European Journal of Political Research **58**: 1001–1018, 2019
doi: 10.1111/1475-6765.12319

On prediction in political science

KEITH DOWDING & CHARLES MILLER
Australian National University, Australia

Abstract. This article discusses recent moves in political science that emphasise predicting future events rather than theoretically explaining past ones or understanding empirical generalisations. Two types of prediction are defined: pragmatic, and scientific. The main aim of political science is explanation, which requires scientific prediction. Scientific prediction does not necessarily entail pragmatic prediction nor does it necessarily refer to the future, though both are desiderata for political science. Pragmatic prediction is not necessarily explanatory, and emphasising pragmatic prediction will lead to disappointment, as it will not always help in understanding how to intervene to change future outcomes, and policy makers are likely to be disappointed by its time-scale.

Keywords: explanation; international relations; philosophy of social science; political science; prediction

Introduction

Prediction has always been central to scientific endeavour. In common parlance, 'prediction' means foretelling something about the future. In the philosophy of science, 'prediction' usually means some implication logically drawn from a theory with no necessary reference to the future. The logical implication might be about something already believed to be the case or something previously unsuspected (see, e.g., Popper 1972: 33; Carnap 1966: 45). While Hempel (1942) believed prediction to be about future events, he still saw a prediction as a logical derivation from a theory and initial conditions. Where the event had already occurred the theory did not predict but rather explained. Predictions in natural science also do not necessarily refer to events. For example, the Higgs boson in physics was predicted by the standard model in the 1960s, but empirically measured only in 2012. Clearly, it existed prior to both its prediction and the confirmation of its existence (Dawid 2013).

There is similar ambiguity in political science over the term 'prediction'. An increasingly important strand in the literature seeks to improve the ability of political science to forecast future events, with some arguing that this ability may constitute a new 'gold standard' for the empirical evaluation of political science research. Yet political scientists also often use the term 'prediction' to refer to the implications of a theory with respect to events that occurred in the (often distant) past. This apparently puzzling state of affairs can be resolved by distinguishing two different types of prediction that political scientists in fact engage upon.

In this article we outline the difference between the two types – 'pragmatic' and 'scientific' prediction – and explain why it is important to separate them for political science. *Pragmatic prediction* is the probabilistic forecasting of types of future events (such as elections, *coups d'état* or civil wars). Assessing how good those predictions are is based on how closely the assigned probabilities match what actually happens. *Scientific prediction* we define as the logical implication of a theoretical model, such as 'the greater the number of veto players in a system the greater the stability of public policy'. Scientific predictions are characteristically produced by formal models – the predictions strictly are entailments of the model. But typically in political science, models are less formal; often they are produced as an extension of other data (induction) and are accompanied by a narrative that is meant to explain the mechanism by which the generalisation is derived. Scientific predictions constitute a necessary component of an explanation in the sense that, if a set of propositions X are meant to explain a phenomenon Y, then those propositions are meant to explain Y to the same degree in all relevantly similar circumstances (Dowding 2016: Chapter 3). Pragmatic predictions forecast future events and need not be explanatory. We argue that while good pragmatic predictions are important and useful, good scientific prediction should remain the principal desideratum for positive political science.[1]

Scientific and pragmatic prediction

This brief definition of scientific prediction requires a further account of what we mean by a theoretical model. Such models come in two types: deductive models, from which predictions are logically derived; and statistical models. Testing the former involves seeing how far outcomes conform to the predictions,[2] with a further demand that the logic of the model describes, at some level of granularity, the mechanism that causally explains the predictions. Whether formal models can be tested in this manner is open to some debate (Clarke & Primo 2012; Dowding 2016: Chapter 5). This abductive process constitutes formally fallacious reasoning, because what is in the conclusion goes beyond what is contained in the premises – but that is just to say abduction is not deduction. In fact, abduction (sometimes called 'inference to the best explanation') is central to scientific methodology (Boyd 1984; Harré 1986; Lipton 1991; Williamson 2007). Abduction involves taking a known outcome and inferring an explanation of that outcome. Most of what we think we know comes through abduction.

Statistical models are created as data are collected where the independent variables are given some functional form that predicts the dependent variable. The independent variables are used to explain the dependent variable by means of an explanatory narrative which, generally speaking, is again supposed to represent the causal mechanism, though sometimes authors note an empirical generalisation that is in need of an explanation. With inductive models, the generalisation produced is projected on to further cases. Here the prediction can be seen as the projection of the curve on to new cases within the same parameters, or the further projection of the curve over new parameters. How far such correlations can be assumed to represent causation is also controversial (Druckman et al. 2011; Morton & Williams 2010). Again, the problem with such abduction, or what in statistics is usually called the 'specification problem', is that we can often come up with several narrative explanations for any given correlation. The specification problem is also the major issue in qualitative case study explanations (Dowding 2016).

There are various critiques of both forms of this theoretical model in political science, but in this article we are concerned only with the one we label 'the pragmatic critique'. According to the pragmatic critique, if the mark of good political science is good predictions, then, at least until recently, political science by its own 'scientific' standards has been very poor indeed; it is very bad at producing accurate pragmatic predictions (Tetlock 2005; Ward et al. 2010). This may reflect a deficiency in our current theoretical models or even, more radically, seem to suggest that we should junk theoretical models altogether and base future political science work on whichever algorithms we find to have the best pragmatic predictive power, irrespective of whether we fully understand how these algorithms arrive at those forecasts. We argue that this would be misguided: a theoretical model with high scientific predictive power can be good and useful with respect to many things, even if its pragmatic predictive power is low; while a model or algorithm with high pragmatic predictive power but little scientific predictive power may not prove very useful for many of the things we expect political science to be able to do.

Scientific predictions might not, initially, even be practically testable, but the worth of a theory is usually judged by how well such predictions can be tested, and its veracity by how well those predictions fare. The strength of a scientific prediction, its empirical content, is derived from what it excludes (Popper 1972, 1983). A theory that says 'anything can happen' excludes nothing, so has low empirical content. A point prediction of a token event excludes much, so has high empirical content. Thus, the worth of scientific predictions can be judged by their empirical content and their predictive success. Some scientific predictions might be low on empirical content and therefore deemed trivial. (We return to this point below.)

Scientific predictions might not be related to future events at all. The entailments of a theory refer to claims that might be known to the theorist at the time the theory was developed or to ones that are unknown. Models which only provide the former are usually termed 'accommodationist' and those which produce the latter 'predictionist'. Scientific prediction in our terms includes accommodation as well as the prediction of unique or unknown discoveries. While it is generally assumed in political science that novel prediction in this sense is epistemically superior to accommodation, this is not universally accepted in the philosophy of science (Hitchcock & Sober 2004; White 2003). While we address this debate below, here we note that in spite of the discipline's general belief in the superiority of prediction to accommodation, much inductive political science is accommodationist. Indeed, a large part of social science is about trying to explain past events, and in that sense cannot be strictly predictionist.

There is an obvious relationship between scientific predictions and pragmatic predictions. After all, scientific predictions are of a conditional nature: if condition C holds, then we expect Y; or if condition C holds, then we expect with some probability p that Y holds. Moreover, the difference between scientific and pragmatic prediction is not necessarily a divide between different political scientists or even different models. Most political scientists who engage in pragmatic prediction also engage in scientific prediction. A theoretical model may be used for either scientific prediction or pragmatic prediction, though of course it may be much better for one than for the other. Neither of these facts, however, imply that pragmatic and scientific prediction are the same thing – the same political scientist can test the same theory using both observational and experimental data, but this does not mean that observational and experimental data are identical. The difference between pragmatic prediction and scientific prediction lies along two axes – reference to the future and explanatory status (Table 1). Pragmatic predictions are about the future, but are not explanatory; scientific predictions are explanatory, but not necessarily about the future.

Table 1. Scientific and pragmatic prediction

	Scientific prediction	Pragmatic prediction
Future-oriented	No	Yes
Explanatory	Yes	No

Explanatory here means that we have an idea of how the outcome y could have been different in some counterfactual situation where condition C would not have obtained, and we can infer what value y would have taken on had a different condition C^* obtained. That this is close to the interventionist definition of causality is no accident,[3] since causality and explanation are closely linked. We do not, however, believe that only causal arguments are explanatory, but we do not pursue this argument further here (see Dowding 2016: Chapter 3). We can note that our table says that scientific generalisations are not future-oriented – of course, we can take a scientific prediction and together with data make a prediction about a future event – that is, our theories with their scientific predictions together with data can generate pragmatic predictions. Analogously, we can unpack successful pragmatic predictions to produce viable explanations. Successful pragmatic prediction is not necessarily explanatory, though, (miracles aside) it ought to be able to be underpinned by scientific prediction. We can reverse engineer the theory from successful predictions to create scientific prediction. But our stipulative distinction between the two types of prediction is designed to elucidate the different roles that predictions play in political science.

Most generalisations in social science are empirical ones. An empirical generalisation is one that tends to hold but is not a strict law or law-like generalisation (Dowding 2016: Chapter 3; Popper 1972). It is not invariant. Empirical generalisations have to be discovered and then the causal mechanism inferred, but usually in the social sciences such generalisations are far from invariant. We often generate empirical generalisations through induction, without necessarily formally modelling a causal mechanism that leads to the empirical generalisation.[4] While scientific prediction can lead to pragmatic prediction, it is possible to pragmatically predict without scientific prediction. As we shall see, big data lends itself to pragmatic prediction without necessarily allowing us to generate scientific predictions. Theoretically specifying mechanisms leads to scientific predictions and purported explanations, but without further predictions beyond known data we might be sceptical that such theories are really explanatory. This is one of the problems of accommodation.

The case for pragmatic prediction

Early work on pragmatic prediction revealed that political science was notably bad at it. The best-known critique is Philip Tetlock's *Expert Political Judgment* (Tetlock 2005), which solicited probabilistic predictions from an anonymous sample of experts about selected political events over a number of years. Tetlock assessed the forecasts on the basis of Brier scores, which measure the absolute distance between the predicted probability of an event and whether it happened. From this survey and metric, Tetlock concludes that most experts fare little better than chance in forecasting future events and some do even worse. A similar test of one of the best-known existing statistical models of civil war onset (Fearon & Laitin 2003) found it to have little forecasting power (Ward et al. 2010).

This early research on pragmatic prediction proved popular with critics of the social sciences. However, the failure of existing social science by the mid-2000s to successfully pragmatically predict spurred a rich vein of literature attempting to do better. Researchers who originally identified the shortcomings in existing scholarship were in the vanguard. They started from the premise that, although social-scientific pragmatic prediction had not yielded impressive results, that did not mean that accurate pragmatic prediction was in principle unachievable. Instead, they sought to understand why past research had gone wrong and to improve upon it.

Researchers cited a number of factors that lead pragmatic prediction astray: cognitive heuristics, faulty incentive structures (specifically a lack of accountability for poor predictions), misplaced theoretical frameworks (e.g., neglecting the importance of networks) and mis-specified statistical models (Tetlock & Gardner 2016; Ward et al. 2010). Rectifying these problems, they claimed, would make better pragmatic prediction achievable. Teams of researchers began to produce public predictions of different classes of political phenomena, such as civil wars, genocides, *coups d'état* and interstate disputes (see, e.g., Goldsmith et al. 2013; Goldstone et al. 2010; Hegre et al. 2013; Nanlohy et al. 2017; Ulfelder 2012; Ward et al. 2013). Crowdsourced forecasts were also produced for a range of events by the Good Judgment Project (Tetlock & Gardner 2016), and by the American Political Science Association for American elections (Campbell et al. 2016). Many of these forecasts demonstrated significant improvements on the earlier work that had 'inspired' the research programme.

The success of the pragmatic prediction research programme has led some of its practitioners to suggest that it represents the future of social science. There are certainly numerous ways in which pragmatic prediction could help to advance social science. First, the development of expertise requires not simply repeated practice, but also direct feedback on performance (Kahnemann 2011). If social scientists are compelled to make specific advance predictions about events and these prove to be relatively inaccurate, this can be taken as an indication that something is wrong with the social scientist's model of the world and it must be revised.

Second, social scientists may differ not only with respect to their theories about the world, but also as to how those theories may be tested. For instance, which type of statistical model should be used? Should a statistical model be used at all or is some other kind of research design more appropriate? Pragmatic prediction, by contrast, implies something about events in the outside world which are verifiable and thus less up for debate. For example, it is easier to argue that one's opponent used the wrong statistical method to test one's theory than it is to argue that Syria did not have a civil war or that Zimbabwe did not have a coup.[5]

Third, political science has tended to accept the argument that successful prediction of some datum D constitutes better evidence for a theory than successful accommodation (King et al. 1994; Schrodt 2014).[6] It is possible, of course, that D existed prior to developing the theory without the researcher being aware of it (Hitchcock & Sober 2004; White 2003). However, this assertion is unverifiable, as it relies on our taking it on trust that the researcher really did not know D when formulating the theory. Pragmatic prediction is therefore a good means of testing theories since we can indeed be assured that the researcher was unaware of D when formulating her theory, if D refers to something which has not yet happened. The case for pragmatic prediction in political science therefore rests partly on the assumption that pragmatic prediction is a tougher test for a theory than accommodation. We shall return to the issue of accommodation below.

Finally, some researchers argue that the social world may be so inherently complex that it is futile to seek to understand 'why' questions. Yarkoni and Westfall (2017) describe the 'machine-learning culture' in statistics where data are assumed to be the result of unknown and potentially unknowable processes and the goal is simply to find a set of inputs which reliably predict a set of outputs. They do not suggest this culture should replace the search for causal mechanisms, but rather that it should play a larger role in the discipline. Similarly, Ward (2016), though he writes admiringly of the ancient Greek device which could accurately predict celestial movements without a detailed theoretical knowledge of Hipparchosian astronomy, does not go so far as to say that prediction without an understanding of causal mechanisms is all social science needs to do.

Ward's and Yarkoni and Westfall's reasoning is instrumentalist: we should not 'worry about whether a theory is a true description of the world, or whether electrons "really, really exist". If a theory enables us to make good predictions, what more can we ask?' (Godfrey-Smith 2003: 184). With machine learning, more social scientists might come to this view. Realists – who are committed to the ontology of a world outside of our perceptions of it – counter instrumentalism through a 'no miracle' account. Models predict, not miraculously but because they represent reality in some way. For realists, such as Daniel Dennett (1998) or Don Ross (2014), what we find are patterns in the data. Theoretical entities posited to explain data are judged in terms of their reality by how well they scientifically predict. Theoretical entities are as real, or more real, than entities that we see with our standard perceptual equipment to the extent that those theoretical entities allow us to predict better than standard perception (Ladyman et al. 2007). Seeing might be believing, but theorising is knowledge to the extent it is predictive. For Dennettian realists, the centre of gravity between two objects is just as real as the two objects themselves. In both cases, what we see are patterns in the data (Ladyman et al. 2007).

However, the instrumentalist-realist debate might be considered irrelevant to working political scientists if we judge the worth of political science simply in terms of pragmatic prediction. Such an approach abandons the attempt to explain and understand the world. However, explanations do not have to be so fine-grained that we can distinguish why some algorithm makes a specific pragmatic prediction. We might lack not theoretical understanding, but specific empirical information. The problem with complexity in systems is not so much not understanding the general form in which complex patterns might emerge, but in the specific empirical details of why, in a particular token example, one rather than another pattern emerges.

The case for scientific prediction

As we explained above, in the philosophy of science 'prediction' does not necessarily refer to a future event, rather, a prediction in a formal model can be seen simply in terms of its logical implications, and how these are projected on to empirical phenomena. Accommodation in this sense is a form of scientific prediction. Karl Popper distinguishes 'scientific prediction' from 'unconditional prophecies' by virtue of the fact that the former is 'ordinarily conditional' – that is, such predictions state that Y will occur if a given condition or set of conditions C pertains (Popper 1989). If one knows those conditions, one can in principle derive a prediction of the future value of Y, but Popper doubts that in the social sciences this is ever possible, at least in the long term. For Popper, scientific prediction is primarily about the consequences of an intervention. We need to know the answer to 'why' questions (i.e, to understand causal mechanisms) so as to allow us to understand which interventions can work and under what type of conditions and also to understand the limits of our interventions. These are questions which, in many respects, pragmatic prediction is not always well placed to answer.

One of the hopes of pragmatic prediction is that it can lead to better public policy advice. However, if we do not understand the causal mechanisms underlying pragmatic predictions, then we cannot know how we should best intervene. Many of the variables giving forecasting power to the new class of pragmatic predictive models do not lend themselves to obvious interventions. For example, infant mortality is one of the variables which consistently predicts *coups d'état* (Ulfelder 2012). Does this, however, imply that an intervention which reduced infant mortality while holding everything else constant would also reduce the risk of *coups*?

The pragmatic predictive finding itself cannot answer this question because it is consistent with a variety of causal mechanisms that have quite different implications about the interventions which might reduce the risk of *coups* and hence how decision makers might spend scarce resources. Should a *coup*-vulnerable leader prevent one simply by spending more money on hospital maternity wards? Most political scientists would think this is unlikely because the relationship between infant mortality and *coups* is causally spurious even if powerful in terms of pragmatic prediction. Something else is causing both infant mortality and *coups*. But what? It could be state capacity, so the recommendation might be to spend more resources on improving tax gathering and the state administrative apparatus. It could be economic or some other kind of inequality, lack of development or pernicious cultural beliefs that in some way cause both *coups* and infant mortality. Only by probing the causal mechanisms can we begin to answer this question and formulate a response. That requires further theoretical reasoning, which leads to other projections that need to be tested on the data on *coups d'état*.

This point is especially important if machine-learning algorithms come to be used routinely in pragmatic prediction. While many machine-learning algorithms lend themselves to intuitive interpretations (e.g., decision-tree-type algorithms) (Yarkoni & Westfall 2017), others rely on complex, nonlinear variable transformations and interactions (see, e.g., De Marchi et al. 2004; Lantz 2013). Even where machine-learning algorithms produce what may be interpreted in a similar way to traditional regression coefficients, two algorithms with the same degree of pragmatic predictive power may produce radically different estimates of these quantities (Mullainathan & Spiess 2017). This all makes it harder, if not impossible, to interpret what effect a given variable (or 'input') has on the outcome the algorithm predicts.

Yarkoni and Westfall argue that in this case analysts can simply note which sets of inputs significantly improve the algorithm's predictive performance. In terms of designing possible interventions, however, this is not sufficient. Suppose we have an input variable X, which significantly adds to the predictive power of machine-learning algorithm A that predicts Y. Now suppose you are trying to use the output of A to design an intervention to reduce the incidence of Y. Knowing that X has an important effect on Y via some unknown process in A is not especially helpful. If the effect of X is nonlinear, then the intervention 'more X' could reduce the incidence of Y at some levels but increase it at others. If the effect is interactive with some other input Z, then 'more X' could similarly lead to less Y for certain values of Z, but more Y for others. A scientific prediction informed by a theoretical model can help us to begin to answer these questions.

Big data and massive computing power sometimes show that an ensemble forecast combining several models of social processes is superior for pragmatic prediction to any of those models used singularly (Hindman 2015). In that sense, the models are not rival. It might be that several causal mechanisms are interacting, or counteracting, with each other, and the big data processes applying different models simultaneously can better calculate the effects. Of course, in such cases we might not know quite how these causal mechanisms interact or counteract, and the complexity of the systems make that discovery, at least in the foreseeable future, problematic. However, this does not demonstrate the inapplicability of scientific prediction. Indeed, some scientific predictions might entail that we cannot pragmatically predict. Critics may object that models which produce scientific predictions of this sort are trivial. We now turn to this claim.

The triviality claim

Some formal models produce scientific predictions that provide no pragmatic predictions. They might seem, in Popper's terms, to have little empirical content, so it might be thought their results are trivial. The folk theorem of PD-like games suggests any finite sequence of conjoint strategies is rational (i.e., scientifically predicted) and that number is enormous in an infinitely repeated game (see, e.g., Binmore 2007: 333–340).[7] In other words, the scientific prediction is that anything can happen. The results themselves provide no guide for pragmatic predictions without further empirical input. We might be able to place probability boundaries on what we expect, but these are unlikely to be generated game-theoretically since they are likely to be determined historically and institutionally. They will be specific to the token historical cases. How likely two people are to cooperate with each other will depend more upon their previous interactions and their cultural or institutional background. Game theory itself does not provide the answer. Here it might be claimed that scientific prediction is trivial, or not empirical (Johnson 2017).

Another example is the McKelvey-Schofield theorem that, for n-dimensional issue space in voting games with large numbers, there is no equilibrium. Riker (1982) interprets this theorem to mean that democratic politics is chaotic because any majority decision can be overturned by a new majority, given voters have agenda-setting powers. Institutions provide stability (Shepsle 1979). Calvert's (1995a, b) game-theoretic account of institutions arising as the equilibriums of underlying repeated games is criticised by Johnson (2017: 16–17):

> Calvert's models are neither empirical nor, for that matter, properly explanatory. Calvert notes as much. His models do not account for the emergence of *any* particular

social or political institution. Instead, they are an analytical exercise aimed at producing an existence result.

The problem is that the games produce multiple equilibriums and so do not say what institutions will actually emerge. When Johnson says the models are not 'properly explanatory', he means they do not provide a full explanation of any token institution. That is undoubtedly true. However, we should not think that only full explanations of token institutions are 'properly explanatory' – that is too much to expect of scientific predictions. Scientific prediction is explanation of types, and only in fully deterministic processes, or where generalisations are invariant, does scientific prediction entail token prediction (and can hence be considered fully explanatory of the token phenomenon). To give an analogy, natural selection does not provide full explanation of any token animal or token species. That, in itself, does not make Darwinian natural selection not properly explanatory. What natural selection does and what political models should do is provide parameters that, together with empirical inputs, can provide pragmatic prediction.

So, we have to be clear about what a given model is purporting to explain. Calvert's examples show us that institutions can emerge, *pace* Riker's claims.[8] Johnson suggests we see this as a conceptual exercise, but it is not, as he claims, unempirical; it is simply low in empirical content. It does empirically exclude, for it suggests that even if preferences do not change, we should expect electoral cycles. For some, that is an important result; for others, it is of little interest. We take it there are empirical implications from the account. One is that, given there are multiple equilibriums, if we want to explain the emergence of any particular equilibrium, any particular institution, we will have to look to historical and contingent factors. What differs in the social sciences from the natural sciences is not some different ontology – the truths of game theory or strategic interaction are as universal as the laws of physics – but rather the specificity of subject matter. Political scientists are interested in the precise strategic moves of agents in our world as well as more general effects of institutions on behaviour and general behavioural tendencies. In just the same way, if we want to find out why a particular species died out in one place and not another, we need to look for particular historical and contingent factors. In both cases, type-level generalisations would be an important element of the explanatory story. We can examine, empirically as well as theoretically, what types of institutions flourish in what types of conditions. This is where the inductive and abductive strategies enter.

Accommodation and prediction reconsidered

We have skirted around the issue of the desirability of pragmatic prediction versus mere accommodation within scientific prediction.[9] It is often held that we should be more prepared to believe a theory that predicts a given datum not known in advance of the theory's construction than one that was merely designed to accommodate known data.[10] This point is well taken. However, while epistemically we should have more confidence, there is no logical reason why the order in which a theory was constructed and the data discovered should affect the truth value of theory. It cannot logically matter if a theory was created on a Tuesday or a Thursday just because the data was analysed on the Wednesday. What matters is whether the theory was designed in order to fit the known data (White 2003).

Roger White offers a fuller justification of (strong) predictionism based on how we select a theory. He argues that if the data are well known, the pool of potential theories is narrowed and hence the evidence that the theory is true (the known data) adds nothing further to our belief that it is true. Predicting unknown data does add something further. In other words, the fact that the theory predicts rather than accommodates data increases our confidence in 'some intermediate factor, which in turn confirms the theory' (White 2003: 674). That intermediate factor is the means by which the theory was generated.[11]

Of course, theories are never constructed without any known data, but to account for some known facts or relationships. The key is whether they also provide new (scientific and/or pragmatic) predictions for unknown data. These might be more precise claims about data already collected, which can then be measured more precisely, or they might include some novel predictions (Popper 1972: 82–3, 1974: 1004–9, 1989: 57–9). The important aspect of this philosophy of science debate for working political scientists is methodological. The methodological problem of accommodation is overfitting (Hitchcock & Sober 2004). Theory A might seem more scientifically predictive because it accommodates the data better than a rival theory B, but that is because theory A was overfitted. Overfitting is a major methodological problem in statistical analysis, and one that can lead us to be overconfident in our accommodative theories, creating epistemic problems. Using a theory to predict unknown data gives us more confidence because it helps overcome overfitting.

We take overfitting to be a methodological issue well specified in statistical work. In more casual contexts, however, overfitting occurs when we create causal stories that place excessive emphasis on idiosyncratic features of specific historical cases. This is a, if not *the*, methodological problem for theory construction using qualitative historical analysis. If you know what happened, it is relatively easy to create a plausible theory-driven causal account of why. Indeed, you may be able to create several such plausible theories.[12]

Predictionism, then, is the view that, other things being equal, we should be more inclined to believe a theory which scientifically or pragmatically predicts some datum D than a theory which simply accommodates D. Predictionism is correct to the extent that the case for it is driven by overfitting. However, that does not mean we should discount theories that only explain known data. Nor that we should necessarily privilege analyses that forecast over ones that do not – that is, we should not privilege them as explanatory. Forecasting models might pragmatically predict, but that does not mean they explain. Explanation entails scientific prediction, but prediction does not entail explanation. Scientific prediction is a necessary condition for explanation; it is not sufficient (Dowding 2016: Chapter 3, 2017: 221–223). We argue that accommodationist theory is often explanatorily powerful, though it is stronger still if it enables pragmatic predictions for novel data.

The more technical issue is an epistemological one. Say we have two theories, T_1 and T_2. The first predicts with high accuracy data known in advance; the second less accurately predicts data not known in advance. Assuming other background conditions that might affect our confidence in the expected accuracy to be equal, should we think T_1 is a more attested and plausible explanation of the data than T_2? The predictionist claim is that it is not, and we agree – because of the danger of overfitting. Accommodation runs the risk of overfitting, and therefore:

> Prediction is at least sometimes better than accommodation, both because it can provide a measure against overfitting, and because successful prediction can provide evidence that overfitting has not occurred. When it is known that the data have been accommodated while guarding against the risk of overfitting, however, both of these advantages are lost. Indeed, since accommodation involves making use of relevant evidence, accommodation will sometimes be superior to prediction. (Hitchcock & Sober 2004: 3)

So, we hold weak predictivism, which is the thesis that 'the difference between prediction and accommodation is epistemically relevant only because it tracks or is symptomatic of other differences that are themselves of evidential import' (Hitchcock & Sober 2004: 4). For instance, White describes two situations: one where Fred states in advance that the lottery is rigged in Jane's favour and Jane then wins; the second where Fred hears that Jane won the lottery and then claims that it was rigged in her favour (White 2003). We would be more inclined to believe Fred in the former case. As White notes, however, this is likely to be because we think Fred knew something about the rigging – perhaps he saw Jane tampering with the lottery machine. But that means it is the content of the claim and its logical connection with evidence that is important, rather than its timing.

With accommodation, one can guard against overfitting, using methods to ensure a model that balances simplicity and fit with the data. Good pragmatic predictions constitute evidence that a good method has been adopted, but what matters is the method adopted, not the prediction itself. There is no logical problem with accommodation, with models that scientifically predict but do not (yet) pragmatically predict, but there are potential methodological problems which can lead us to epistemically prefer some theories to others, even if they do not explain the extant data so well.

The most pragmatically predictive model might be less true than a model for which pragmatic predictions prove impossible, for the reason that, if all the factors that have some causal effect were to be placed in a model, then the calculations might prove intractable. However, a simpler model, albeit less accurate since it ignores some factors which exert some causal influence, may provide reasonably good forecasts. Sometimes pragmatic prediction and full explanation (in the sense of including all causally relevant elements) might need to be traded off. While some political scientists, such as Bruce Bueno de Mesquita et al., produce both successful scientific and pragmatic predictions, the work tends to be rather different in character, utilising different methods. The scientific predictions of Bueno de Mesquita et al.'s selectorate theory are tested against past data on economic growth, government spending, war outcomes and leadership survival among other things, whereas their pragmatic predictions on, for instance, nuclear proliferation are derived from more complex models featuring a far larger number of actors and extensive qualitative input from area specialists (Bueno de Mesquita et al. 2005; Bueno de Mesquita 2010). What is important about pragmatic prediction is that it is pragmatic. It is do-able.

The no-miracle claim of realism is that if theories are pragmatically predictive, then they (and their theoretical elements) should describe reality, otherwise the accurate predictions would be a miracle. Hitchcock and Sober (2004) argue that no fitted model will be more predictively accurate overall than the true fitted model; however, if fit is the criterion by which we judge models, then the true fitted model is epistemologically inaccessible (we cannot know we have found it simply by fit). So, while the no-miracle claim is ontologically important, we cannot rely upon it simply in terms of predictive accuracy, because less true theories might be pragmatically more predictive; there are dangers in judging theories only by their pragmatic predictions. Hitchcock and Sober show that unbiased estimates of the predictive accuracy of a model can be obtained by considering both its fit-to-data and its complexity measured by the number of adjustable parameters – the higher its AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) (Schwartz 1978) scores, the greater the information (Hitchcock & Sober 2004). Then a model must fit the data sufficiently better to compensate for loss in simplicity. In this way, we can control for overfitting when comparing rival models. Of course, historical stories with all their detail are overfitted because each element is chosen to illustrate the case being made.

A similar argument may be made with respect to p hacking. P hacking, when researchers try out several statistical analyses and/or data eligibility specifications and then report those that produce statistically significant results, is a major problem in a number of scientific disciplines and has resulted in a widespread discussion as to whether 'science is broken' (Head et al. 2015). Pragmatic prediction is one possible solution to the problem of p hacking: a p hacked statistical finding is unlikely to result in accurate pragmatic predictions of future data. As with overfitting, however, pragmatic prediction is not the only solution. Norms whereby journals require researchers to deposit their data and code online provide a means both to detect and deter p hacking, while statistical methods such as Bayesian model averaging (Montgomery & Nyhan 2010) and extreme bounds analysis (Sala-i-Martin 1997) provide means through which researchers can reassure sceptics of the robustness of their results, while still 'only' accommodating known data. An essential desideratum of a scientific prediction is that it should be both robust and replicable.

We have assumed social science is probabilistic. Is this simply through lack of knowledge, or is it an ontological feature of society? For our argument thus far, the answer to this question does not matter. Given the strategic nature of agency, and supervenience of agency on different types of brain states and social conditions, however, the probabilistic nature of social science is, we suggest, ontological. For example, individual actors are strategic: they respond to incentives around them. So, if someone realises that their actions have been modelled within some parameters, they can refute a pragmatic prediction by changing their behaviour. *Coups d'état* are a major focus for pragmatic prediction. Suppose that the *coup*-forecasting literature advances to the stage of making very accurate predictions (Beger & Ward 2017; Ulfelder 2014, 2015). It is quite likely that coup-vulnerable leaders will take these forecasts on board and adjust their behaviour accordingly – firing, executing or ramping up surveillance of senior military officers; beefing up parallel military forces or intelligence services; and/or boosting military pay. If pragmatic predictions of events such as *coups* achieve the potential many believe possible, then we should expect policy makers to start to pay attention to them, in the same way as they currently do to, for instance, economic forecasts or opinion polls. Once they begin to condition their behaviour accordingly, a problem emerges for pragmatic prediction.

Do poor pragmatic predictions condemn a theoretical model?

A supporter of pragmatic prediction might respond that a theoretical model is fine and useful but should both lend itself to pragmatic predictions and be judged on the accuracy of those pragmatic predictions. But how good is pragmatic prediction as a benchmark for judging the quality of a theory? A good theoretical model may still perform poorly in terms of pragmatic prediction, because the data which we observe in the world are partially truncated by strategic selection. For instance, leaders who back down in international crises are no more likely to lose office than leaders who do not (Schultz 2001). While the model from which this conclusion is drawn was not pragmatically predictive, any model which uses crisis behaviour to predict a leader's tenure is likely to perform poorly. Does this imply that a leader trying to remain in office should not care whether or not she backs down in a crisis? Schultz argues convincingly that this is not so. Only leaders who believe their tenure in office is relatively secure should be willing to back down in a crisis. This conclusion is theoretically driven. A pragmatic predictive model would simply note that backing down in international crises does not add predictive power to a forecasting model of leadership tenure.

For similar reasons, theoretical models can give us some guidance as to what we might expect the marginal effect of a variable to be at levels which are not found in an empirical sample. For instance, it is quite plausible that decision makers set the level of the minimum wage carefully so as to avoid triggering unemployment. Minimum wage levels therefore might not add much predictive power to a model of unemployment. It would be rash, however, to draw the inference that the minimum wage can be set at any level without affecting unemployment.

Theoretical models based on structural considerations may be especially prone to combining low pragmatic, but high scientific pragmatic power. Consider, for instance, Tsebelis' (2002) veto players model. This theory seeks to explain variation in policy stability across polities with reference to the number of veto players in the system. Veto player theory implies, for instance, that labour market reforms would be more likely to pass into law in Britain than in Germany because the former has fewer veto players. Veto player theory, however, would be quite underspecified when it comes to predicting whether a *particular* labour market bill proposed in Germany is likely to become law. This may depend on short-run, time- and country-specific factors such as the content of the bill, the views of key veto players towards it, specific interpersonal dynamics (scandals, personality clashes) within the ruling coalition in Berlin and much else. A 'superforecaster' who incorporates this information (what Tetlock and Gardner (2016) term 'the inside view') will produce a more accurate pragmatic prediction than someone who simply applies veto player theory. If, however, one were drafting a constitution for one newly independent country, or attempting to gauge the limits of legal reforms in another – that is, designing an intervention or assessing the success of one – Tsebelis' explanatory theory would prove more useful, because we do not know what specific, short-term factors may obtain in the distant future, but we do know that on average there will be greater policy stability in polities with more institutionalised veto players.

Similarly, Rapoport's well-known tit-for-tat strategy in the iterated prisoner's dilemma does not easily lend itself to pragmatic prediction (Axelrod 1984). In order to get to a pragmatic prediction, it would be necessary to identify a future situation which approximates an iterated prisoner's dilemma, identify an actor or group of actors whose strategy was analogous to tit-for-tat and identify some measurable payoff for them. Even then, of course, the tit-for-tat strategy may not produce especially good pragmatic predictions, since it omits so many potentially relevant factors. Nonetheless the insights generated by a model of such abstraction can help inform the design of institutions across a range of issues approximating the iterated prisoner's dilemma, from international trade agreements to competition law.

Conversely, the fact that a predictive model based on theory A outperforms one based on theory B does not mean necessarily that theory A is superior to B. Success at pragmatic prediction is not a guarantee of usefulness, at least with respect to the main desiderata of social science. Some successful pragmatically predictive models in political science have been critiqued on the grounds that they include as predictors variables which are effectively a lag of the outcome the model is seeking to predict. Predictive models of American presidential elections which include the president's approval rating or voting intentions fall into this category, as do predictive models of civil or international conflict which include media reports of low-level violence as predictors of higher levels of violence (Jäger 2016). Such models are still useful for many purposes (such as early warning) but are less useful for intervention than a good theoretical model, even if the latter has a less impressive performance by pragmatic predictive accuracy. The advice 'if you don't want a civil war, don't have rival factions fighting each other on the streets' is not useful if one accepts the plausible assumption that, by the time fighting has broken out on the streets, things have escalated to the point at which an intervention is no longer going to work. Even if an intervention is still feasible at this stage, it is not clear, based on the pragmatic prediction alone, what this intervention would be.

To emphasise pragmatic prediction as the justification for political science leaves us vulnerable to claims that we do not really explain anything, and that our models are trivial extensions of known data. The other danger is disappointment caused by the short time horizons of good pragmatic prediction. When government agencies solicit predictions of the future they are often talking about the world decades ahead (see United States Office of the Director of National Intelligence's Global Trends: www.dni.gov/index.php/global-trends-home). For instance, high-technology military equipment often takes decades to move from conception to entry into service and can be expected to remain operational for decades. Similarly, choices about force posture and composition – Do we spend more on the army, the navy or the air force? Do we want an army focused on counterinsurgency or conventional combat? – usually take effect years rather than months after they have been made. At present, pragmatic predictions have only been good over much shorter time scales, usually one year or less.

Predictions on a longer time scale are harder, partly because they involve forecasting future values of the independent variables as well as the dependent ones. Consequently, such pragmatic predictions are rare.[13] Moreover, the record suggests that there may be something intrinsically harder about forecasting on this scale – Tetlock's superforecasters, for instance, while reasonably accurate on predictions with a one-year lead, perform no better than chance when the lead time is even five years into the future (Tetlock & Gardner 2016). While pragmatic prediction on a one-year time scale is impressive, government agencies will often want more, leading potentially to frustration and unfulfilled expectations.

We can see therefore that, while pragmatic prediction is an important desideratum for political science, it is neither necessary nor sufficient as a marker for good political science theories.

Conclusion

Scientific prediction does not necessarily lead to good pragmatic prediction. For one, scientific prediction in theoretical models is directed at types of phenomena and can only be turned to pragmatic prediction when the conditions are filled in. That might require data not currently or practically available. Failure of scientific prediction to provide pragmatic predictions does not necessarily mean we should question the science. Weather forecasting, notoriously inaccurate prior to the 1990s, has become increasingly accurate over the past few decades (Silver 2012; Ward 2016). Yet the value of physics as a science was not widely questioned on the grounds it was unable to lend itself well to weather forecasting. The basic physics underlying the weather was in fact generally well understood; the difficulty lay partly in having the computing power to apply these laws at a sufficient level of granularity to devise acceptably accurate forecasts (Silver 2012). If we are to condemn political science today for not having the ability to forecast at a similar level of granularity, we are applying to it a stricter set of standards than those we apply to the natural sciences.

We can hope and expect that some pragmatic predictions are possible from scientific ones and strive to achieve them. However, we cannot simply judge models on the grounds of pragmatic prediction alone. First, algorithms might be too complex to understand; second, their strongest predictors might be proxies for something else; third, their nonlinear and stochastic nature means they may not provide guidance for policy intervention. Nor should we pin too much on successful pragmatic prediction, as its short time horizons might well lead to disappointment among policy makes. Theory can help in all these regards.

Pragmatic prediction can help with the problems of overfitting but is not the means by which we can guard against overfitting. The methodological problems of much of quantitative political science should not be confused with the debate over the relevant advantages of scientific and pragmatic prediction. This article is a defence of standard approaches to political science, understanding political science as explanatory by producing theoretical models to try to understand how empirical generalisations are produced by causal mechanisms. It is not meant in any way as a critique of forecasting or pragmatic prediction, but rather strikes a cautionary note that these methods should be seen not as an alternative to our past standards, but as an important addition that can help us to improve what we can achieve in political science. Scientific prediction is defensible even when only accommodationist. Properly understood, political science fares far better than many believe in terms of scientific prediction. Pragmatic prediction, when underscored by scientific prediction is important both methodologically and heuristically, but there are dangers in placing too much emphasis upon pragmatic prediction.

Acknowledgements

A version of this article was presented at the SPIR research seminar at the ANU; we thank the participants for their questions and comments. We also thank Will Bosworth, Patrick Dunleavy, Ben Goldsmith, Bruno Jahn, Peter John and Enzo Lenine for their written comments and suggestions.

Notes

1. We prefer the term 'pragmatic prediction' to 'forecasting' since the latter term tends to be used in a slightly narrower context. Our argument is that scientific prediction in formal models are akin to logical implications or entailments, but again we prefer 'scientific prediction' to either of those terms since our usage also includes inductive and abductive inferences which have a more equivocal logical status. Furthermore, our article is designed to highlight issues concerned with the ambiguous use of the term 'prediction' in political science.
2. Our argument is agnostic as to whether a hypothesis test should be interpreted as involving the acceptance or rejection of the hypothesis or simply an update of the prior odds of the hypothesis' truth or falsity – we do not take a stand between Bayesian and frequentist philosophies of statistics and our argument does not depend on accepting one or the other, although below we do suggest some techniques for overcoming problems with scientific prediction that do depend upon Bayesian techniques.
3. An interventionist account defines causation in terms of differences made through interventions and is based upon experimental methods (see, e.g., Woodward 2003).
4. Formal modelling usually implies some mathematical proof. However, the logic of an argument can be set out in prose – what matters is how definitive the prediction is. Non-formal models are often suggestive, but do not exclude as much as more formal ones in their predictions.
5. Though we recognise that coding decisions have to be made, so these issues are not always as straightforward as might first appear.
6. Note that this is true of scientific and not just pragmatic prediction – that is, it might predict that some sub-atomic particle must exist even though it is currently unknown, not that the particle must exist at some point in the future. The prediction-accommodation dispute is related but orthogonal to the scientific-pragmatic prediction distinction.
7. They are rational since they constitute Nash equilibria. The rationality assumption is for prediction, not anything substantially more than that.
8. To be clear, we are not claiming that this means the Calvert story is explanatory because it is like natural selection. The analogy is only meant to suggest that we cannot claim something is not properly explanatory because it only produces type-level explanation.
9. Some political scientists, following Hempel, tend to use the term 'explanation' to refer to what we call 'accommodation' (King et al. 1994; Schrodt 2014; Waltz 1979), but this is too narrow an understanding of explanation (Dowding 2016: Chapter 3). Our use of the term 'accommodation' reflects contemporary usage in the philosophy of science (see Hitchcock & Sober 2004; White 2003).
10. See White (2003: 655) for a formal definition of accommodationist and predictionist theory.
11. Note that White never assumes the predictions are generated without an explanatory theory – in that sense the philosophers of science are defending pragmatic prediction as part of, not in place of, scientific prediction.
12. Unless we make it clear that by 'overfitting' we mean 'statistical overfitting', we use this term in this more expansionist sense.
13. Hegre et al. (2013) is one of the few exceptions.

References

Axelrod, R. (1984). *The evolution of co-operation*. London: Penguin.

Beger, A. & Ward, D. (2017). Lessons from near real-time forecasting of irregular leadership changes. *Journal of Peace Research* 54(2): 141–156.

Binmore, K. (2007). *Playing for real: A text on game theory*. Oxford: Oxford University Press.

Boyd, R. (1984). The current status of scientific realism. In J. Leplin (ed.), *Scientific realism*. Berkeley, CA: University of California Press.

Bueno de Mesquita, B. (2010). *The predictioneer's game: Using the logic of brazen self-interest to see and shape the future*. New York: Random House.

Bueno de Mesquita, B., Smith, A., Siverson, R.M. & Morrow, J.D. (2005). *The logic of political survival*. Cambridge, MA: MIT Press.

Calvert, R.L. (1995a). Rational actors, equilibrium and social institutions. In J. Knight & I. Sened (eds), *Explaining social institutions*. Ann Arbor, MI: University of Michigan Press.

Calvert, R.L. (1995b). The rational choice theory of social institutions: Cooperation, coordination and communication. In J.S. Banks & E.A. Hanushek (eds), *Modern political economy*. Cambridge: Cambridge University Press.

Campbell, J.E. et al. (2016). Symposium: Forecasting the 2016 American national election. *Political Science and Politics* 49(4): 649–690.

Carnap, R. (1966). *Philosophical foundations of physics: An introduction to the philosophy of science*. New York: Basic Books.

Clarke, K.A. & Primo, D.M. (2012). *A model discipline: Political science and the logic of representations*. Oxford: Oxford University Press.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge: Cambridge University Press.

De Marchi, S., Gelpi, C. & Grynaviski, J.D. (2004). Untangling neural nets. *American Political Science Review* 98(2): 371–378.

Dennett, D.C. (1998). Real patterns. In *Brainchildren: Essays on designing minds*. Harmondsworth: Penguin.

Dowding, K. (2016). *The philosophy and methods of political science*. Basingstoke: Palgrave Macmillan.

Dowding, K. (2017). So much to say: Response to commentators. *Political Studies Review* 15(2): 217–230.

Druckman, J.N. et al. (2011). *Cambridge handbook of experimental political science*. Cambridge: Cambridge University Press.

Fearon, J.D. & Laitin, D.D. (2003). Ethnicity, insurgency and civil war. *American Political Science Review* 97(1): 75–90.

Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science*. Chicago, IL: University of Chicago Press.

Goldsmith, B.E. et al. (2013). Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003. *Journal of Peace Research* 50(4): 437–452.

Goldstone, J.A. et al. (2010). A global model for forecasting political instability. *American Journal of Political Science* 54(1): 190–208.

Harré, R. (1986). *Varieties of realism*. Oxford: Blackwell.

Head, M. et al. (2015). The extent and consequences of p-hacking in science. *PLoS Biology* 13(3): e1002106.

Hegre, H. et al. (2013). Predicting armed conflict, 2010–2050. *International Studies Quarterly* 57(2): 250–270.

Hempel, C. (1942). The function of general laws in history. *Journal of Philosophy* 39(1): 35–48.

Hindman, M. (2015). Building better models: Prediction, replication and machine learning in the social sciences. *Annals of the American Academy of Political and Social Science* 659(1): 48–62.

Hitchcock, C. & Sober, E. (2004). Prediction versus accommodation and the risk of overfitting. *British Journal for the Philosophy of Science* 55(1): 1–34.

Jäger, K. (2016). Not a new gold standard: Even big data cannot predict the future. *Critical Review* 28(3–4): 335–355.

Johnson, J. (2017). Models-as-fables: An Alternative to the Standard Rationale for Using Formal Models in Political Science. Paper presented at Midwest Political Science Association, Chicago.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus & Giroux.

King, G., Keohane, R. & Verba, S. (1994). *Designing social inquiry*. Princeton, NJ: Princeton University Press.

Ladyman, J. et al. (2007). *Every thing must go*. Oxford: Oxford University Press.

Lantz, B. (2013). *Machine learning with R*. Birmingham, AL: Packt.

Lipton, P. (1991). *Inference to the best explanation*. London: Routledge.

Montgomery, J. & Nyhan, B. (2010). Bayesian model averaging: Theoretical developments and practical applications. *Political Analysis* 18(2): 245–270.

Morton, R. & Williams, K. (2010). *Experimental political science and the study of causality: From nature to the lab*. Cambridge: Cambridge University Press.

Mullainathan, S. & Spiess, J. (2017). Machine learning: An applied econometric approah. *Journal of Economic Perspectives* 312(2): 87–106.
Nanlohy, S., Butcher, C. & Goldsmith, B.E. (2017). The policy value of quantitative atrocity forecasting models. *RUSI Journal* 162(2): 24.
Popper, K. (1972). *The logic of scientific discovery*, rev. edn. London: Hutchinson.
Popper, K. (1974). Replies to my critics. In P.A. Schilpp (ed.), *The philosophy of Karl Popper: Library of living philosophers*. La Salle, IL: Open Court.
Popper, K. (1983). *Realism and the aim of science*. London: Routledge.
Popper, K. (1989). *Conjectures and refutations: The growth of scientific knowledge*, rev. edn. London: Routledge.
Riker, W.H. (1982). *Liberalism against populism: A confrontation between the theory of democracy and the theory of social choice*. San Francisco, CA: WH Freeman & Co.
Ross, D. (2014). *Philosophy of economics*. Basingstoke: Palgrave Macmillan.
Sala-i-Martin, X. (1997). I just ran two million regressions. *American Economic Review* 87(2): 178–183.
Schrodt, P. (2014). Seven deadly sins of contemporary quantitative political science. *Journal of Peace Research* 51(2): 287–300.
Schultz, K.A. (2001). Looking for audience costs. *Journal of Conflict Resolution* 45(1): 32–60.
Schwartz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics* 6(2): 461–465.
Shepsle, K.A. (1979). Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science* 23(1): 27–59.
Silver, N. (2012). *The signal and the noise: Why so many expert predictions fail – but some don't*. New York: Penguin Press.
Tetlock, P. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton University Press.
Tetlock, P. & Gardner, D. (2016). *Superforecasting: The art and science of prediction*. London: Random House.
Tsebelis, G. (2002). *Veto players: How political instutions work*. Princeton, NJ: Princeton University Press.
Ulfelder, J. (2012). The watch list: Predicting state failure isn't as hard as you think. *Foreign Policy*, 18 June. Available online at: http://foreignpolicy.com/2012/06/18/the-watch-list/
Ulfelder, J. (2014). Coup forecasts for 2014. Available online at: https://dartthrowingchimp.wordpress.com/2014/01/25/coup-forecasts-for-2014/
Ulfelder, J. (2015). Statistical assessment of coup risk for 2015. Available online at: https://dartthrowingchimp.wordpress.com/2015/01/17/statistical-assessments-of-coup-risk-for-2015/
Waltz, K. (1979). *Theory of international politics*. Boston, MA: McGraw-Hill.
Ward, M.D. (2016). Can we predict politics? Toward what end? *Journal of Global Security Studies* 1(1): 80–91.
Ward, M.D., Greenhill, B.D. & Bakke, K.M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47(4), 363–375.
Ward, M.D. et al. (2013). Learning from the past and stepping into the future: Toward a new generation of conflict prediction. *International Studies Review* 15(4): 473–490.
White, R. (2003). The epistemic advantage of prediction over accommodation. *Mind* 112(448): 653–683.
Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell.
Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
Yarkoni, T. & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science* 12(6): 1100–1122.

Address for correspondence: Keith Dowding, School of Politics and International Relations, Australian National University, Haydon-Allen Building 22, Acton ACT 0200, Australia. Email: keith.dowding@anu.edu.au