Uses and Abuses of Formal Models in Political Science

Jack Paine and Scott A. Tyson

The Use of Models

Formal political theory is a methodological approach—common in domestic politics, comparative politics, and international relations—that is characterized by its strong commitment to logical rigor as well as its conceptual and analytical clarity.[1] One of formal political theory's core strengths is that it confronts foundational questions about politics. For instance, who shapes policy? what strategies do they use? and what informational and incentive constraints affect political interactions? Pioneering insights from formal political theory to each of the subfields go beyond these basic questions to precisely articulate the mechanisms responsible for the political outcomes we observe, often by untangling countervailing effects and isolating clear counterfactual comparisons.

A formal political theory is usually comprised of at least two components. First, a logical (often mathematical) structure representing the critical individuals, decisions, constraints, and information that make up the substantive question. Second, and just as important, an interpretation of that logical structure that gives substantive meaning to the aspects and results of the model. These two components are critical (Rubinstein 2012), and they also introduce a flexibility in the questions formal political theory can address. The diverse ways in which formal political theory can contribute to understanding politics has also engendered considerable disagreement about how scholars can most productively use formal models as an analytical tool. In fact, there is a great deal of disagreement among formal theorists regarding what qualities make for a good (or insightful) model, the relationship between theory and empirical work, and what kinds of questions formal models are most appropriate for answering.

In this chapter, we present a novel distinction between two common approaches to formal models in political science. First is the *phenomenon* perspective, which seeks to relate a formal model to descriptive empirical patterns, and the second is the *experimental* perspective, which views formal models as an explication of a causal mechanism.[2] To illustrate the strengths of each of these perspectives (relative to the other), we consider the typical concerns a theorist confronts when developing a formal model from each perspective. We focus in particular on how each perspective approaches a *comparative static comparison*, which examines a comparison from changing one factor, while all other factors remain "static." A comparative static analysis focuses on an "all else equal" comparison by changing a single factor, holding all other aspects of the model fixed, and looking at the change in some outcome (perhaps simply equilibrium strategies).

An ideal model from the phenomenon perspective addresses three empirical considerations. First, what patterns in the real world motivate the need for a formal model? Second, do real-world actors perceive tradeoffs that correspond with key assumptions in the model setup? Third, do the model's comparative static predictions match empirical relationships? Although phenomenon-driven

---

[1] We interchangeably refer to "formal political theory," "game theory," and "formal theory."

[2] See Cox (1990) for a similar distinction applied to statistical models.

models are not realistic in the sense of providing a literal description of the real world, the setup and implications of these models do attempt to match attributes of the real world.

Many approaches to model construction in political science draw elements from the phenomenon approach, whether they espouse combining models with quantitative evidence (Morton 1999; Granato and Scioli 2004), qualitative evidence (Bates, Greif, Levi, Rosenthal and Weingast 1999; Goemans and Spaniel 2016; Lorentzen, Fravel, and Paine 2017), or a combination (Laitin 2003). Furthermore, in practice, many scholars attempt to provide insight into real-world phenomena when writing models, therefore implicitly adopting at least some elements of the phenomenon approach. Lorentzen, Fravel, and Paine (2017) surveyed every game theory article in six prominent political science journals between 2006 and 2013 that examined topics in international relations or comparative politics. They found that of the 182 articles, 128 (70%) included either a quantitative or a qualitative empirical component. The extent of this evidence differs from article to article, ranging from brief anecdotes in the introduction, to regression analysis of experimental or other originally collected data, and detailed case studies. But even sparse discussions of empirical evidence aim to convince the reader that aspects of the model are "realistic", and descriptively reflect substantive cases.

The real world is messy and complicated, and sometimes the best approach to understanding how it works is to analyze things in isolation. Because there are always substantive features which, although known to be important real-world considerations, are nevertheless superfluous for explaining the core political mechanism. This observation motivates the experimental approach to writing a formal model, which focuses on isolating and understanding substantive mechanisms. Ideally, an experimental-driven model is intentionally parsimonious because the priority is on viewing a particular causal mechanism in isolation. Consequently, introducing extraneous features into the model, although more descriptively realistic, is counterproductive because either such features add no additional insights, or worse, they create confusion. Instead, the more focused the model, the more focused the comparison, and the more general the insight (Banks 1990).

Comparing the experimental approach to formal political theory to actual experimental design highlights its goals and virtues (Ashworth, Berry, and De Mesquita 2015; Haavelmo 1944). The classic setup of an experiment considers different levels of a "treatment" and compares average outcomes between a treatment group and control group. Holding all else equal is precisely the goal of models from the experimental perspective, and consequently, there is less concern with accounting for the full panoply of substantive factors because—from the theorist's perspective— these additional things are not a critical part of the analysis.[3] A key strength of this approach is that by focusing on a particular mechanism, the analysis can reveal and understand the nuts and bolts of a substantive case, regardless of whether the mechanism of interest actually operates in isolation in the real world.

A particularly important component of comparative static analysis from the experimental perspective concerns *indirect* effects. Changing a single parameter can affect an outcome of interest through direct and informational channels. For instance, to understand the influence of political mobilization on government policies (through voting, protest, or other means), scholars generally study two effects. First is a direct effect: mobilized dissent can create various problems that a

---

[3] In experiments, all else equal is accomplished by randomization of treatment assignment (Kempthorne 1977; Rosenbaum 2017).

government is forced to deal with regardless of the reason for the disruption to society. Second, mobilized dissent is generally considered to communicate dissatisfaction among members of the citizenry with the government's policies. This leads to a conceptually distinct, informational, channel through which mobilized dissent influences government policy. From the phenomenon perspective, indirect effects can be a nuisance because they obstruct clean directional predictions from the model. However, from the experimental perspective indirect effects are often the most interesting aspect of the model, because they demonstrate the character and importance of the strategic considerations.

Below we provide numerous examples of the phenomenon and experimental approaches in applied research, distinguished by model motivation, setup, and comparative statics. We then discuss common critiques of formal models based on empirical applicability or lack thereof, and illustrate the differences in how the two approaches handle critiques. We discuss two influential debates. First, redistributive political transition models posit that economic inequality affects prospects for democratization by affecting demands for redistribution (Acemoglu and Robinson 2000, 2001, 2006; Boix 2003). Second, selectorate theory examines how institutional variation in leadership selection affects a range of outcomes, including public good provision and international war (Bueno de Mesquita, Smith, Morrow and Siverson 2005). We conclude with implications for research and training. Specifically, we emphasize how graduate game theory courses, by incorporating crucial philosophical and conceptual issues, could demonstrate how models can address substantively interesting questions in addition to teaching the technical structure of models.

Table 1 summarizes the defining elements of both approaches. Importantly, these approaches are not mutually exclusive, and most published formal modeling articles contain elements of each. However, explicating this distinction is critical for understanding how to use formal models to advance knowledge of political phenomena, and how to avoid common critiques that may be pertinent to one approach but not the other. Our conceptual distinction between different perspectives has largely been overlooked and is useful for all political scientists who might otherwise neglect some contributions of formal models.

Table 1: Key Differences Between Approaches

	Phenomenon	Experimental
Motivation	Explain descriptive patterns	Isolate mechanisms
Model setup	Assumptions should correspond with tradeoffs perceived by real-world decision-makers	Assumptions should be parsimonious to yield conceptual clarity
Comparative statics	Sign of key comparative static predictions (usually the total effect) should match statistical relationship or actions/outcomes in empirical cases	Comparative statics are used to isolate substantive channels

The Phenomenon Perspective

Motivation

The phenomenon perspective is often motivated by empirical patterns or a set of observations, which can be either quantitative or qualitative, that existing research does not convincingly explain.

Sometimes, the researcher presents a single pattern that raises strategic questions. For example, Slantchev (2010, 357) asks a question about a particular case: "During the last days of September 1950, the U.S. administration faced a momentous decision about what to do in Korea: should American forces stop at the 38th parallel, as originally planned, or should they continue into North Korea, and turn the conflict from a war of liberation into a war of unification?" He then presents a model in which an optimal response to such dilemmas depends on the opponent's incentive to "feign weakness." Miller and Schofield (2003) demonstrate that U.S. states won by the Republican presidential candidate William McKinley in the 1896 election nearly perfectly corresponds with states won by the Democratic presidential candidate Al Gore in the 2000 election, motivating their model on how party agents can push platforms that over time yield party realignment. Acemoglu and Robinson's (2006) book begins with narratives from Britain, Argentina, Singapore, and South Africa to highlight four regime trajectories that differ on whether democratization occurs and its stability. Their model explains how economic inequality shapes the equilibrium behavior of elites and the masses, which creates varying regime trajectories.

Other articles juxtapose disparate patterns and argue that they share a common strategic logic. For example, Powell (2012, 620) posits "three striking features or stylized facts about both interstate and civil war" based on quantitative and qualitative evidence in existing research: "(1) there are often periods of persistent fighting, (2) fighting commonly ends in negotiated settlements as well as in militarily decisive outcomes, and (3) fighting sometimes recurs." He argues that shifts in the distribution of power and actors' strategic fighting decisions (to forestall adverse shifts) yield equilibrium behavior consistent with all three patterns. Paine (2018) contains a section before the model setup that presents regression tables to highlight a mixed empirical pattern: higher country-level oil production covaries with less frequent center-seeking civil wars, whereas higher regional-level oil production covaries with more frequent separatist civil wars. The model highlights two main countervailing effects of oil production on the likelihood of civil war onset, and explains why these mechanisms vary in magnitude depending on the opposition's optimal civil war aims.

Model Setup

To explain empirical phenomena, the model setup should incorporate important tradeoffs that real-world actors perceive when making choices. Although explicitly motivating assumptions using real-world examples is somewhat less common than motivating examples or testing comparative statics predictions, Lorentzen, Fravel, and Paine's (2017) survey shows that 23% of game theory articles in their sample contained explicit evidence for assumptions. For example, Svolik (2009) studies an interaction between a dictator that seeks to concentrate power and a ruling coalition that attempts to maintain a power-sharing arrangement. He assumes that the dictator's strategic action to concentrate power sends an informative (but imperfect) signal to the ruling coalition, who may react by staging a coup. Svolik demonstrates the empirical relevance of this assumption by providing examples in which leaders' attempts to consolidate power generated observable signals to ruling coalition members. In the Soviet Union, Lavrenty Beria merged formal ministries after Josef Stalin's death to concentrate power in his hands. In Iraq, Saddam Hussein gradually replaced earlier supporters with loyalists from his hometown. In these cases, subordinates gained information that was consistent with attempted power concentration, but they were unsure of the true motives of the dictator—which corresponds with the core assumptions of Svolik's (2009) model.

The motivating puzzle in Nalepa (2010) is that in the late 1980s, many communist regimes in Eastern Europe negotiated democratic transitions with the opposition. Gaining assurances that

communist agents would not face punishment following a regime transition, they resigned peacefully in cases such as Poland, Hungary, and Czechoslovakia. This is puzzling when considering that the communists should seemingly have expected the opposition to break these promises. But, empirically, the new democratic leaders held their promises, which is also puzzling given widespread desire to punish the communists. Nalepa (2010) studies a signaling model and explains that these promises were credible because of communists' private information about transgressions committed by the opposition as informants during communism—i.e., their ``skeletons in the closet.'' But this mechanism is only empirically relevant if the real-world actors did indeed perceive this information asymmetry, which she confirms using evidence from interviews. For example: "The communists attempted to exploit this informational advantage by trying to convince the opposition that it was highly infiltrated. One of the dissidents representing Solidarity in the roundtable negotiations recalled: 'When I met Kwasniewski, he said, 'Do not mess with those files, let them be—the agents were mostly your own people''' (349-350).[4]

Comparative Statics

Whatever the initial motivation for presenting and setting up a formal model, the analysis generates comparative static predictions that researchers can evaluate either with statistical or qualitative evidence. This is a central element of the influential "Empirical Implications of Theoretical Models (EITM)" approach to political game theory (Signorino 1999; Morton 1999; Signorino and Yilmaz 2003; Granato and Scioli 2004) and also receives support from methodological research on combining game theory and qualitative methods (Bates et al. 1999; Goemans and Spaniel 2016; Lorentzen, Fravel, and Paine 2017). Lorentzen et al.'s (2017) aforementioned survey of game theory articles in political science shows that 63% of game theory articles provided either statistical tests, cross-case comparisons, or case studies to evaluate comparative static predictions.

For example, Conrad and Ritter (2013) examine the effects of international human rights treaties on incentives for domestic leaders to exercise repression. First, these treaties increase the likelihood of domestic protests in reaction to repression, increasing the need to exercise repression to retain power. Second, international human rights treaties increase the probability that repressive rulers will face litigation, which increases the costs of repression. Their formal analysis shows that the magnitude of the first effect depends on other aspects of the leader's job security. The first mechanism is relatively small in magnitude for secure leaders because they are unlikely to experience mass unrest regardless of the presence of an international treaty. However, the first mechanism is large in magnitude if the ruler is insecure, and dominates the second mechanism. This analysis yields a clear implication about an empirically observable interaction effect. Conrad and Ritter provide regression evidence that international human rights treaties are uncorrelated with repressive behavior in states with insecure leaders, but covary with lower repression in states governed by secure leaders.

As another example, Paine (2016) examines two countervailing implications of oil production: it raises the value of capturing the state for a rebel group, but it also increases government revenues to spend on patronage distribution and coercion. Untangling these distinct effects yields an implication about conventional practice in the empirical conflict literature. Standard conflict models include both oil production and income per capita on the right-hand side of the regression, and usually find that more oil production covaries with higher civil war frequency. The motivation

---

[4] Less frequently, scholars motivate key model assumptions using statistical evidence (e.g., Paine 2018).

for controlling for income per capita is that this is a strong predictor of civil war onset. However, the logic of the model highlights the problem with this control variable, which many argue proxies for government revenues. By controlling for income, the regression implicitly answers the largely irrelevant question of what the effect of discovering oil in countries like Saudi Arabia would have been if discovering oil did not increase government revenues. Revised regression specifications that incorporate this consideration demonstrate empirical results inconsistent with conventional wisdom about a conflict resource curse.

As an example of using qualitative evidence, Dunning (2008) highlights a set of conditions where resource wealth can promote democratic stability. High rents enable the government to provide public goods to the masses without needing to soak wealthy elites for tax revenues—mitigating class conflicts that would otherwise arise under a democratic regime. Using evidence from Venezuela, he shows that when oil rents were high in the 1970s, elites did not object to the high levels of public benefits provided to the masses because these public goods did not require high taxation (163-166). By contrast, as oil rents fell, Dunning (2008) shows that politics became polarized around classes and redistributive conflicts and ultimately facilitated the rise of the populist Hugo Chavez (166-183).

These examples also highlight the value-added of the *formal* analysis for deriving empirically testable comparative statics. In all three examples, the model analysis highlights two countervailing effects of a particular stimulus. The formal model facilitated rigorously examining the interaction between the two mechanisms and the conditions in which one should dominate the other. In each case, the analysis yielded novel empirical predictions that the researcher could take to data and check for directional congruence.

The Experimental Perspective

Motivation

The goal of experimental-driven models is to study specific attributes of strategic tradeoffs, such as individual motivations, information frictions, and other strategic issues that shape politics. For instance, institutional constraints like voting rules, the timing of elections, or the rules determining how legislation must be proposed dramatically influence various aspects of democracy (Diermeier and Krehbiel 2003; Dewan and Shepsle 2011). Other examples include how political accountability differs from standard contracting problems (Ashworth 2012), and the importance of communication in bureaucracy (Gailmard and Patty 2012).

As another example, Di Lonardo and Tyson (2018) study the interaction of domestic political threats and the logic of deterrence, which they approach from an experimental perspective. In particular, they first present the baseline crisis bargaining model of Fearon (1994) and Schultz (1998), and use this model to formally articulate the conventional logic of deterrence. Then, they introduce domestic political threats into this framework similar to Bueno de Mesquita et al. (2005) and Baliga, Lucca and Sjostrom (2011): domestic support is necessary for a leader to keep power. To isolate the effect of domestic political threats on the logic of deterrence, it is important when adding domestic politics *to hold constant* all other aspects from the benchmark model. In addition, although the benchmark crisis bargaining model suffers from some shortcomings, the contribution of Di Lonardo and Tyson (2018) would not be clear had they started with a nonstandard benchmark model of an international crisis.

Model Setup

The goal of experimental-driven models is not to attempt to approximate the real world, but instead to only include in the model elements needed to elucidate the core mechanism. For example, Tyson (2018) studies a central problem with exercising repression in authoritarian regimes: the dictator requires the cooperation of her security apparatus. However, the very need for a security apparatus creates an agency problem: if the leader loses power, then she cannot completely fulfill promises made to members of the repressive apparatus. Tyson (2018) explicitly removes other agency problems from the model, like moral hazard and adverse selection, even though such features are unarguably present in reality. Tyson (2018) does not include these aspects in his model in order to study implications resulting *exclusively* from the agency problem that arises from the leader's tenuous hold on power.

As another example, Banks and Duggan (2006) study determinants of public policy in legislatures with majority rules. They adopt a bargaining approach that assumes different members of the legislatures interact over time, and each can be randomly selected to make a policy proposal. If a majority adopts a proposed policy, then it becomes the new policy. By contrast, if a majority rejects a proposal, then the status quo remains in place. The goal of the model is to examine the implications of changing *one* key assumption from existing models: each legislator prefers any settlement to the status quo policy, i.e., the status quo is necessarily bad. This change implies that legislators may view the status quo policy favorably, making legislators more reluctant to vote for a new policy. Although real-world legislatures contain many additional features that Banks and Duggan (2006) do not incorporate into their model, making the setup more realistic would distract from their goal of changing a single substantive feature from existing models.

Comparative Statics

The purpose of comparative static exercises in experimental-driven models is to highlight the distinct channels through which a single factor causes a change in an outcome of interest, including equilibrium actions or their substantively relevant consequences. In most cases, there are numerous channels that correspond to separate mechanisms. The primary goal of the experimental approach is to elucidate each mechanism.

As a canonical example, suppose that different values of a treatment are represented by different values of x that directly influence an outcome, but may also provide information to decisionmakers, i.e., by changing their beliefs. In this case, the substantive outcome of interest, $Y(x,\beta)$, depends on x and beliefs, β. Supposing that everything is differentiable, then the total derivative with respect to x equals the sum of the direct and informational effects:

$$\frac{dY}{dx} = \frac{\partial Y}{\partial x} + \frac{\partial Y}{\partial \beta}\frac{d\beta}{dx}$$

The first term reflects the direct influence of x on Y. The second term combines the direct effect of beliefs on the outcome and the effect of x on beliefs. These distinct effects may pose a nuisance if the goal is to yield a predicted relationship between x to Y to take to the data. However, from the experimental perspective, the goal of the model is to *untangle these distinct mechanisms*. Sometimes these results are counterintuitive and "surprising" from the perspective of existing theories, and this is a key strength of models emerging from the experimental perspective.

For an experimental-driven formal theorist, perhaps the most interesting aspect of a comparative static relationship is the indirect effects that arise as a result of the strategic context. To highlight the importance of informational effects, consider for example, the classic jury problem, where an important substantive question regards whether jury verdicts reflect people's sincere opinions gathered from the facts of the case (Austen-Smith and Banks 1996; Feddersen and Pesendorfer 1998; Persico 2004).

To clarify this point, suppose there are N jurors and two collective outcomes, guilty (G) and innocent (I). Suppose also that convicting (i.e., choosing G) requires unanimity. There are also two equally likely states of the world: the defendant is truly guilty, or she is truly innocent, represented by $\omega \in \{G, I\}$, respectively. Jurors want to convict the guilty and to acquit the innocent, and their payoffs are represented by

$$u(G,G) = u(I,I) = 1 \text{ and } u(G,I) = u(I,G) = 0.$$

Each juror attends the trial, but despite their common preferences, each interprets the evidence and arguments slightly differently. To capture this, each juror receives an informative signal, where guilty signals are more likely to be seen when the defendant is guilty, and innocent signals are more likely to be seen when the defendant is innocent. Formally, juror i receives a signal s_i that equals either G or I, and $\Pr(s_i = \omega | \omega) = q \in (0.5, 1]$.

Will all jurors vote sincerely in line with their signal? Consider the problem from the perspective of an individual juror, who truly wants to convict only the guilty and to acquit only the innocent. Imagine this juror has seen a signal suggesting that the defendant is innocent. However, she also knows there is some probability that her signal is wrong and the defendant is guilty, and moreover, other jurors' signals may differ from hers.

The juror in this example is driven by an informational concern. There is a direct effect that follows from her signal, namely, an innocent signal suggests that the defendant is innocent. However, there is an important indirect effect that follows from the structure of the jury problem, namely, the voting rule. Specifically, a juror considering whether her vote is pivotal in the ultimate verdict, and who is considering voting to acquit, knows that the only case in which her vote will make a difference is when all other voters have cast guilty votes. But if all these voters have voted sincerely, it means that they have all received guilty signals—an extremely unlikely event when the defendant is in fact innocent. Consequently, it is not a best response for the juror to vote sincerely. More broadly, this example illustrates how indirect informational concerns influence decisions in political contexts.

The Abuse of Models

Most formal political theory articles contain elements of both the phenomenon and experimental approaches. A formal political theory formulated from one perspective is motivated by a distinct set of concerns that the other perspective does not necessarily share—nor should it. But because the distinction between phenomenon and experimental kinds of models has not been articulated previously, concerns that are important ingredients from one perspective are often unintentionally used to obstruct the other. For example, an experimental-driven model, on the surface, appears to be far more stylized than one written from the phenomenon perspective. However, it is important to stress that this superficial kind of "artificiality" is intentional, and constitutes one of the key

strengths of this theoretical approach. The experimental theorist is driven not by a desire to include as many factors as possible, but instead, needs to ensure that mitigating influences, with respect to the main factor of interest, are suppressed. To accomplish this theoretically, the theorist intentionally omits factors, even though they might be important in the real world. These omitted factors are precisely the things that an empiricist controls for, but for a formal model to keep such things fixed, the theorist *must omit* them from the model.

A common critique of game theoretic models is that their implications are unimportant because they rest on unrealistic assumptions (for example, Elster 2000; Green and Shapiro 1996).[5] To illustrate the difference between the phenomenon and experimental approaches, consider how a theorist from each perspective might respond to this criticism. A phenomenon-driven theorist should respond by modifying the assumptions to better reflect reality, whereas the experimental-driven theorist would allege that such a complaint reflects a misunderstanding of the question their model was designed to address.

Scholars have also debated the role and importance of empirical evidence in validating a model's predictions. On the one extreme, the *American Journal of Political Science* proposed briefly in the early 2000s a submission policy in which the editors would desk-reject any formal modeling manuscript that lacked an accompanying empirical test (Hill 2005). On the other extreme, Clarke and Primo (2012) argue that empirically testing models misunderstands their purpose. Instead, they argue that the only purpose of models is what we call the experimental approach. With regard to this controversial debate, the difference between the phenomenon and experimental approaches to formal political theory is crucial for understanding the source and relevance of these different points of view. Confusing philosophical positions with quality judgments tends to obscure the discussion, leading scholars to talk past each regarding thingz that are largely orthogonal to substantive issues. To illustrate our point, we present two examples from prominent models that exemplify these distinctions.

Redistributive Political Transitions

The idea that inequality and prospects for economic redistribution affect incentives to seek or to resist democratization has a long pedigree in political science. More recently, Acemoglu and Robinson (2000, 2001, 2006) present a parsimonious formal framework to explain these incentives, where a commitment problem is the key mechanism. Acemoglu and Robinson's (2006) core model analyzes an interaction between a representative rich elite that sets policy under a dictatorship, and a representative agent of the poor masses that sets policy under democracy. Each actor seeks to maximize its own consumption by affecting the tax rate. Because of the assumed wealth disparity, elites prefer no taxes whereas the masses prefer a positive tax rate. Furthermore, economic inequality determines the extent to which the two actors disagree about taxes, as higher inequality causes the masses to prefer a higher tax rate. Although the elite unilaterally determines the tax rate under dictatorship (de jure power), the masses may be able to force higher tax rates by

---

[5] The inherent complexity of the social world requires imposing some simplifying assumptions to construct a model of political behavior, and thus, all models simplify, formal or not (Clarke and Primo 2012). Friedman (1953) presents an extreme view that models should be assessed solely for their predictive ability, and that the assumptions that generate these predictions are entirely unimportant. On the other end of the spectrum, Bates et al. (1999, 14) argue that "the assumptions [should] fit the facts" for a model to have empirical applicability, which is perhaps also too extreme.

staging a revolution (de facto power). The elite has three options to stave off revolution: temporary concessions, repression, or democratization.

One key mechanism that Acemoglu and Robinson's (2006) model elucidates is the effect of economic inequality on the likelihood of democratization. They derive a non-monotonic relationship in which democratization only occurs if inequality is intermediate. At low levels of inequality, there is low demand by the masses for democracy because the amount of wealth held by elites that the masses could redistribute to themselves in democracy is low. At high levels of inequality, democratization does not occur because the elites use repression instead. The amount of redistribution under democracy would be so high that elites prefer to use costly repression to retain power. However, if inequality is intermediate, then mass demand for democratization is high enough that negotiated concessions are insufficient to prevent revolution, but the elites' fate under democracy is not dire enough for them to use repression.

Subsequent research criticizes numerous assumptions of the model setup. Some scholars allege that class differences between rich and poor is usually not the primary political cleavage that drives political transitions (Epstein, Leventoglu, and O'Halloran 2012; Haggard and Kaufman 2012; Ansell and Samuels 2014). Others argue that, at least in the post-colonial world since 1945, economic elites do not usually exercise political control. For example, the military usually does not act as a proxy for the wealthy (Slater, Smith, and Nair 2014). Some posit that revolutionary threats rarely provide a stimulus for democratization and that other factors appear more important for explaining manhood suffrage in most European countries (Collier 1999; Lizzeri and Persico 2004; Llavador and Oxoby 2005), womanhood suffrage (Przeworski 2009), or internationally driven transitions in recent decades (Levitsky and Way 2010; Haggard and Kaufman 2012). Finally, many democracies do not redistribute en masse either because they lack infrastructural capacity (Slater, Smith, and Nair 2014), or because elites exert considerable influence even under democracy (Albertus and Menaldo 2018).[6]

Are these critiques relevant? From the phenomenon perspective, many of these are pertinent critiques that require a sustained theoretical and empirical dialogue. Given Acemoglu and Robinson's (2006) stated goal to explain empirical instances of democratic transitions, it is important for the model to incorporate key tradeoffs that real-world policy makers faced. Correspondingly, models written in response to these critiques have yielded numerous insights by altering aspects of the original setups to more closely capture particular empirical settings (Dower, Finkel, Gehlbach, and Nafziger 2018).

From an experimental perspective, these critiques are less relevant because the key contribution of Acemoglu and Robinson (2000, 2001, 2006) was to take existing non-formal theories of democratization to understand the strategic interaction among social classes.[7] Moreover, perhaps

---

[6] Others examine empirical contexts in which the core assumptions of Acemoglu and Robinson's original redistributive political transition theories exhibit greater empirical plausibility. Paine (2019b) argues that post-1945 European settler colonies in Africa fit the scope conditions of a rich and politically dominant European elite that fears the revolutionary potential of the non- European majority and demonstrates statistical evidence consistent with Acemoglu and Robinson's and Boix's prediction that high inequality should yield high repression and revolution.

[7] Some responses by the authors adopt a mechanism-based defense. Discussing the original model, Acemoglu, Naidu, Restrepo and Robinson (2013, 2, 16) state that "once one relaxed the simple poor versus rich nature of political conflict in their original models as well as the restriction of policy instruments, the nature of the comparative statics with respect to inequality in the basic model changed. Put simply, if the groups in conflict were not rich versus poor, but for example

the most important contribution of these models is in identifying how democratization can result from a commitment problem that arises when elites lose (even temporarily) de facto political power. The models of Acemoglu and Robinson (2000, 2001, 2006) also generate several counterintuitive predictions. For example, Acemoglu and Robinson show that if the masses can only mobilize infrequently to stage a revolution, then eventual democratization becomes *more* likely. This result follows because infrequent mobilization enhances the masses' bargaining power in periods they can organize for revolution, since their future valuation of the status quo regime is low. As another example, Boix (2003) shows that high inequality does not cause elites to resort to repression when asset liquidity is low. If elites can move their assets abroad, then they do not fear high taxes under democracy, hence highlighting a subtle mitigating effect in the inequality-democratization relationship. Furthermore, highlighting the value of mechanism-based contributions to spurring future research and empirical insights, Paine (2019*a*) extends the asset liquidity mechanism in a dynamic model to help explain the empirical relationship between oil production and separatist civil wars.

Selectorate Theory

The experimental perspective to formal political theory faces a different kind of criticism. Specifically, does the setup of the model, including the underlying assumptions, isolate clear causal mechanisms? Whereas ensuring that all relevant factors are incorporated into the model is a mark of quality from the phenomenon perspective, it is often a sign of conceptual confusion from the experimental perspective. Likewise, having a clean, streamlined, and focused model is ideal for the experimental approach, a scholar motivated by the phenomenon perspective typically has a skeptical view of such a model's conclusions.

As an illustration, consider the selectorate theory presented in Bueno de Mesquita, Morrow, Siverson, and Smith (1999) and Bueno de Mesquita et al. (2005). A simple observation motivates selectorate theory: every leader relies on the support of some specified set of individuals, called the selectorate, which is designated by a country's institutions. As a result of this, leaders cannot sustain their hold on power without adequately compensating their winning coalition, the proportion of the selectorate needed to keep them in office. When the selectorate is small, as in autocratic regimes, this is most effectively accomplished through providing private goods. By contrast, when the selectorate is large, as in democracies, this is most effectively accomplished through public good provision. Numerous implications follow from this core insight, including why democracies do not fight each other, which they confirm with numerous statistical tests.

Like redistributive political transition models, selectorate theory has attracted considerable criticism. Gallagher and Hanson (2015) critique three main aspects, all of which reflect a phenomenon perspective. First, in reality, there is no clear distinction among winning coalition members, selectorate members, and non-selectorate members. Second, existing measures of these concepts are flawed, rendering Bueno de Mesquita et al.'s (2005) statistical tests invalid.[8] Third, Gallagher and Hanson (2015) criticize selectorate theory's core assumptions, arguing that the

---

based on ethnic, religious or regional cleavages, it was not necessarily true that increasing inequality, in the sense of a higher Gini coefficient, would exacerbate conflict between groups. It might just result in increased redistribution within groups."

[8] This critique also relates to Clarke and Stone's (2008) re-analysis of Bueno de Mesquita et al.'s (2005) data.

theory treats selectorate members as homogeneous, conflates rulers with regimes, and mischaracterizes the relationship between public goods and political rights.

Once again, the response to these critiques depends on one's philosophical perspective. From the phenomenon perspective, it is important to improve the descriptive accuracy of the assumptions and to incorporate additional elements into the original model. These considerations have motivated several extensions to the original model that include revolutions, purges, and other forms of authoritarian ruler turnover (Bueno de Mesquita and Smith 2009, 2017); the effects of natural disasters (Flores and Smith 2013); and leader health shocks (Bueno de Mesquita, Smith et al. 2018).

However, viewed from the experimental perspective, a deeper concern with the core selectorate theory model is that it may attempt to be *too* realistic. The baseline selectorate model presented in Bueno de Mesquita et al. (2005, Chs. 2 and 3) contains more than ten choice variables, plus a number of exogenous parameters and an infinite horizon. The core mechanism of the model, however, can be expressed more clearly by *removing* most of these moving pieces. For instance, Bueno de Mesquita (2016, Ch. 11) presents a simplified version of selectorate theory that isolates the effects of the core mechanism—winning coalition size—and shows how it affects public good provision and foreign policy aggression.

Implications for Research and Training

Many debates about specific formal models in political science, and the modeling enterprise more generally, draw from what we term the *phenomenon* and *experimental* approaches. But because scholars have not previously articulated these distinct perspectives, we often talk past each other—including those actively engaged in the formal theory enterprise and those who are not. Perhaps the most important takeaway from our discussion is that neither the phenomenon perspective nor the experimental perspective is inherently flawed. Instead, scholars often combine them effectively, if only implicitly, and insights from each has unique strengths that have improved the scholarly understanding of politics.

Importantly, the phenomenon and experimental approaches to formal models are not mutually exclusive, and most published models contribute to both approaches. However, most authors typically frame their contribution emphasizing one perspective over the other, which generally leads the overall contribution to be overlooked. Compared to the experimental approach, for many, the phenomenon approach is more intuitive when writing and thinking about models in political science because it more closely corresponds to historical and qualitative approaches. However, the experimental perspective has been gaining ground in all the social sciences—and political science is no exception.[9] Consequently, the experimental approach to formal political theory will become more useful as it more naturally connects to research designs focusing on causal relationships as well as lending insight into the issues that are at the heart of these empirical studies.

In addition to the direct implications for conducting and evaluating formal political theory research, the phenomenon and experimental distinction also carries important implications for future formal political theory training in graduate programs. Formal political theory's key strengths lie in its ability to bring conceptual clarity to substantive issues by transparently articulating the

---

[9] This movement gained substantial momentum following Leamer (1983).

relationships that drive broader scholarly debates. But introductory courses in formal political theory focus almost exclusively on "tools" or "skill-building," which has the unintended consequence of leaving some important philosophical and conceptual issues unaddressed. Of course, correctly solving a formal model is necessary, but it is not sufficient for making a contribution to political science using formal political theory. Instead, articulating the distinct virtues of the phenomenon and experimental approaches highlights the diverse contributions of formal political theory, and it is our hope that explicitly highlighting distinct philosophical perspectives to formal political theory can bring clarity into the general discussion.

References

Acemoglu, Daron and James A. Robinson. 2000. "Why did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4):1167–1199.

Acemoglu, Daron and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91(4):938–963.

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson. 2013. "Democracy, Public Policy, and Inequality." *Newsletter for the Comparative Democratization Section of the American Political Science Association* 11(3):2, 16–20.

Albertus, Michael and Victor Menaldo. 2018. *Authoritarianism and the Elite Origins of Democracy*. New York, NY: Cambridge University Press.

Ansell, Ben W. and David J. Samuels. 2014. *Inequality and Democratization: An Elite Competition Approach*. Cambridge, UK: Cambridge University Press.

Ashworth, Scott. 2012. "Electoral accountability: recent theoretical and empirical work." *Annual Review of Political Science* 15:183–201.

Ashworth, Scott, Christopher Berry and Ethan Bueno De Mesquita. 2015. "All Else Equal in Theory and Data (Big or Small)." *PS: Political Science and Politics* 48(1):89–94.

Austen-Smith, David and Jeffrey S Banks. 1996. "Information aggregation, rationality, and the Condorcet jury theorem." *American political science review* 90(1):34–45.

Baliga, Sandeep, David O Lucca and Tomas Sjostrom. 2011. "Domestic political survival and international conflict: is democracy good for peace?" *The Review of Economic Studies* 78(2):458–486.

Banks, Jeffrey S. 1990. "Equilibrium behavior in crisis bargaining games." *American Journal of Political Science* pp. 599–614.

Banks, Jeffrey S. and John Duggan. 2006. "A General Bargaining Model of Legislative Policy-Making." *Quarterly Journal of Political Science* 1(1):49–85.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal and Barry Weingast. 1999. "Analytic Narratives.".

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge University Press.

Bueno de Mesquita, Bruce and Alastair Smith. 2009. "Political Survival and Endogenous Institutional Change." *Comparative Political Studies* 42(2):167–197.

Bueno de Mesquita, Bruce and Alastair Smith. 2017. "Political Succession: A Model of Coups, Revolution, Purges, and Everyday Politics." *Journal of Conflict Resolution* 61(4):707–743.

Bueno de Mesquita, Bruce, Alastair Smith, James D Morrow and Randolph M Siverson. 2005. *The Logic of Political Survival*. MIT press.

Bueno de Mesquita, Bruce, Alastair Smith et al. 2018. "Political Loyalty and Leader Health." *Quarterly Journal of Political Science* 13(4):333–361.

Bueno de Mesquita, Bruce, James D Morrow, Randolph M Siverson and Alastair Smith. 1999. "An institutional explanation of the democratic peace." *American Political Science Review* 93(4):791–807.

Bueno de Mesquita, Ethan. 2016. *Political economy for public policy*. Princeton University Press.

Clarke, Kevin A. and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. Oxford University Press.

Clarke, Kevin A. and Randall W. Stone. 2008. "Democracy and the Logic of Political Survival." *American Political Science Review* 102(3):387–392.

Collier, Ruth Berins. 1999. *Paths Toward Democracy: The Working Class and Elites in Western Europe and South America*. Cambridge, UK: Cambridge University Press.

Conrad, Courtenay R. and Emily Hencken Ritter. 2013. "Treaties, Tenure, and Torture: The Conflicting Domestic Effects of International Law." *Journal of Politics* 75(2):397–409.

Cox, David R. 1990. "Role of models in statistical analysis." *Statistical Science* 5(2):169–174. Dewan, Torun and Kenneth A Shepsle. 2011. "Political economy models of elections." *Annual Review of Political Science* 14:311–330.

Di Lonardo, Livio and Scott A. Tyson. 2018. "Political Instability and the Failure of Deterrence." *Mimeo: University of Rochester* .

Diermeier, Daniel and Keith Krehbiel. 2003. "Institutionalism as a Methodology." *Journal of theoretical politics* 15(2):123–144.

Dower, Paul Castañeda, Evgeny Finkel, Scott Gehlbach, Steven Nafziger. 2018. "Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms." *American Political Science Review*, 112(1): 125–147.

Dunning, Thad. 2008. *Crude Democracy: Natural Resource Wealth and Political Regimes*. Cambridge, UK: Cambridge University Press.

Elster, Jon. 2000. "Rational Choice History: A Case of Excessive Ambition-Analytic Narratives." *American Political Science Review* 94(3):685–695.

Epstein, David, Bahar Leventoğlu and Sharyn O'Halloran. 2012. "Minorities and Democratization." *Economics & Politics* 24(3):259–278.

Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American political science review* 88(3):577–592.

Feddersen, Timothy and Wolfgang Pesendorfer. 1998. "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political science review* 92(1):23–35.

Flores, Alejandro Quiroz and Alastair Smith. 2013. "Leader survival and natural disasters." *British Journal of Political Science* 43(4):821–843.

Friedman, Milton. 1953. "The methodology of positive economics.".

Gailmard, Sean and John W Patty. 2012. "Formal models of bureaucracy." *Annual Review of Political Science* 15:353–377.

Gallagher, Mary E. and Jonathan K. Hanson. 2015. "Power Tool or Dull Blade? Selectorate Theory for Autocracies." *Annual Review of Political Science* 18.

Goemans, Hein and William Spaniel. 2016. "Multimethod Research: A Case for Formal Theory." *Security Studies* 25(1):25–33.

Granato, Jim and Frank Scioli. 2004. "Puzzles, Proverbs, and Omega Matrices: The Scientific and Social Significance of Empirical Implications of Theoretical Models (EITM)." *Perspectives on Politics* 2(2):313– 323.

Green, Donald and Ian Shapiro. 1996. *Pathologies of rational choice theory: A critique of applications in political science*. Yale University Press.

Haavelmo, Trygve. 1944. "The probability approach in econometrics." *Econometrica: Journal of the Econometric Society* pp. iii–115.

Haggard, Stephan and Robert R Kaufman. 2012. "Inequality and Regime Change: Democratic Transitions and the Stability of Democratic Rule." *American Political Science Review* 106(3):495–516.

Hill, Kim. 2005. "AJPS rejection without review for theory papers." Posted on H-POLMETH24 March 2005. Accessed 12/13/13.

Kempthorne, Oscar. 1977. "Why randomize?" *Journal of Statistical Planning and Inference* 1(1):1–25.

Laitin, David D. 2003. "The Perestroikan Challenge to Social Science." *Politics & Society* 31(1):163–184.

Leamer, Edward E. 1983. "Let's take the con out of econometrics." *The American Economic Review* 73(1):31–43.

Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes after the Cold War*. Cambridge University Press.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's "Age of Reform"." *Quarterly Journal of Economics* 119(2):707–765.

Llavador, Humberto and Robert J. Oxoby. 2005. "Partisan Competition, Growth, and the Franchise." *Quarterly Journal of Economics* 120(3):1155–1189.

Lorentzen, Peter, M. Taylor Fravel and Jack Paine. 2017. "Qualitative Investigation of Theoretical Models: The Value of Process Tracing." *Journal of Theoretical Politics* 29(3):467–491.

Miller, Gary and Norman Schofield. 2003. "Activists and Partisan Realignment in the United States." *American Political Science Review* 97(2):245–260.

Morton, Rebecca B. 1999. *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science*. New York, NY: Cambridge University Press.

Nalepa, Monika. 2010. "Captured Commitments: An Analytic Narrative of Transitions with Transitional Justice." *World Politics* 62(2):341–380.

Paine, Jack. 2016. "Rethinking the Conflict "Resource Curse": How Oil Wealth Prevents Center-Seeking Civil Wars." *International Organization* 70(4):727–761.

Paine, Jack. 2018. "A Theory of Strategic Civil War Aims: Explaining the Mixed Oil-Conflict Curse." Working paper, Department of Political Science, University of Rochester. Accessed 12/21/18.

Paine, Jack. 2019*a*. "Economic Grievances and Civil War: An Application to the Resource Curse.". *International Studies Quarterly*, forthcoming.

Paine, Jack. 2019*b*. "Redistributive Political Transitions: Minority Rule and Liberation Wars in Colonial Africa." *Journal of Politics*, forthcoming.

Persico, Nicola. 2004. "Committee design with endogenous information." *The Review of Economic Studies* 71(1):165–191.

Powell, Robert. 2012. "Persistent Fighting and Shifting Power." *American Journal of Political Science* 56(3):620–637.

Przeworski, Adam. 2009. "Conquered or Granted? A History of Suffrage Extensions." *British Journal of Political Science* 39(2):291–321.

Rosenbaum, Paul R. 2017. *Observation and experiment: an introduction to causal inference*. Harvard University Press.

Rubinstein, Ariel. 2012. *Lecture notes in microeconomic theory: the economic agent*. Princeton University Press.

Schultz, Kenneth A. 1998. "Domestic opposition and signaling in international crises." *American Political Science Review* 92(4):829–844.

Signorino, Curtis S. 1999. "Strategic interaction and the statistical analysis of international conflict." *Amer ican Political Science Review* 93(2):279–297.

Signorino, Curtis S and Kuzey Yilmaz. 2003. "Strategic misspecification in regression models." *American Journal of Political Science* 47(3):551–566.

Slantchev, Branislav L. 2010. "Feigning Weakness." *International Organization* 64(3):357–388.
Slater, Dan, Benjamin Smith and Gautam Nair. 2014. "Economic Origins of Democratic Breakdown? The

Redistributive Model and the Postcolonial State." *Perspectives on Politics* 12(2):353–374.
Svolik, Milan W. 2009. "Power Sharing and Leadership Dynamics in Authoritarian Regimes." *American Journal of Political Science* 53(2):477–494.

Tyson, Scott A. 2018. "The Agency Problem Underlying Repression." *The Journal of Politics* 80(4):1297– 1310.