Prediction, history and political science

Abstract
Political science usually requires either prediction or contextual historical work to succeed. Because of the difficulty of prediction, the main focus should often be contextual historical work. Both of these methods favor narrow-scope explanations. I illustrate, via an example, the role that this still leaves for theory. I conclude by assessing the scope for political science to offer policy advice. All of this tells against several practices that are widespread in the discipline.

1. Introduction and main claims

What methods are appropriate for political science?[1] I argue for prediction and contextual historical work, and that these two methods each favor narrow-scope explanations. The arguments apply to field sciences generally, but I concentrate on political science.[2]

The structure of the paper is as follows: in this section, I state some basic distinctions and then my main theses and targets. In section 2, I make the case for what has been called 'scientific prediction'. In sections 3 to 7, I examine the implications for scientific prediction of a series of methodological challenges, namely underdetermination, causal fragility, and noise. In section 8, I discuss what role is left for theory, and in section 9, I illustrate this role via an example. Finally, in section 10, I consider the scope for political science to offer policy advice.

1.1 Different types of prediction

The word 'prediction' is ambiguous. Two distinctions are important:

1) *Forward-looking* prediction: predictions about future data; versus
Retrospective prediction: predictions about past data.

2) *Simple* prediction: the attempt simply to predict future or past outcomes; versus
Scientific prediction: the attempt to empirically test particular theories or models. Often, scientific prediction concerns the impact of varying just one focal variable, holding other variables constant or in some other way controlling for confounders. Scientific prediction may also concern conditional predictions, such as what will happen if a policy-maker intervenes in a certain way.

I borrow the term 'scientific prediction' from Dowding and Miller (2019).[3] Dowding and Miller discuss only the second of the above distinctions explicitly, although they

[1] My focus in this paper is on epistemological rather than ideological or political considerations.
[2] By 'field sciences' I mean non-laboratory investigations of systems that are not engineered artefacts.
[3] Dowding and Miller's distinction between scientific and non-scientific prediction is fruitful. (They label non-scientific prediction 'pragmatic'.) As they note, within philosophy of science this distinction was noted by Popper (1989), among others. See also Salmon (1981) and Watkins (1968).

implicitly nod to the first distinction as well by defining non-scientific prediction to encompass only forward-looking prediction. I keep the two distinctions distinct.

To clarify the relations between the different kinds of prediction: first, scientific prediction is a subset of simple prediction. A simple prediction is in addition scientific only if it concerns data gathered in certain epistemologically propitious conditions, namely those conditions suitable for testing a theory or model. Second, scientific prediction, and simple prediction generally, can be either forward-looking or retrospective. Retrospective simple prediction is just description of the past.

1.2 Different types of explanation
There is a difference of degree between:
Wide-scope explanations, which apply to many cases; versus
Narrow-scope explanations, which are very local or contextual and that, in the limit, may apply only to a single case.[4]

An explanation might or might not be derived from a theory or model (section 8). Theories and models themselves vary in scope, but what matters for our purposes is the scope of explanations (section 8).

Throughout, I have in mind *causal* explanations. This is not because I rule out other kinds of explanation but rather is because causal explanations are the kind of explanation usually offered in political science.[5] For the most part, the relevant models are causal too.

1.3 Main theses
I argue that, in political science:
Thesis 1) While any kind of prediction is desirable, scientific prediction is especially so.
Thesis 2) Usually, the purpose of scientific prediction is fulfilled only by forward-looking prediction or by contextual historical work.
Thesis 3) Usually, narrow-scope explanation is favored.

1.4 Targets
Some widespread practices in political science fall foul of the above theses:
1) A lack of emphasis on forward-looking prediction.
2) Forward-looking predictions, when they are made, being wide-scope.
3) Wide-scope retrospective testing, such as much large-n statistical work.
4) Resources being devoted to 'pure theory', in other words devoted to building up a repertoire of wide-scope models divorced from frequent empirical application.

2. Scientific prediction is desirable

Causal knowledge is central to political science, as it is to most sciences, because it is the key to explanation, intervention and extrapolation. The standard template for causal

[4] Exactly how we should individuate 'cases' here is no doubt itself contextual.
[5] There is a literature on the relation between causal and *structural* or *functional* explanations. I do not discuss that here.

inference is to change one variable while keeping all else equal. Scientific prediction typically concerns the results of just such changes, and therefore delivers causal knowledge.

Simple prediction, by contrast, is concerned purely with actual outcomes, which in field environments are typically the result of many different variables varying all at once. Therefore, simple predictions that are non-scientific do not deliver causal knowledge even when accurate. Thus, Thesis 1: Scientific prediction is especially desirable.

Confirmation of theory is useful for purposes other than causal inference too (Dowding and Miller 2019). Therefore, because it is the means to get confirmation of theory, scientific prediction is useful for these other purposes too.

3. Underdetermination

Suppose that candidate X wins an election. There are likely many plausible explanations of why X won. This creates a problem of underdetermination: the mere occurrence of the same headline fact, namely that X won, cannot by itself discriminate between the many different explanations of that fact. This problem is likely to be especially acute if an explanandum is qualitative, as when we seek to explain merely who won an election rather than by exactly how much. But even in quantitative cases, as when we seek to explain X's margin of victory, the bar is often lowered, because only a partial or approximate fit with the data is demanded: a model may be endorsed even though it 'explains' only some but not all of the relevant variation.[6] Such a lowered bar might be cleared by many models, and so the underdetermination problem remains.

One solution to underdetermination is forward-looking prediction. If you have to stick your neck out in advance, that removes the possibility of fudging awkward outcomes after the fact.[7] Many models or pundits might be able to explain the outcome of one election, even of five elections, relatively plausibly. But few are able to predict five correctly in advance. Lucky guesses, although possible of one election, are much less likely of five, especially if the target of prediction is quantitative. Insisting that accurate prediction be forward-looking discriminates between competing models more effectively than does allowing accurate prediction to be retrospective.[8]

Logically speaking, indefinitely many models fit any given body of evidence. But methodologically speaking, the key issue is whether those different models, in addition to being logically possible, are also plausible or to be taken seriously. It is at this

[6] Not all explanations are derived from models (section 8), but the points in the text carry through *mutatis mutandis*.

[7] Howson and Urbach (1993) and Worrall (2014), among others, give formal Bayesian demonstrations of this point. A similar conclusion can be demonstrated in non-Bayesian ways too.

[8] At least for the most part. The advantage is epistemic and contingent rather than logically necessary. But, while there are cases in which forward-looking prediction is not favored, in political science it usually is.

methodological level that the confirmatory asymmetry between forward-looking and retrospective prediction carries bite.[9]

4. Contextual historical work

Forward-looking prediction is not the only solution to the underdetermination problem. A second solution is if there are *no plausible alternative* explanations, in which case the mere fact of retrospective fit is decisive: if no other plausible theory or model can accommodate the past evidence then the fact that your model can tells strongly in its favor. In political science though, this situation applies only rarely. It is hard to prove this claim in a non-anecdotal way but, as with the election example, in political science there are usually many plausible explanations after the fact (Dowding 2016).

A third solution to the underdetermination problem is more promising: we may gather *additional evidence* that favors one explanation over others (Northcott 2019a).[10] If so, there is no need to rely on forward-looking prediction to break an epistemic tie. Additional evidence may favor, say, one explanation of an election result over others. This evidence might take the form of post-election interviews of voters, or of comparisons of vote shares in different districts cross-referenced by potentially explanatory demographic and economic variables.

Happily, gathering such additional evidence is usually possible.[11] It is what historians do all the time. Such additional evidence is often idiographic; that is, it often concerns causal relations that are sui generis and local. The explanations that result thus tend to be specific to the particular case, appealing to local details. So, contextual historical work leads to narrow-scope explanations. Sui generis local details are typically not included in wide-scope models or theories, which seek to capture factors that recur across contexts.

Thus, combining this and the previous section, Thesis 2: Usually, the purpose of scientific prediction is fulfilled only by forward-looking prediction or by contextual historical work. We therefore face a choice: either to make forward-looking predictions, or else to engage in historian-like detailed sifting of the evidence.

Thus, also, one ‘half’ of Thesis 3: Usually, contextual historical work favors narrow-scope explanations. (For why forward-looking prediction does too, see section 5 below.)

This endorsement of contextual historical work brings with it a challenge: is reliable causal inference possible in singular historical cases? The ubiquity of historical controversies suggests that the problem of underdetermination is still present. Formal techniques of causal inference, such as experiments or statistical analysis, are usually

[9] A recent movement in political science seeks to prioritize forward-looking prediction (see Dowding and Miller 2019 for references). As per section 2, I agree with Dowding and Miller that forward-looking prediction should in addition be scientific.

[10] Formally, this is a version of the no-plausible-alternatives solution, but now with respect to an augmented body of evidence.

[11] Formally, such gathering of additional evidence enables retrospective scientific prediction because the additional evidence can be used to test one explanation against another.

inapplicable. In reply, how skeptical do you want to be? Take Holocaust denial, which implies a denial of many singular causal claims. Setting aside the ugly moral and political dimensions, the skepticism behind Holocaust denial is untenably extreme *epistemologically*, at least with respect to many basic facts of the case. Total skepticism with respect to mundane causal inferences about individual human actions or the social world, is similarly extreme: such inferences are, like observation itself, although fallible, usually reliable. Historians' causal inferences, usually backed up by copious archival and other evidence, and by well-supported background knowledge, are just extensions of these everyday procedures.

Some historical controversies do persist nevertheless. If, after contextual historical work, it remains underdetermined which explanation is correct, then the only epistemic tie-breaker left is forward-looking prediction: which explanation's predictions are borne out? Often, this tie-breaker is unavailable, perhaps because the relevant events occurred long ago. If so, we are stuck: we must concede that we do not know which explanation is correct.

Total skepticism about narrow-scope historical causal inference would negate this section's argument for narrow-scope explanation. (I present other arguments for narrow-scope explanation shortly.) A more reasonable partial skepticism leaves the endorsement of contextual historical work less widely applicable rather than negated, and therefore this section's argument for narrow-scope explanation also less widely applicable. What is at stake is whether contextual historical work can obviate the need for forward-looking prediction when confirming an explanation. When forward-looking prediction is very difficult, as it often is (section 6), what is at stake is therefore whether political science can succeed. If we deny the possibility of narrow-scope historical causal inference, we deny most of political science.

5. Causal fragility

Consider, for a moment, an example from physics. Coulomb's Law describes the electrostatic force between charged particles. Although a particular charged particle's trajectory may not be predicted accurately by Coulomb's Law because of the presence of other forces, Coulomb's Law does identify one of the forces present, and so it does explain the charged particle's motion partially.[12] How is this explanatory claim warranted? By the predictive success of Coulomb's Law in a different context, namely a laboratory experiment, combined with a *stability* assumption that the causal relation demonstrated in the other context is still present in our context (Cartwright 1989).

Label the opposite of such causal stability, *causal fragility*. It tells against wide-scope explanations twice over.

[12] See (Northcott 2012, 2013) for the relevant sense of partial explanation.

The first reason is metaphysical: causal fragility means that causal relations themselves are not wide-scope. Therefore, explanations based on those causal relations likewise cannot be wide-scope.[13]

The second reason is epistemological: if causal relations are fragile, then even if empirical warrant is achieved in one context, fresh empirical warrant is required again for each new context. Empirical warrant can no longer be deferred; it must be prioritized continuously. A wide-scope explanation cannot automatically be imported, in the manner of Coulomb's Law, on the back of success in other contexts. A model's claim to have identified a causal relation in the field is not established by that model's empirical success in the laboratory.

Thus, if causal fragility is usual in political science, then Thesis 3: Usually, narrow-scope explanation is favored.

Causal relations are often stable in sciences such as laboratory physics, Coulomb's Law being a paradigm example. What about in field sciences such as political science? Any answer is inevitably somewhat anecdotal, but it seems that in field sciences causal relations are indeed usually fragile: they tend not to generalize easily.

Consider political elections. The causal relations between demographic variables and voter preference change from election to election, and even during elections and between different regions. The causal relations between demographic variables and voter enthusiasm change frequently too, as do those between economic variables and election outcomes, and between likelihood to vote and answers about that to opinion pollsters (Northcott 2019b).

Field scientists' own practice often implicitly assumes that causes are fragile. A famous study showed that, in one circumstance, raising the minimum wage increases employment (Card and Krueger 1994). But in other circumstances, it does not: say, when the minimum wage is already very high, when it is raised by a large amount, or when economic conditions are different. In response, crucially, rather than search for countervailing causes that outweigh the original employment-increasing one, instead researchers just assume that the original employment-increasing cause no longer obtains (Reiss 2008, 173-6). This response is arguably typical in economics, and it implicitly assumes causal fragility.

Causal fragility extends beyond just the social sciences. It is arguably typical of ecology (Elliott-Graves 2018, Sagoff 2016), of field biology more widely (Dupré 2012), and of data science too (Pietsch 2016). In complex systems generally, an explanation that works today often does not work tomorrow. A significant predictor of causal fragility seems to be just that we are in a context of unshielded field phenomena.

[13] This is true even if the causal relations *in a model* are wide-scope. A causal model being wide-scope does not imply that those causal relations are wide-scope in the world, nor that an explanation derived from the model is wide-scope (section 8).

Further evidence for the ubiquity of causal fragility is the ubiquity of the problem of external validity. In social science, external validity can rarely if ever be assumed (Levitt and List 2007; Reiss 2008, 92-6): results from laboratory experiments are notoriously unreliable in the field, presumably because of the huge range of new contextual cues and inputs in a field environment. In other words, causal relations discovered in the laboratory are fragile. Again, scientists' own practice often implicitly concedes the point. When field operation of a mechanism really needs to be ensured, extensive contextual testing and simulation is demanded, as in the design of the US government auctions of electromagnetic spectrum in the 1990s (Alexandrova 2008, Alexandrova and Northcott 2009).

Similar remarks apply not just to external validity but also to extrapolation generally. The causal relations underpinning policy interventions or field trials in one context typically cannot just be assumed to carry over to another; they are too fragile for that (Cartwright 2019, Khosrowi 2019, Cartwright and Hardie 2012, Steel 2008).

6. Noise

A familiar difficulty in field sciences is to distinguish signal from noise. Unshielded environments are typically 'noisy' in the sense that it is hard to isolate the impact of one factor alone. A model might posit that X causes Y, so to test the model we must measure the impact of X on Y. But if other factors A, B, and C also impact on Y, then scientific prediction is more challenging than simply predicting Y after a change in X, because we need to shield off or control for the influence of A, B, and C too.[14] In field settings, such shielding off is difficult, and therefore so is scientific prediction.

Political science is not a laboratory science; it is concerned with noisy, unshielded field environments, and so scientific prediction is difficult. There are some well-known workarounds. One is natural experiments, when processes outside the investigator's control happen to divide a sample into treatment and control groups in the same way that an experimenter would have. Others include quasi-experiments, randomized field trials and laboratory experiments (Northcott 2019a). For the many occasions when experiments are not possible, an array of statistical techniques have been developed to try to infer causes from non-experimental data. These various workarounds have different strengths and weaknesses, but they all share two serious difficulties. The first difficulty is scope: practical and ethical limitations mean there is only a limited range of political questions that experiments can usefully elucidate, while statistical methods require large samples and so are difficult to apply to explanatory claims that are narrow-scope. The second difficulty is external validity (section 5): only rarely do causal inferences from an experimental or other context extrapolate reliably to a new context.

The problem of noise therefore has several consequences. First, scientific prediction can be achieved only by forward-looking rather than retrospective prediction. In a shielded environment, matters can be set up so that only one (salient) model or theory is under

[14] It is because field environments are typically noisy that simple prediction is typically not scientific: simple prediction takes account of the impact on Y not just of X but also of A, B, and C.

test, as in controlled experiments, thereby overcoming the problem of underdetermination, thereby enabling retrospective scientific prediction. But political science rarely concerns shielded environments. Neither can statistical surrogates often fill the gap. Thus, retrospective scientific prediction cannot be achieved, and the first disjunct of Thesis 2 is reaffirmed: forward-looking prediction is favored.

Second, because forward-looking scientific prediction is difficult, usually we must fall back on the second disjunct of Thesis 2, namely contextual historical work, which in turn (section 4) favors narrow-scope explanations. Thus, Thesis 3: usually, narrow-scope explanation is favored.

Third, narrow-scope explanation is favored via another route too. Wide-scope explanations inevitably miss sui generis local causal relations, and so typically identify only some of the causal relations present in a given field situation. There is no shame in that: in noisy environments, such partial explanations are often the best that can be hoped for. But a partial explanation requires empirical warrant just like any other explanation, and given such explanations' empirical inaccuracy this warrant must be imported from empirical success elsewhere. Applications of Coulomb's Law outside the laboratory, for example, import warrant from the Law's empirical success inside the laboratory. Causal fragility threatens this 'imported warrant' strategy because empirical warrant from elsewhere may not travel (section 5). The ubiquitous difficulty that noise creates for accurate scientific prediction now threatens the 'imported warrant' strategy in a new way: it is likely that there is no full empirical success *any*where, and thus no empirical warrant anywhere available to *be* imported. There is no analogue of the successful Coulomb laboratory experiment (Northcott 2017, MS).

To illustrate the difficulty: many competing models seek to explain election results. Suppose that one model says election results are caused by GDP growth in the preceding year. In any actual election, likely many other factors too are causally relevant, so the GDP model will be explanatory at best only partially. If we could somehow tweak a polity so that only GDP was altered and then see how this tweak impacted an election result, then we could test the GDP model and *only* that model. Obviously, such an election experiment is impossible though: which is precisely the point. Because of noise, there is no successful scientific prediction anywhere, either in the context at hand or in other contexts. The GDP model's partial explanation is left without warrant (Northcott 2015).[15]

The only solution is empirical accuracy in the case at hand. In a noisy environment with many ever-changing and sui generis causal relations, this implies a causal description – and thus an explanation – that is narrow-scope.[16]

[15] The GDP model's wide-scope prediction is not empirically confirmed in the particular case. Because of noise and causal fragility, this is the frequent fate of wide-scope predictions. In principle, evidence collected across many elections could favour some wide-scope models over others, thus alleviating underdetermination. But this method would have to assume that the relevant causal structures are stable across many elections, which is dubious

[16] Another way to see the same point: noise means that any empirically accurate explanation is likely multi-factorial. Any multi-factorial explanation is more likely than a single-factor explanation to be narrow-

Overall, the problem of noise therefore provides new support for Thesis 2: either forward-looking prediction or contextual historical work is required. It also provides new support, in two ways, for Thesis 3: narrow-scope explanation is favored.

The problem of noise is closely related to the problem of *overfitting*. In fields such as machine learning and statistics, a model normally has a number of free parameters, which leaves considerable flexibility when fitting the model to data. The problem of overfitting is that tweaking parameter values to ensure maximal fit with every idiosyncratic detail of past data often reduces predictive accuracy with respect to future data. How close a fit, then, should we aim for? It is hard to know. At root, overfitting is an underdetermination problem caused by noise: because data are noisy, a precise fit with a model is implausible.

A standard solution to overfitting is to test competing models on data not used in those models' formulation or calibration. If free parameters must be fixed in advance of this testing, then 'cheating' is made impossible, i.e. the free parameters cannot be adjusted to fit outcomes after the fact.

This solution is effective. But because the data used to test between competing models need only be independent of model estimation, there is no necessity for the relevant predictions to be forward-looking. Does the solution therefore undercut Thesis 2, which favors forward-looking over retrospective prediction? No. If an explanation is narrow-scope, then usually there is not a large stock of relevant past data available, and so the only predictions available for countering overfitting are forward-looking ones. In political science, successful explanations usually are indeed narrow-scope. Therefore, in political science Thesis 2's favoring of forward-looking over retrospective scientific prediction is (usually) endorsed.

7. Laboratory versus field sciences

Laboratory sciences avoid the problems discussed in this paper. Shielded, controlled experiments avoid the underdetermination and noise problems, and the causal relations that laboratory sciences deal with seem to be less fragile.[17] As a result, wide-scope theories and models succeed empirically, as do wide-scope explanations derived from them. Retrospective prediction is sufficient. Explanations derived from Newtonian force models are a paradigm case: they are wide-scope and they can be satisfactorily tested by retrospective evidence.

Just the opposite is true of field sciences: the problems of underdetermination and noise are not avoided, and causes tend to be fragile. As a result, forward-looking prediction is favored. If forward-looking prediction is too difficult, then only contextual historical

scope, because it contains more factors (and interactions between them) that are potentially sensitive to a change in context.

[17] This may well not be coincidence: sciences whose causal relations are stable have more to gain from investigating via laboratory experiments. But I do not explore that suggestion here.

work will do. Only narrow-scope explanations succeed. Theses 2 and 3 apply. And they apply to political science because political science is predominantly a field science.

8. The role of theory

A hallmark of science is that it has ambition beyond singular explanations: it also aims for wide-scope theory. But if forward-looking scientific prediction is usually infeasible because of noise and causal fragility, then usually we must turn to contextual historical work, for which, in unshielded field environments, wide-scope theory is ill-suited. What role, then, can be salvaged for theory?

We should not be blinded by famous physics: not all theories are universal regularities written in mathematical form. Theoretical work in political science is better understood via the *toolbox* view, according to which theories are individual items in an overall repertoire or toolbox. No theory is thought to apply universally or across a whole sample, but any one or more theory might apply in any given case (Cartwright 1999). In political science, a 'theory' in this sense will typically be a causal model or mechanism. In a complementary vein, scientific explanation is not taken to require universal laws; instead, explanation requires only causal relations, whose scope may sometimes be very local (Woodward 2003).

Theory development consists in the expansion and refinement of this toolbox. This expansion and refinement cannot be done in an empirical vacuum: by applying models from the toolbox to real cases, we both sharpen our sense of when a particular model is likely to be applicable, and also sharpen the model itself by learning from experience what aspects of it gain empirical traction (Ylikoski 2019). Such work is essential. Insulation from empirical application is seriously harmful.[18]

Many times, models from the toolbox are putatively wide-scope. Does this vitiate the conclusion that explanations in political science are usually narrow-scope? No, because even if a model is wide-scope, it may still deliver an explanation that is true of only a few cases, i.e. that is narrow-scope. It is true that contextual-historical work inevitably draws on background knowledge, and this background knowledge may in turn draw on wide-scope theory. But any explanation is a particular combination of such knowledge, and may apply only narrowly.

Because a given toolbox model typically applies only to some or a few cases in a sample, the use of large-n statistical methods that assume otherwise is problematized. I cannot do full justice to this issue here. But it is a mistake, for example, to use a statistical regression to test simplistically whether a toolbox model 'is confirmed' in a sample as a whole.

[18] A common criticism of orthodox economic theory is precisely that it has been developed too remotely from empirical application (Northcott 2018, Northcott and Alexandrova 2015, Farmer 2013). Some strands of theory in political science may be vulnerable to the same criticism.

Within the toolbox view, there is an important distinction between two ways in which a model leads to an explanation. On the *causalist* view, relations between terms in a model correspond to causal relations in the world, and so causal explanations can be read directly off the model, at least in successful cases. On the rival *heuristicist* view, the role of a model is more indirect. A model or models may helpfully suggest new categories or lines of enquiry to explore, but supplementary empirical work is required to develop an eventual causal hypothesis that is not itself derivable from the model or models. It is this eventual causal hypothesis that furnishes the explanation. The original model or models are not themselves tested; instead, the demand for testing is transferred to the eventual causal hypothesis.[19]

When should we adopt a causalist view and when a heuristicist one? The less idealized and the more contextual a model, the more likely it is we can read off actual causes from it directly, and so the more likely it is that a causalist view is appropriate. And the more contextual a model, the more likely that an explanation derived from it is narrow-scope. Explanations achieved via the heuristicist route, meanwhile, are usually narrow-scope too, because of the reliance on contextual empirical investigation over and above the original (typically wide-scope) model. Either way, usually we end up with explanations that are narrow-scope, in accordance with Thesis 3.

In summary so far: because of underdetermination and causal fragility, and because both experiments and accurate forward-looking prediction are difficult, investigation in field sciences should usually be contextual-historical. Explanations are usually narrow-scope (sections 2 to 7). This tells against theories understood as applying everywhere. But it does not tell against theories understood as per the toolbox view, which are endorsed so long as they are not developed in isolation from empirical application. In this way, there is still a role for theory, for scientists usefully to develop it, and for cross-contextual scientific achievement.

9. An exemplar of theory at work: della Porta on political violence

Donatella della Porta's *Social Movements, Political Violence, and the State* (1995) is a highly influential study of political violence in 1960s and 1970s Italy and Germany.[20] It is well known for, among other things, emphasizing the non-ideological determinants of violent actors' behaviors. Della Porta's primary goal is well-evidenced causal explanations and to this end she adopts, in effect, a contextual-historical approach. Despite this, her book is famous for being innovative theoretically. How so? Like game-changing work in history and field sciences generally, she provides new categories and outlooks that successors are obliged to consider. A bedrock of her approach is explanatory pluralism, by which she means a willingness to incorporate multiple

[19] The heuristicist view was originally inspired by cases of successful auctions in which the eventual auction mechanism is not derived, or derivable, from auction theory alone, but rather requires extensive supplementary experimental and practical development (Alexandrova 2008, Alexandrova and Northcott 2009).

[20] Over 1500 citations, according to Google Scholar.

theoretical approaches, and to add new ones of her own, whenever these pay their way by enabling new causal explanations to be identified.[21] This is the toolbox view in action.

For example, what explains the behavior of violent groups? Some previous work focused on broad sociological determinants, such as the scope within a polity for expression of political frustrations; other previous work focused on rational-choice explanations of what tactics might best achieve a group's ideological goals. Della Porta deviates from both of these. She examines organizational dynamics at the group rather than society level, and even though those dynamics are 'irrational' in the sense of not being driven by the groups' ostensible ideological goals (1995, 116-33). Her analysis begins with arrests by police. These arrests disproportionately weaken those groups that are organized loosely, creating a selection effect in favor of groups that are more centralized and compartmentalized. This leads to reduced recruitment, and so to subsequent evolution becoming dominated by internal factors. Targets are chosen to achieve internal goals such as discipline or self-defense (robberies, shoot-outs during arrests, punishment of 'traitors') rather than, as earlier, external goals such as propaganda or campaigning (actions against unpopular factories or businesses). The emphasis on self-defense rather than recruitment leads to tactics becoming increasingly lethal and bloody. Ideology evolves accordingly, becoming decreasingly comprehensible to outsiders, with less emphasis on propaganda for external consumption and more emphasis on internal integration. The more underground and sealed off a group becomes, finally, the less effectively it influences wider society, because of its isolation.

The toolbox approach informs all of della Porta's book. Throughout, theory is developed and sharpened via detailed empirical engagement with her Italy and Germany case studies. One fruitful new category of hers is the *policing of protests* (1995, 56). Policing tactics serve as a downstream proxy for deeper state factors and institutional features, such as police organization, the nature of the judiciary, law codes, and constitutional rights. This simplifies the empirical tracking of the state's influence on the (already theorized) 'political opportunity structure', because the connection between policing and social movements is conveniently direct. It also enables policing itself to be analyzed in a subtler way than before. Policing tactics became more hardline often not because of internal dynamics within the police but rather because of external political decisions (which the police tried to resist), contrary to much previous theory (1995, 77-8). Other political explanations too are revealed or supported. Examples include (1995, 76-8): how hardline state and police attitudes rose and fell with the attitude and strength of the moderate 'old left'; how political polarization strengthened the hand of hardliners on both sides; how, in the long run, hardliners declined in influence; and how the tactics of the protestors influenced the tactics of the police.

[21] Della Porta employs *mixed methods* in a similar spirit. Individual actors' life histories, in the form of qualitative analysis of interviews, form part of her evidence base; she uses various quantitative data too, for instance about the number of acts of violence; in addition, she uses archival research, such as consultation of official records. These various forms of evidence each pay their way by supporting particular causal inferences in the service of della Porta's larger explanatory ambitions.

This rich explanatory detail is made visible by Della Porta's theoretical innovations. Her theories are not formal models; rather, they are qualitative and verbal. Her use of these theories is often heuristicist rather than causalist, bringing into view new categories or ways of seeing things rather than specifying causal hypotheses directly.[22]

Della Porta is explicitly against the possibility of universal theory or wide-scope explanation (1995, 210). 'Political violence' and 'radicalism' enter into causal relations that are fragile. These causal relations vary with: leftist versus rightist protest movements; democratic versus authoritarian political environments; class versus ethnic bases; and different organizational models, forms of action, and ideologies and goals. Large-n studies would fatally gloss over these heterogeneities, and so would miss many causal explanations. The implicit aim of such large-n studies, namely to confirm or discover a wide-scope causal generalization, is futile in this domain (1995, 14-20). Della Porta takes her own explanations to apply only to leftist, class-based groups in a democratic environment. At the end of her book, she cautiously examines how well these explanations might transfer to the case of the ethnically based civil rights movement in 1960s and 1970s USA. Again, detailed empirical engagement is the only way to tell (1995, 210-15).

10. Policy advice

Policy advice is inevitably (in part) forward-looking. Therefore, it requires forward-looking prediction, yet in field sciences forward-looking prediction is difficult. Does this imply a counsel of despair? Not always.

To warrant an intervention, we require a confirmed causal model. There are two routes to that. The first route is induction: perhaps the most convincing warrant for forward-looking predictive confidence is past forward-looking predictive success, combined with confidence that the context is sufficiently stable. At the macro level, such as predicting civil wars, the predictive record in political science is disappointing (Tetlock 2005, Ward et al 2010). This mandates caution, except when there is a record of success.

The second route to predictive confidence is local knowledge. Combined with relevant background knowledge, local knowledge can warrant forward-looking predictions when it is detailed enough to establish that there are few significant unmodelled causes, that there is sufficient causal stability, and that outcomes are predictable at all in the sense of not being too sensitive to unknowable details (Sterelny 2016).[23] This route obviously favors local-level predictions. For example, local knowledge may warrant a confident prediction of who will win an election in a particular new district, even if at the national level the election result is in doubt.

[22] "Recent studies on social movements *provide the main categories* for the explanatory model of political violence in Italy and Germany that I am going to develop here." (Della Porta 1995, 9, emphasis added)

[23] There is principled reason to expect predictability to require local knowledge. The literature on the extrapolation of a model from one context to another, including from the past to the future, concurs that such extrapolation requires detailed knowledge of the target context (Cartwright 2019, Khosrowi 2019, Cartwright and Hardie 2012).

Generally, we should expect warranted interventions usually to be narrow-scope. This is because they require confirmed causal knowledge, which we should expect usually to be narrow-scope for the same reasons that causal explanations are.

11. Conclusion

There is hope, but only via empirical success. Usually, that means via contextual history or via forward-looking prediction. Usually, it also means narrow-scope explanations, and not seeking to confirm wide-scope explanations via retrospective prediction. Warrant for policy interventions will usually be narrow-scope too.

References

Alexandrova, A. (2008). 'Making Models Count', *Philosophy of Science* 75, 383-404.

Alexandrova, A., and R. Northcott (2009). 'Progress in Economics', in D. Ross and H. Kincaid (eds) *Oxford Handbook of Philosophy of Economics*, 306-337. Oxford.

Card, D., and A. Krueger (1994). 'Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania', *American Economic Review* 84, 772–93.

Cartwright, N. (1989). *Nature's Capacities and their Measurement*. Oxford: Oxford University Press.

Cartwright, N. (1999). *The Dappled World*. Cambridge: Cambridge University Press.

Cartwright, N., and J. Hardie (2012). *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford.

Cartwright, N. (2019). *Nature, the Artful Modeler*. Open Court.

della Porta, D. (1995). *Social Movements, Political Violence, and the State*. Cambridge.

Dowding, K. (2016). *The Philosophy and Methods of Political Science*. Basingstoke: Palgrave Macmillan.

Dowding, K., and C. Miller (2019). 'On prediction in political science', *European Journal of Political Research*, online February 2019.

Dupré, J. (2012). *Processes of Life*. Oxford: Oxford University Press.

Elliott-Graves, A. (2018). 'Generality and Causal Interdependence in Ecology', *Philosophy of Science* 85.1: 1102-1114.

Farmer, D. (2013). 'Hypotheses non fingo: Problems with the scientific method in economics', *Journal of Economic Methodology* 20.4: 377-385.

Howson, C., and P. Urbach (1993). *Scientific Reasoning: The Bayesian Approach* (2nd edn). Open Court.

Khosrowi, D. (2019). *Extrapolating Policy Effects*. PhD dissertation, University of Durham.

Levitt, S., and J. List (2007). 'What do laboratory experiments measuring social preferences reveal about the real world?', *Journal of Economic Perspectives* 21, 153-74.

Northcott, R. (2012). 'Partial explanations in social science', chapter 7 in H. Kincaid (ed) *Oxford Handbook of Philosophy of Social Science*, 130-153. Oxford.

Northcott, R. (2013). 'Degree of explanation', *Synthese* 190.15, 3087-3105.

Northcott, R. (2015). 'Opinion polling and election predictions', *Philosophy of Science* 82, 1260-1271.

Northcott, R. (2017). 'When are purely predictive models best?', *Disputatio* 9.47, 631-656.

Northcott, R. (2018). 'The efficiency question in economics', *Philosophy of Science* 85.5, 1140-1151.

Northcott, R. (2019a). 'Prediction versus accommodation in economics', *Journal of Economic Methodology* 26.1, 59-69.

Northcott, R. (2019b). 'Big data and prediction: four case studies', *Studies in History and Philosophy of Science*, online September 2019.

Northcott, R., and A. Alexandrova (2015). 'Prisoner's Dilemma Doesn't Explain Much', in M. Peterson (ed) *The Prisoner's Dilemma*, 64-84. Cambridge: Cambridge University Press.

Pietsch, W. (2016). 'The Causal Nature of Modeling with Big Data', *Philosophy and Technology* 29, 137-171.

Popper, K. (1989). *Conjecture and Refutations*, revised edition. London: Routledge.

Reiss, J. (2008). *Error in Economics: Towards a More Evidence-Based Methodology*. Routledge.

Sagoff, M. (2016). 'Are there general causal forces in ecology?' *Synthese* 193, 3003-3024.

Salmon, W. (1981). 'Rational prediction', British Journal for the Philosophy of Science 32.2, 115-125.

Steel, D. (2008). *Across the Boundaries* (New York: Oxford University Press).

Sterelny, K. (2016). 'Contingency and history', *Philosophy of Science* 83.4, 521-539.

Tetlock, P. (2005). *Expert political judgment*. Princeton: Princeton University Press.

Ward, M., Greenhill, B., and K. Bakke (2010). 'The perils of policy by p-value: predicting civil conflicts', *Journal of Peace Research* 47.4, 363-375.

Watkins, J. (1968). 'Non-inductive corroboration', in I. Lakatos (ed) *The Problem of Inductive Logic*, 61-6. Amsterdam: North-Holland Publishing Co.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford.

Worrall, J. (2014). 'Prediction and accommodation revisited', *Studies in History and Philosophy of Science* 45, 54-61.

Ylikoski, P. (2019). 'Mechanism-based theorizing and generalization from case studies', *Studies in History and Philosophy of Science* 78, 14-22.