics. You should use whatever programs you are comfortable with.⁴¹

When you start discovering analytical results, it is time to write them down in your draft paper. This is where the first lemmas and propositions will make an appearance. There might not be a lot connecting them yet, but the skeleton of the argument is being constructed as you learn from your model. I also like to typeset them in LaTeX immediately because the math looks beautiful, the proofs are easier to read and the text is readily useful. (It is also easier to make global notational changes.) I also write explanations of the intuition behind these results as if I am talking to an audience unfamiliar with the model. These often make it in some form into the final draft and are particularly useful to keep the argument running in my head.⁴² Do not wait until the analysis is complete to write – write as you go along. You will end up with multiple drafts of various models and partial solutions as the record of your research endeavors, and you will have the basic draft of the formal exposition ready when you complete your analysis.

³⁸ An example from economics would be the notion that unemployment represents the workers' free choice of leisure without a job over working at previous pay, as the real business cycle theory would have one believe. See Robert Lucas and Leonard A. Rapping, 'Real wages, employment, and inflation', *Journal of Political Economy*, 77, 5 (1969), pp. 721–54.

³⁹ Robert Solow, 'The state of macroeconomics', *Journal of Economic Perspectives*, 22, 1 (2008), pp. 243–9.

⁴⁰ Drew Fudenberg and Jean Tirole, *Game Theory* (The MIT Press, Cambridge, 1991).

⁴¹ I use *Mathematica* for the analytical derivations (I prefer doing them by hand and checking the result), and I use *Gauss* for simulations. The only reason for these choices is that these are the programs I learned in graduate school. The numerical exploration approach is advocated in Catherine Langlois and Jean-Pierre Langlois, 'From numerical considerations to theoretical solutions: rational design of a debtor-creditor agreement', *Peace Economics, Peace Science, and Public Policy*, 22, 4 (2016), pp. 403–12.

⁴² Keeping notes for yourself is not as strange as it sounds. If you stop working on a model for a few weeks, by the time you come back you will have forgotten all the details. At this point, I usually have to re-derive everything from scratch, so it really helps to keep extensive notes and not rely on memory for anything. This is a practice I inherited from my days as a computer programmer, when staring at my own code a couple of months after I wrote it taught me to document it as if I was addressing a partially lobotomized monkey.

Throughout all of this, you must be ready to be taught by the model. Or, rather, you must understand the intuition behind the results you are getting, and you must be willing to either jettison the model if it does not represent your argument properly, or accept that your original intuition might have been incorrect or incomplete. In the end, learning from your model is the largest payoff from formalizing your argument.

Exposition: Learning from Your Model and Telling Others

Now that the analysis is done and you have your main results, what next? If you followed the advice to write the intuition behind each step in your argument, you have an excellent handle on how it works. Your goal now is to convey this to others and to persuade them.

Your final draft will not track either your research progress or the complexity of the argument itself. Instead, the paper should focus on how your results answer the question you posed (which may have been restated several times as your thinking has evolved while solving various models), and it should convey that connection efficiently and intuitively. The paper must lead the audience to your conclusions, not rely on it making its own inferences. This means exposing all necessary steps in the argument without getting bogged down in technical detail. It might be painful to relegate 50 pages of hard-won mathematical results to an appendix very few will ever read, but this is what you must do. The body of the paper should include just enough mathematical detail to carry the argument in plain prose.

If you cannot explain the behavior of your agents without reference to equations, you have a problem. You are telling a story, which means that your agents should have intuitive (given incentives and constraints) behaviors. Nobody will care about uninterpreted statements that refer to impenetrable mathematical conditions, no matter how correct they are. Nothing should remain ‘counterintuitive’ after you have presented your argument. Presenting the paper to colleagues is a very effective way of fleshing out the rough spots in the write up. What is obvious to you might be totally opaque to others. What you think is trivial might be crucially important to others. It is very difficult to put yourself in the position of an audience that has not spent any time on your research, so do not do it. Instead, go with a real audience for that. I never miss an opportunity to present ideas and have never turned down an invitation to do so, especially if the audience is not academic.

Word count limits prevent me from spending more time on advice about crafting the paper. Fortunately, there are excellent books about how to write elegantly and concisely. Some of them are even specific to formal work.⁴³ Follow their advice. Read widely and emulate writers you admire. Like any skill, writing is made perfect with practice. And do not default to the dry, pedantic and, frankly, boring tone characteristic of academic papers.

Remember, your model is an argument, and persuasion hinges on how it is presented, on rhetoric. Strive for readable prose. Do not be afraid to be slightly imprecise when the alternative is a detour into technical detail. Use historical cases to illustrate your points (but do not pretend that they are some sort of ‘tests’ of your results). It is fine to be entertaining. It is you who must

⁴³ I have found two very useful for general prose: Helen Sword, *Stylish Academic Writing* (Harvard University Press, Cambridge, 2012); Francis-Noël Thomas and Mark Turner, *Clear and Simple as the Truth: Writing Classic Prose* (2nd ed.) (Princeton University Press, Princeton, 2011). For formal writing, one is indispensable: William Thomson, *A Guide for the Young Economist: Writing and Speaking Effectively about Economics* (The MIT Press, Cambridge, 2001).

provide the justification and the interpretation of the model. It is you who must explain the argument. It is you who must hold your audience's attention and persuade it. Leaving any of these steps to others is a guarantee that your modeling efforts will be for naught. And we would not want that, now, would we?