The Art of Modeling[*]

John W. Patty[†]

January 22, 2025

Abstract

In this chapter, I provide some practical advice for how to build a model. The overall message is "keep it as simple as possible."

---

[*] I thank Scott de Marchi, Sean Gailmard, Jenny Kim, Skip Lupia, Maggie Penn, Ian Turner, and Scott Tyson for many useful conversations over the past years. This chapter is dedicated to the memory of Richard McKelvey, without whom this chapter would have been impossible to write.

[†] Departments of Political Science and Quantitative Theory & Methods, Emory University. Email: *jwpatty@gmail.com*.

1 Introduction

I have had many great experiences as an instructor in EITM. Perhaps my personal favorite is something I call "model-thon." For many years now, I have had the pleasure of building formal models for students at the workshop, in some cases doing so on the fly at the whiteboard in front of the students. I do not say that in a bragging way. I don't know if my models are any good — I really just love building models.

One of the values in doing this — and something I have found through my teaching in general — is that it has given me a better perspective on how to build a model *generally*. I like to think it also has increased some young scholars' appreciation of not only the value of "having a model," but also the relative ease with which one can get a simple model up and running. More provocatively, it has convinced me that the best models are the simplest models. It's not a particularly original recognition (essentially, another instantiation of the principle of Occam's razor), but I think it is an important one, especially for graduate students and untenured faculty facing real (and tighter than some realize) deadlines. It also mirrors similar advice my more empirically-focused friends and mentors have given me through the years: more complicated models (regardless of whether they're formal or statistical) are less robust than simpler ones, *ceteris paribus*, because you necessarily have to make more assumptions for the (again, theoretical or empirical) results to be *valid*.

2 The Value(s) of A Simple Model

"All models are wrong" is an old saw. I would modify it just a bit: all *useful* models are wrong. After all, any phenomenon X is a *perfect* model of X. The reason we build some other model of X, M, is to *simplify* X. Simplification has at least two related virtues. First, simplification clears "the weeds" that might be obscuring the answer to some question one has about X. Second, simplification helps us "solve" the model of X, M. I'll talk more about solving models below (Section 4) and instead focus on the first virtue of simplification in this section.

Simple Models are Robust. An implication of the first virtue of simplicity is the benefit of *robustness*. For example, suppose that I have a model of X, M_1, that contains two components: a and b, and suppose that another model, M_2, contains only component a. Suppose that some aspect of X, α, is explained by the smaller model, M_2. It follows that α can also be explained by model M_1. However, we also know that explaining α does not require component b. Model M_2 is more robust than model M_1 because its explanation requires fewer assumptions (*i.e.*, the explanation does not need any additional assumptions about b).

Models Explain Logic. Of course, there may be another aspect of X, β, that is explained by M_1, but not by model M_2. Many models can be nested like M_2 is "nested within" M_1. Now, suppose that you want to explain both aspects α & β. One good approach to presenting the analysis leverages the nestedness of M_1 and M_2. Here's a simple guide:

1. Describe phenomenon X and the relevant aspects, α and β.
2. Define model M_2 and analyze it, focusing only on aspect α.
3. Define model M_1 as an extension of model M_2 and explain to readers the additional assumptions required (*i.e.*, b).
4. Analyze model M_1, focusing only on aspect β.
5. Review the findings, noting again the relative explanatory roles of components α and β.

Consider the classic model of turnout due to Riker and Ordeshook (1968). The question motivating their study is a classic one: why do people turn out to vote? This can lead to several ancillary predictions, such as *who* turns out to vote, what *kind of elections* do they turn out to vote in, what kind of candidates do they turn out to vote for (or against), ..., and so forth. Regardless, central aspects of their model are:

1. An individual voter who cares about:
   (a) which candidate wins the election (valued at $B \geq 0$),
   (b) the instrumental cost of turning out to vote ($c \geq 0$), and
   (c) the intrinsic appeal of voting ($d \geq 0$), and
2. A probability that one's vote will be *decisive* ($p \in [0, 1]$).

The model is very simple and based on methodological individualism — the individual chooses whether to vote ($v = 1$) or not ($v = 0$) — and the individual's preferences are based on only his or her decision, v, and the winner of the election. The model's central conclusion is that the individual should vote whenever

$$p \cdot B + d - c > 0,$$

and *not* vote whenever

$$p \cdot B + d - c < 0.$$

Simple as this is, it transparently illustrates the logic of Riker & Ordeshook's argument: because $B \geq 0$ (by definition), individuals should vote in elections where their vote is more likely to be decisive (*i.e.*, for larger values of p), when the net reward from their most-preferred candidate is higher (*i.e.*, for larger values of B), when the cost of voting, c, is small, and/or when their intrinsic value of voting, d, large.

Before moving along, note that the theory has made a lot of simplifications. To name only three, the theory does not (explicitly) include the following:

- How many other people turned out to vote in the election?
- What are the platforms of the candidates running for office?
- How is the economy doing?

By not including these concepts within their theory, Riker & Ordeshook should not be read as saying that "these things don't matter" but, rather, that they are seeking to explain only one basic logic of an individual's turnout decision.[1]

One could extend the model by either including one or more of these factors explicitly within the model or, a little less obviously, by "modeling" these effects as affecting one or more of the four parameters (p, B, c, and d). The first approach raises at least two challenges. As mentioned above, the first challenge is that such an extension would be less robust than Riker & Ordeshook's model. More concrete is the second related challenge: how do these factors actually affect the individual's payoff from turning out to vote? There are infinite ways that one can include any of them. I discuss this issue in more detail in Section 3.3 and also in Section 7.

Models Connect Seemingly Disparate Phenomena. Another advantage of simplicity/robustness of a model is that its logic can "travel more widely." Some scholars might say that this amounts to simple arbitrage, with the potential implication being that doing so is less impressive than developing an entirely new and/or more complicated model. In my mind, this line of argument ignores that a model is simply a tool, and versatility of a given tool is generally considered a virtue.

There are several examples of versatile models in political science and economics. For example, the alternating offers bargaining framework of Rubinstein (1982) has been applied to phenomena as seemingly disparate as what kind of legislation will emerge in a legislature (*e.g.*, Baron and Ferejohn (1989)) and when nations will go to war (*e.g.*, Fearon (1995), Reiter (2003)), to name only two broad topics. Similarly, the cheap talk model of Crawford and Sobel (1982) has been applied to understand how legislatures organize themselves (*e.g.*, Gilligan and Krehbiel (1987)), what types of advisors can be credible to political executives (*e.g.*, Gailmard and Patty (2012, 2019*a*), Patty (2024)), and international conflict (*e.g.*, Ramsay (2011), Joseph (2021)). As a final

---

[1] Whether this means that Riker & Ordeshook believed that these factors *do not* — or *should not* — play a role in individual turnout decisions is another story, but beyond my scope here.

one of several other examples, the unidimensional model of spatial competition, due to Hotelling (1929), has been applied to understand what types of positions candidates adopt in elections (*e.g.*, Downs (1957)), what types of coalitions emerge in collective choice (*e.g.*, Axelrod (1970)), and when policy-making is stymied by institutions such as supermajority rule and veto powers (*e.g.*, Krehbiel (1998), Monroe, Patty and Penn (2018)).

Models *qua* Research Design. The versatility of simple models embodies another of their virtues. Specifically, comparing models is the starting point for statistical inference.[2] In layman's terms, discovering the simplest model that explains two separate phenomena most clearly illuminates their common foundations.

For example, returning to the discussion above of two nested models of phenomenon X above (*i.e.*, about models M_1 & M_2), suppose that another phenomenon Y is explained by both models M_1 and M_2. Then the *simpler* model (M_2) provides more insight into the connections between phenomenon X and Y than model M_1 does. In this sense (and others), every model of any phenomenon *is* a research design. The model excludes those factors that — at least for the purpose of explaining the aspect of the phenomenon of interest — should *not* be required for systematic explanation. Less aggressively, including a factor in the model acknowledges that it *might* have a systematic role in the offered explanation. To use linear regression as a reference point, just note that the a model that claims that factors a and b have an impact on some dependent variable y, but excludes factor c, is equivalent to a restricted model that "includes" factor c:

$$y = \gamma_0 + \gamma_a \cdot x_a + \gamma_b \cdot x_b + \varepsilon = \gamma_0 + \gamma_a \cdot x_a + \gamma_b \cdot x_b + 0 \cdot x_c + \varepsilon.$$

Models have another key connection to research design: *causality*. While the above discussion of the linear regression example avoided the question, note that in the simple regression model given by

$$y = \gamma_0 + \gamma_a \cdot x_a + \gamma_b \cdot x_b + \varepsilon,$$

if we assume that ε is independent of a and b (as typically done), then the factors a and b are each causal with respect to the expected value of y:

$$E[y \mid \gamma, x] = \gamma_0 + \gamma_a \cdot x_a + \gamma_b \cdot x_b + E[\varepsilon]. \tag{1}$$

It's important to pause here for a moment: the equality in Equation (1) *is* causality: if the sum on the left side of (1) increases (or decreases), then the right hand side increases (or, respectively,

---

[2] This is particularly true if the models are nested, but the concept can travel more broadly with a little (or, sometimes, a lot) of work. This is not coincidentally related to the difficulty of comparing the empirical performance of two or more statistical models of the same phenomenon.

decreases) as well.

But Correlation Isn't Causation! [3] My example above is meant to be a little provocative — Equation (1) looks like a regression equation, because it is. My potentially provocative point is that *any* equality is making a causal claim. Consider two individuals, a and b, perfectly balanced on a seesaw. If person a is launched off the seesaw, person b will fall to the ground with force and, similarly, if person b falls to the ground with force, then person a will be launched off the seesaw. Did a cause b to fall, or did b cause a to launch into the air?

Of course, there are ways to answer this question, but — at least in my opinion — there is no clear answer without more context. It is in this sense that I say that the effects identified in any formal model "is causal": while we can use mathematical notation to try and describe richer theories of causality, the model is unable to establish what one might call "real-world causality."

We can see this emerge in the context of Nash equilibrium. Consider the matching pennies game, an example of which is illustrated in Table 1.

	$a_2 = L$	$a_2 = R$
$a_1 = L$	$\rho, -\sigma$	$-\rho, \sigma$
$a_2 = R$	$-\rho, \sigma$	$\rho, -\sigma$

Table 1: A Matching Pennies Game ($\rho > 0$, $\sigma > 0$)

The game in Table 1 has a unique Nash equilibrium in which each player chooses L with probability $1/2$. However, there is no clear causal relationship in this (again, unique) prediction. This is for several reasons. The first is that the equilibrium is identified by equalities: as one learns in the first couple of weeks of game theory, we find this equilibrium by setting one player's probability of choosing L so as to "make the other player indifferent" with respect to his or her own choice between L and R. Slightly more subtly, this equilibrium is a very good example of how Nash equilibrium can be thought of in terms of player's choices of action (*i.e.*, each player must be optimizing with respect to his or her choice, given the other player's strategy) **or** as an "equilibrium in beliefs" (see Harsanyi (1973)). Regardless of which interpretation one takes, causality is unclear: does player 1 induce player 2 to mix with equal probability, or does player 2 induce player 1 to do so? But the slightly philosophical point I am raising here is that the fact that there exist two different "causal stories" that are consistent with Nash equilibrium implies — to me at least — that game theory alone does not establish strong positive grounds for a specific causality claim.

Moving back to the regression example above, a common first step in interpreting a regression equation as causal is to presume that the right hand side variables (x_a, x_b, and ε) are exogenous, which is often reasonably interpreted as "they are realized prior to the realization of y." Of course,

---

[3] I thank Sean Gailmard and Maggie Penn for very helpful discussions about this point.

this is fine but, strictly speaking, it is an additional assumption above and beyond the information contained in Equation (1), once again demonstrating the agnostic nature of "=" about causality. I've spent enough time on this soapbox — let's start (talking about) building a model!

3 The Building Blocks

Years ago, I was talking to a senior colleague who — in a friendly, teasing way — said, "well, you're a theorist, you can just make it up," to which I responded, "yeah, but we *have to make it up*." He laughed, which I optimistically took to mean that he got the joke. The hardest part of modeling is "making it up." I think of it as analogous to empirical scholars worrying about measurement, choosing and constructing measures and indices, and so forth. While there is no one correct way to build a model, my approach begins with three types of ingredients: *actors, actions, and payoffs*.

3.1 Actors: Who's invited to the party?

The basic building block of most social science models is the set of actors of interest. In a formal sense, these are simply placeholders at first — the details of the actors will be our next two steps — but practically it is important to think about "who" the actors are in order to prepare for the definitions of their actions and payoffs.

There are three broad classes of actors: *strategic actors*, *non-strategic actors*, and *observers*. Strategic actors are the most common type of actor: these actors will not only have choices to make, they will do so in a purposive way. Whether they are "rational" or not is not important at this point — for example, the model might be an agent- or rule-based model — but it is important to think about their payoffs.[4]

Non-strategic actors, on the other hand, will have choices to make, but these actions will not necessarily be driven by the actor's payoffs. In many cases, you do not even need to specify the payoffs for this type of player. An example of such an actor in economic models is sometimes referred to as a "noise trader." Noise traders are, in a nutshell, investors who buy and sell in an apparently random fashion. In political economy models, it is not uncommon to include "noise voters" (*e.g.*, Cho and Kang (2015), Zarazaga (2016)) This type of actor is included in a model essentially as a representation of actors who are not the principal focus of the model, but with whom the strategic actors interact and/or worry about in some fashion.

---

[4] I don't offer much discussion of this in this chapter, but one of the reasons it is important (in my mind) to think about payoffs even when one assumes behavioral rules for actors is that the model otherwise is silent on welfare impacts or proposed changes to behavioral or institutional rules, for example.

Finally, observer actors may not make any choices at all. I refer to them as observers because they might respond in some way to the actions of the strategic and non-strategic actors. For example, the "bargaining before an audience" model due to Groseclose and McCarty (2001) represents "the public" as a single voter, who has an essentially hard-wired "approval function" that the two strategic actors (a president and a unitary Congress) are in part seeking to manipulate. Another example is provided by Gailmard and Patty (2017), who include "society" as an observer — none of the other three strategic actors in their model actually care about society: this observer is included in order to consider the impact of the strategic actors' decisions on "social welfare."

3.2 Actions: What kind of games can they play?

After choosing the actors for the model, you need to define what choices they can make. Obviously, the details of how to do this will depend on what kind of situation you are modeling. In order to offer some advice, then, I will focus a bit on the technical side of the model — backward inducting, if you will, the step of solving the model, as I discuss in Section 4, below.

- **Don't Give the Actors Too Many Choices.** In line with my discussion of the value of simplicity above, try to focus your – and your readers' — attention on the choices that are most important.[5] A good rule of thumb in my mind is "start the model with binary choices." For example, consider the Riker and Ordeshook (1968) model of turnout discussed above. On election day, real people face more choices than simply whether to vote or not, such as "what would I do instead of going to the polls?" The simple fact is that, in general, we don't really care what an individual does that day if they don't vote — we only case about whether they vote or not.[6]

- **Does Actor A Care What Actor B Does?** Particularly in a strategic model, it is important to think about whether (and, possibly, how) each actor cares about what the other actors do. This provides another reason to keep the action set of any actor as simple as possible, especially at first. This is another aspect of modeling that mirrors empirical modeling. If your model has n strategic actors, each of whom has only 2 choices, you will need to model preferences over 2^n potential outcomes, which is analogous to the curse of dimensionality in empirical analysis. Viewed a slightly different way, each combination of individual choices will require n "parameters" to represent the preferences of the actors over these outcomes.

---

[5] This is also in line with my advice for knowing when a model is "done" (Section 7).

[6] Paraphrasing important advice that my friend and coauthor Sean Gailmard told me years ago in the (old) Palmer House bar about reviewers wanting verisimilitude in models, "some people in the real world wear socks, does that mean I should include a variable for whether each actor is wearing socks?"

3.3 Payoffs: Why are they here?

The final initial step of building a model is to specify individuals' payoffs (or, if the model is rule- or agent-based, start thinking about how each actor's rule will be represented). If one or more actors's preferences depend on what one or more other actors choose, then the model involves *strategic behavior*. If, on the other hand, every actor cares only about his or her own decision, then the model is *decision-theoretic* in nature (Riker and Ordeshook (1968) is an example of such a model). Decision-theoretic models are not always sufficient to explain behavior and incentives, but when such a model is sufficient, I recommend using that model.[7]

The Importance of Trade-Offs. Regardless of the type of model one is building, if the model includes assumptions about payoffs (or, more generally, motivations), it is important to consider what kinds of *trade-offs* the actors are facing. In game theoretic terms, for example, an actor who has a *dominant strategy* is an example of an actor who does not face any trade-offs within the model: regardless of what the other actors choose (or are believed by the actor to be choosing), the actor's optimal choice is unchanged. In most cases, such an actor is not particularly central to the model's logic — even though a strategic actor, his or her strategic calculations are in some sense trivial unless one or more of the other actors do not realize that the actor has a dominant strategy.[8]

In strategic models, the presence of trade-offs can lead to multiple and/or mixed strategy Nash equilibria. Consider the Crawford and Sobel (1982) cheap talk model in which one player (the "sender") has private information that he or she can try to share with a second player (the "receiver") by choosing a free, unverifiable message to send. Crawford & Sobel show that, if the sender does not have the same preferences as the receiver, then the sender will always at least partially misrepresent his or her information in equilibrium. This reflects the trade-offs the sender faces — which emanate from the fact that the receiver's preferences differ from the sender's. In the interesting cases (namely, when the two players' preferences differ only by a little), the sender can send a "noisy but informative" signal to the receiver and the receiver uses as much information as he or she can from the received message to make her own choice of action. If the sender did not insert any noise into the message (*i.e.*, if the sender was "always honest" with the receiver), then the receiver would believe the sender's message with certainty, which would give the sender an incentive to always lie to the receiver. However, if the sender always lied to the receiver, there would still be no noise in the signal, and the receiver would correctly adjust for the sender's lie.

---

[7] In slightly more detail, and mimicking one of many important things that my coauthor Maggie Penn has taught me, if a decision-theoretic model can explain part of the logic of a strategic model, you should begin the (oral or written) presentation of the model by walking through the logic in /of the simpler setting. This will not only make the overall logic of the bigger model more transparent, it also can sometimes ironically help you shorten the presentation.

[8] This could be for several reasons: it could be that the actor has private information that reveals he or she has a dominant strategy, or one or more of the other actors is not optimizing his or her choices correctly.

Accordingly, any informative equilibrium must have some *endogenous* noise in it unless the sender and receiver have identical preferences.

The fundamental nature of this kind of trade-off can be seen by the breadth of applications of this model in political economy. And this, to me, is the most important role of simple models: they unite our understanding of seemingly diverse topics *through their simplicity*. Going farther, I believe that the best simple models are those that not only have actors who face trade-offs, but in which the trade-offs are limited in scope for each actor. This is one of the reasons I advised above to not give the actors any more choices than necessary: if an actor faces a choice between three options, X, Y, and Z, then the choice of, say, X, generally does not provide much evidence about whether the actor values Y more than Z or, even more disturbingly, even whether the actor actually prefers X to both Y and Z. An example of this is provided by the logic of both sophisticated voting (Farquharson (1969), Shepsle and Weingast (1984), Penn, Patty and Gailmard (2011)) and strategic voting (Austen-Smith and Banks (1988), Myerson and Weber (1993), McKelvey and Patty (2006), Patty, Snyder and Ting (2009)): when engaged in plurality rule voting over more than two options,[9] one's own preference over the options is important, but it can be optimal to vote for any of the options *other than one's least-preferred option*. This is because one should consider how likely it is that one's vote will be decisive and, if so, decisive for what.[10]

4 Equilibrium or Not?

After you build a model, a natural next question is "how should I solve it?"[11] The most common approach used over the past 30 years or so applies non-cooperative game theoretic arguments and focuses on *Nash equilibria* and its refinements as the prediction concept. As with any solution concept, this is not without its problems. For example, while Nash equilibrium has a clear basis in non-cooperative terms — each strategic actor in the model is presumed to make his or her choices optimally, given his or her information and beliefs about the preferences and behavior of the other actors. The epistemological foundations of the solution concept — one of its principal theoretical/philosophical *strengths* — is arguably the principal *weakness* of Nash equilibrium as a predictive tool in social science. How this problem emerges in practical terms depends on the context to which the concept is applied, but one way to summarize most of the gaps between the solution concept and empirical reality is as follows. *In a Nash equilibrium, the actors are assumed to know much more about the situation than the analyst does.*

---

[9] Sophisticated voting describes the logic one should use when the voting proceeds sequentially over the alternatives. Strategic voting describes the analogous logic when all of the options are voted on simultaneously.

[10] The second question is relevant for sophisticated voting, the first question is the heart of strategic voting reasoning.

[11] Note I say "should" here, because it is sometimes a very different question than how one actually solves it.

As partially alluded to above, I view this weakness as also a potential strength of Nash equilibrium as a theoretical tool: if something cannot happen in Nash equilibrium then, if the model is correctly specified, there is no reason to expect to see the behavior in practice. For example, *subgame perfect Nash equilibrium* is logically complete in a very elegant fashion: it describes all Nash equilibria that are based on Nash equilibrium behavior in all cases.[12] Put another way, a subgame perfect Nash equilibrium describes rational behavior based on the actors' having correct beliefs in all situations. Of course, this elegance comes at a price as a predictive concept. In addition to the fact that individuals might not have correct beliefs, it can be very difficult for someone trained in game theory to derive a subgame perfect Nash equilibrium in even moderately complicated situations and requires that individuals think not only about what they themselves would do in every possible scenario, but also about how all other actors would behave and "think about" every such situation as well.

I am a big fan of subgame perfection as a refinement (most refinements used by scholars these days share a version of this property). That said, it is only one of many possible ways to "solve" a model. Indeed, some models do not even have a clear analogue to Nash equilibrium or any of its refinements. For example, agent-based models often do not have direct analogues for individual preferences and/or beliefs and instead explicitly or implicitly specify rules by which actors make their choices. There is no reason that these rules need to be in equilibrium (though, in practice, they often are). Similarly, cooperative models typically focus on concepts that are distinct from individual optimization. For example, the core is a solution concept that considers collective "preference" over outcomes by focusing on coalitions of actors within the model.[13]

Linking Solution Concepts with Empirical Methods. Every usable solution concept involves making assumptions. For example, Nash equilibrium involves assuming that the players have correct beliefs about each other's strategies, subgame perfect Nash equilibrium imposes the additional assumption that it is common knowledge that all players will use strategies that are rational in all situations that might occur in the game.

In my mind, this is analogous to the assumptions empirical scholars make when using or developing any estimator. As alluded to above, applying the ordinary least squares (OLS) estimator to Equation (1) requires assuming that x_a and x_b are each independent of ε. Interpreting/making inferences about the coefficients typically requires more assumptions, such as that the realizations of ε are independently and identically distributed according to a Normal distribution. Such assumptions are required before one can even offer an answer to simple questions such as "how much

---

[12] This differs from Nash equilibrium because it rules out "non-credible threats" or, perhaps, "mistakes," being the basis for actors' beliefs about counterfactuals.

[13] There are many similar but distinct solution concepts in this vein, such as the Pareto set, the uncovered set (Miller (1980)), and the Banks set (Banks (1985)).

information is contained in this data set?"

Regardless of whether one is choosing a solution concept or an estimator, it is almost certainly the case that you will need to justify the assumptions embodied in your choice and discuss the degree to which they are likely to hold in practice in the phenomenon being studied. This is one reason that I think of choosing one's solution concept as part of the modeling process, rather than simply a "step in the analysis" of the model. The model and solution concept applied are — to me at least — inextricably linked in the best research.

Choosing a Solution Concept as Modeling. One's choice of solution concept is inherently a theoretical choice. Nash equilibrium is arguably a good solution concept whenever individuals have unilateral discretion over their actions. However, it suffers from normative and predictive problems of its own: there is no reason to suppose that a Nash equilibrium is socially efficient (*e.g.*, the Prisoner's Dilemma) and, in general, there may be multiple Nash equilibria.[14] This is seen most acutely in the case of infinitely repeated games (situations in which actors interact in the same situation repeatedly into the future), where the *folk theorem* (Friedman (1971)) tells us that there may be infinite (non welfare-equivalent) Nash equilibria.

The choice of how to solve the model represents one of many selection choices that a modeler must make. The central question, in my mind, when choosing how to solve a model is the nature of the argument that one wants to offer. Nash equilibrium and its refinements are robust with respect to individual incentives, cooperative concepts tend to be responsive to collective welfare, and agent/rule-based model are arguably more easily calibrated with respect to the empirical regularities of individual behavior. A simple conclusion from this variety is that the is no one "true" way to solve a model. I would personally argue that the non-cooperative approach embodied by Nash equilibrium is the best "first guess" because it not only embodies methodological individualism, it also can speak to questions of collective welfare and the rules (*i.e.*, strategies) by which individuals "should" behave in a given context. But, again, that's my personal belief.[15]

In the best of all worlds, every model would be solved in every way—the various behavioral/welfare/outcome differences implied by different solution concepts are also themselves an aspect of models *qua* research design. Take the Baron-Ferejohn bargaining setup as an example (Baron and Ferejohn (1989)). A set of n individuals must choose how allocate a unit of utility among themselves using a random recognition alternating offers bargaining protocol. While the model has an essentially unique equilibrium with proper refinements of Nash equilibrium (such

---

[14] Indeed, this multiplicity has led to the creation of various refinements of Nash equilibrium, such as subgame perfection, Markov perfection, and so forth.

[15] Paraphrasing what my mentor and coauthor Richard McKelvey once told me, the predictions of a model in which individuals' decisions do *not* satisfy the properties of a Nash equilibrium would be immediately suspect if anyone modeled in the model reads the paper! Whether that is a real worry is, of course, an empirical matter.

as subgame perfection, stationarity,[16] and elimination of weakly dominated voting strategies) in which a minimal winning coalition (*i.e.*, a bare majority) of the individuals receive a positive allocation of utility, it is frequently forgotten that, if we do not apply these refinements and instead consider all Nash equilibria, *any division of the utility can be supported in equilibrium.* Note that I do not see this as a weakness of Baron & Ferejohn's model, which represents a straightforward extension of the Rubinstein alternating offers model (Rubinstein (1982)),[17] the "shadow of the future" provides the analyst an incredible range of strategies that individuals *could* use to "enforce" all sorts of allocations, including many that are quite pathological. For example, *there are Nash equilibria of both the Rubinstein (1982) and Baron and Ferejohn (1989) models that involve no bargain ever being struck.* Rather, I view it as rather amazing that both models can be essentially uniquely solved by *any* non-cooperative solution concept.

The fact that there is an essentially unique equilibrium in subgame perfect, stationary, and weakly undominated strategies indicates that the incentives offered within the model (*i.e.*, as produced by the bargaining protocol) is even more informative once one notes that removing *any* of these refinements immediately results in there being an *infinity* of allocations that can be supported in a Nash equilibrium satisfying the other two refinements. In other words, if we wish to provide anything resembling a useful positive prediction about what should emerge from such a bargaining situation, we need each of the refinements: the solution concept used by Baron & Ferejohn is arguably as "slim" as possible to reach its prediction.

This raises the question of whether Baron & Ferejohn's model is "realistic." Empirical evidence about the model's validity is weak. Some have argued that this is because the model is unrealistic in substantive ways including the assumptions that individuals are "selfish" and indifferent about which other players also get positive allocations from the pie, that all individuals are rational, or that the proposer in each period is drawn at random in an independent fashion across time. But this criticism is wrong-headed in my opinion. If we ask not whether the observed allocation is consistent with the stationary, weakly undominated subgame perfect Nash equilibrium but instead whether it is consistent with some Nash equilibrium, the empirical validity of the model i arguably impossible to reject based only on the final allocation of the pie![18] Accordingly, the "unrealistic" part of Baron and Ferejohn (1989) is not the structure of the model itself. Rather, the solution

---

[16] Stationarity assumptions come in slightly different flavors, but most such assumptions can be described as "people always behave the same way in "the same situation" throughout the game.

[17] And, to be clear, Rubinstein's model has the same property: the model has a unique subgame perfect Nash equilibrium, but any allocation can be supported as the outcome of a Nash equilibrium.

[18] Given the aim of this article, I am being a bit sloppy here. One could expand the data being analyzed to include the history of rejected proposals (for example) or even collect "more data" about individual-level characteristics, such as the players' beliefs about how other players would respond to various proposals. That said, this is often infeasible and, regardless, one would need to make more assumptions about the data generating process (*e.g.*, that players truthfully reveal their beliefs about other players' responses to other — unobserved — proposals).

concept they ultimately use is arguably unrealistic.

This brings us back again to robustness: the solution concept that Baron & Ferejohn apply can be described as fragile in several distinct but complementary respects. But *the model itself* is clearly realistic in several important ways and incorporating more realism within it will probably not undermine the fundamental severe multiplicity of Nash equilibria. Furthermore, most widely applied cooperative concepts offer little to no predictive purchase in the "divide the dollar" setting (*e.g.*, McKelvey (1986), Penn (2006)). While one could use agent-based models to generate predictions (*e.g.*, De Marchi and Page (2014)), such an approach will be similarly brittle (*e.g.*, there is no *a priori* reason to suppose that the predictions of such a model will be robust to even slight modifications of the described rules for any single agent) and, ultimately, the complexity of describing the model itself will reduce the transparency of the model's fundamental logic.

5 Notation and Figures

One of the most serious challenges for scholars interested in building their first model is learning and understanding notation. Obviously, there are lots of jokes about the Greek letters, but the real challenge in my mind is conceptual. A practical way to reach a reasonable initial level of familiarity is common to writing, music, and art: find the model(s) you want to emulate and copy them (with attribution, of course)! In addition to being practical, this approach is especially useful to the degree that one agrees with the idea that the best models span multiple topics and applications. As the examples discussed here and there earlier, "workhorse models" are central to social science. As the old saw goes, there's no need to "reinvent the wheel."

When borrowing another scholar's model, I personally recommend that *you adopt his or her notation* (at least at first). This makes it a lot easier to both check your own work as well as — and this is very important when you're starting out (and even thereafter) — sharing your work with others. Getting the model right is a combination of design choices and solving the model correctly, as described above, but it also requires proper interpretation and presentation.

The "soft middle" of modeling is really an exercise in measurement: who do your actors represent? What are the real world actions that the choices they face in the model represent? What do the payoffs, information structure, and other various parameters of the model represent? A partial way to help understand this is very much in the spirit of EITM: how would one "test" your model's predictions? This form of translation is, in my experience, the most important challenge when trying to get your research published in a peer-reviewed journal.[19] A first step toward making this

---

[19] Similarly, in my experience, this is even more true in political science than many other disciplines, because political science is only lightly "subfield-specific" relative to many other disciplines (particularly quantitative ones). Your "median reviewer" will often not be working on the specific topic that your research is focused on. This is part

challenge less onerous is making a proper choice of notation.

Mnemonics and Conventions. There are several notational norms that are wise to at least keep in mind. Some of them are typographical in origin. For example, one rarely sees the lower case L, l, as a variable. This is at least in part because l looks a lot like the number 1.[20] Similar norms suggest not to use the parameter e if you're going to give it an exponent, because e represents Euler's constant in mathematics.[21]

Before continuing, I will be honest: in line with the analogy between modeling and art, one's notational choices are very personal. That said, here are some pieces of advice about the quotidian details.

1. N is a set of actors. Note that this is not the same as $\mathbf{N}$, which denotes the set of natural numbers, $\{1, 2, 3, \ldots\}$.

2. i is an arbitrary actor. To be honest, I also typically reserve j and sometimes k to represent "actors other than i." Also, for reasons similar to the discussion of lower case L above, the notation I (capital I) is generally good to avoid.

3. u_i is actor i's payoff function (v_i is also used — particularly in dynamic settings in which players make choice over time — but I find that v_i is very useful for, say, denoting the **v**ote choice of individual i when appropriate). This is obviously social science specific, where u naturally (*i.e.*, mnemonically) represents "utility" and v similarly represents "value."

4. Try to avoid superscripts "as arguments": this one is tough to achieve all of the time (in part because I prefer u_i^j to $u_{i,j}$, but the question of whether one should use $u_i(j)$ or $u(i, j)$ as opposed to u_i^j is more of a case-by-case call). There are two reasons to avoid superscripts if possible: the first is that the result can easily look overly grandiose/complicated to your reader. The second, and more of the reason I avoid using them when I can is that superscripts represent exponentiation (*i.e.*, "raising to a power"). For example, u_i^1 and u_i^2 could each represent free parameters, or it might be the case that (by definition) $u_i^2 = u_i^1 \cdot u_i^1$.

5. In general, "backward induct" your notation. One way to do this is to prepare your model as slides, laying out the notation, and putting a word label for each: make your slides as easy

---

of the reason, in my opinion, that so many scholars (including myself) describe their research as being focused on "representation" and/or "accountability."

[20] And, to be clear, *l* (*i.e.*, a lower case L) looks a lot like 1 (*i.e.*, the number 1 italicized).

[21] That said, for some reason, we have accepted that π is a variable, as opposed to the ratio of a circle's circumference to its diameter. And, as I mention below, i is a *very* common indicator for agents in a social science model, as opposed to the square root of -1.

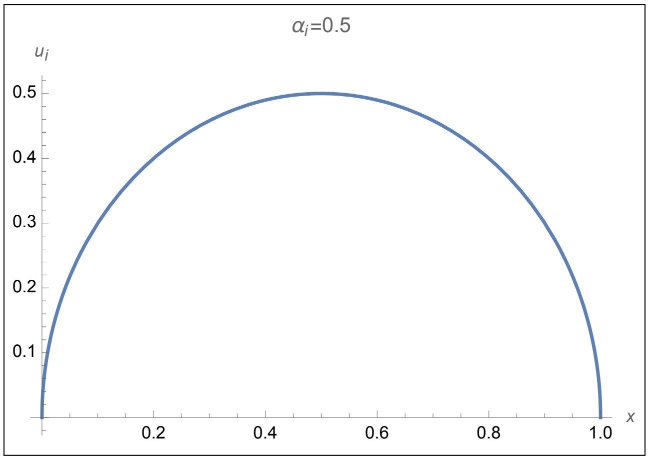

Figure 1: Cobb-Douglas Payoffs from $x \in [0, 1]$, with $\alpha_i = {}^1\!/_2$

as possible to "not get lost in": practicing presenting the model to others — particularly as you are building it — is a really good way to road-test the choices.

An overall piece of advice is to always remember that *good models clarify the logic of your argument*: don't make notational choices that obscure the role of your parameters and design choices in your argument. There's a lot in here: not only in your article, but in your presentations of the work, there will sometimes be a need to present an equation from the paper: good choices of notation make this much more powerful. For example, it is hard to overestimate the degree to which people are familiar with the basic concepts — **e**nergy and **m**ass — in Einstein's famous equation, $e = m \cdot c^2$ (and, again, refer to my argument against superscripts as arguments, above).

Notation is important, because it is the syntax of the language you're developing for your model. A related question is the role of pictures or, more formally, "figures." "A picture is worth a thousand words" is another old saw. I agree, and think it is often an underestimate. The simple fact is that even the most simple models typically have more than single parameter of interest. Just to provide a simple example of this, consider a model in which an actor, i, is choosing how much time to work $x \in [0, 1]$, with the remainder $(1 - x)$ going to goofing off (or, as the British say, "leisure"), and i has a Cobb-Douglas utility function of the following form:

$$u_i(x) = x_i^{\alpha} \cdot (1 - x)^{1-\alpha_i},$$

where $\alpha \in [0, 1]$ is an exogenous parameter. This is a very convenient and widely used functional form (it loosely represents Jack Nicholson's character in *The Shining*: "all work and no play makes Jack a dull boy"), but its shape is not immediately obvious to most people. The following figure illustrates what it looks like when $\alpha = {}^1\!/_2$. Figure 1 clarifies that x has a non-monotonic effect on u_i, that it is a smooth effect (it is continuously differentiable in x), and it is maximized at $x = {}^1\!/_2$.

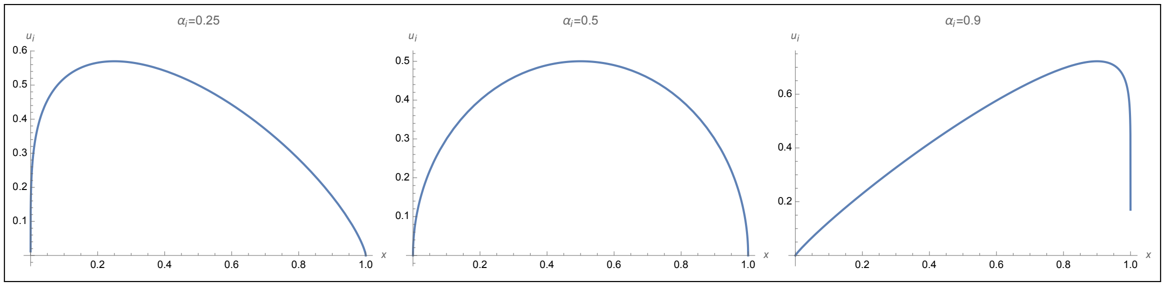

Figure 2: Cobb-Douglas Payoffs from $x \in [0, 1]$, for $\alpha_i \in \{0.25, 0.5, 0.9\}$

The value of such figures does not stop there: below, I have included figures showing the effect of x on u_i for three arbitrary values of $\alpha_i \in \{0.25, 0.5, 0.9\}$. Figure 2 partially displays not only the comparative statics of u_i with respect to x, it also shows (most importantly) the effect of α_i on the value of x that maximizes u_i (*i.e.* the optimal value of x) for a few different values of the exogenous parameter, α_i.

An Undervalued Graphic: The Extensive Game Form. When presenting a decision-theoretic or strategic model, it is *crucial* to include a "picture of the model."[22] In many cases, this picture is the extensive game form. Drawn well, this can display a simple decision- or game-theoretic model in a very flexible and transparent way. The flexibility of a picture of the game form is that the reviewers/readers can find answers to various different questions within it, making the choice of language to describe the model less onerous. The transparency of a game form is a function of how one chooses to draw it, but there are several norms in presenting these figures that greatly simplify the choices when one is drawing the figure.[23]

One of the subtle aspects of a model, in my experience, is its *information structure*. Essentially, who knows what, and when? For example, *moral hazard* and *adverse selection* are closely related, but distinct, problems in agency situations. A key distinction between the two is *what is unobserved* by the actors. In moral hazard problems, one or more actor cannot directly observe the action chosen by some other actor in the model. In adverse selection problems, one or more actors know some other aspect of the model (*e.g.*, the actor's "type").[24] This timing is central to the more general question of *credible commitments* within the model being analyzed: problems of moral hazard and/or adverse selection are "solved" (if solved at all) with different timings and/or information structures.

As a very simple example, suppose that you are modeling two players, "1" & "2," who face a

---

[22] This also comes from lessons taught to me by Maggie Penn, who is an expert at coming up with the perfect figure to make even complicated points.

[23] That said, as the model gets more complicated, these choices also become more complicated in several ways, including the obvious problem of "fitting it all into one picture."

[24] Of course, real agency problems typically have both moral hazard *and* adverse selection aspects.

"Pareto-ranked" coordination problem as displayed in Table 2.

	$a_2 = L$	$a_2 = R$
$a_1 = L$	100,100	0,0
$a_2 = R$	0,0	1,1

Table 2: A Pareto-Ranked Coordination Game

When the players choose simultaneously (*i.e.*, in ignorance of each other's decisions: a symmetric, dual-sided moral hazard problem) the game has three Nash equilibria: (L, L), (R, R), and the mixed strategy equilibrium in which each player chooses L with probability $\frac{1}{101} \approx 0.01$ and chooses R with probability $\frac{100}{101} \approx 0.99$. This can be inconvenient for various reasons (including, but not limited to, using the model for structural estimation). There are at least two ways to fix (or, really, sidestep) this problem.

Applying Selection/Refinements. Suppose that you are interested in considering only Pareto efficient equilibria (which in this case is unique: (L, L)). Then one can apply selection arguments, such as *payoff dominance* (Harsanyi and Selten (1988)).[25] One of the challenges of selection (or refinement) arguments is that they tend to be fragile and limited in scope in terms of what kinds of models they produce the desired results in. Payoff-dominance works well in the game pictured in Table 2, but it does not provide the same purchase in the structurally similar model pictured in Table 3.[26]

	$a_2 = L$	$a_2 = R$
$a_1 = L$	100,1	0,0
$a_2 = R$	0,0	1,100

Table 3: An Asymmetric Coordination Game

The game pictured in Table 3 has *two* payoff dominant equilibria: (L, L) and (R, R), demonstrating the importance of players having somewhat aligned preferences for payoff dominance to provide a unique selection.

Change the Timing and Information Structure. Returning to the interaction as modeled in Table 2, a second approach to justify focusing on the Pareto efficient Nash equilibrium is to alter the timing and information structure. If Player 1 chooses first, knowing that Player 2 will choose

---

[25] Or, as is often done in signaling models such as Crawford and Sobel (1982), one can simply apply an *ad hoc* refinement to Pareto dominant equilibria. Part of my point here is that such a selection criterion is not necessarily any more *ad hoc* than other of the other many modeling choices one has already made.

[26] The game pictured in Table 3 is often referred to as either "battle of the sexes" or "Bach or Stravinski."

after observing Player 1's choice (hence eliminating Player 1's moral hazard problem with respect to Player 2, but not the reverse), the game can be easily and transparently drawn as in Figure 3.

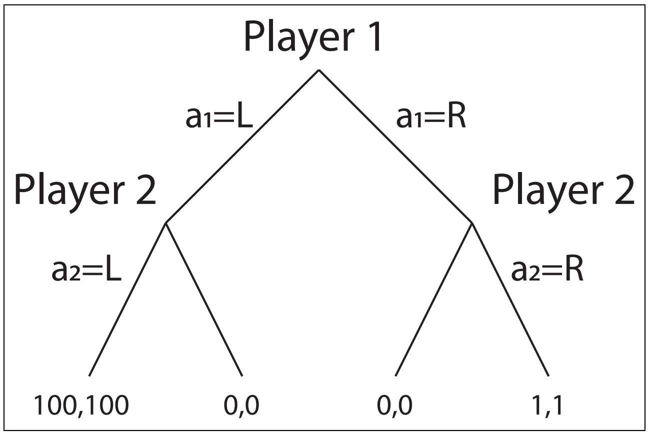

Figure 3: Sequential Pareto Coordination Game

Because Player 2 could in principle use a behavior strategy that involves randomizing after either or both of Player 1's choices, the modified model pictured in Figure 3 actually has more Nash equilibria than the simultaneous-move version of the game in Table 2.[27] However, it has a unique subgame perfect Nash equilibrium: (L, L).

Unlike payoff dominance, subgame perfection is well-defined in all finite games — the game as pictured in Table 2 has only one subgame, while the sequential version in Figure 3 has three distinct subgames,[28] so each of the three Nash equilibria for the game in Table 2 is (trivially) subgame perfect. To show one aspect of the value of extensive form figures, Figure 4 displays an extensive form that is strategically equivalent to the game pictured in Table 2.

The dashed line connecting Player 2's decision nodes — Player 2's *information set* —represents the assumption that Player 2 cannot observe Player 1's choice of $a_1 = L$ or $a_1 = R$ prior to making his or her own choice of $a_2 \in \{L, R\}$. In some cases, it can greatly help a reader to understand the impacts of, and reasoning behind, the assumptions you make in your model to see and compare Figures 3 and 4.

---

[27] The game has an uncountable infinity of Nash equilibria, but only 3 if player 2 is restricted to using a pure strategy.

[28] A subgame corresponds to a "game within a game," and serves as the foundation for subgame perfect Nash equilirium. In the game pictured in Figure 3, one of its three subgames corresponds to the "whole game" (Player 1's initial choice and both of Player 2's potential choices), one corresponds only to Player 2's choice following Player 1 choosing $a_1 = L$, and one corresponds only to Player 2's choice after Player 1 chooses $a_1 = R$.

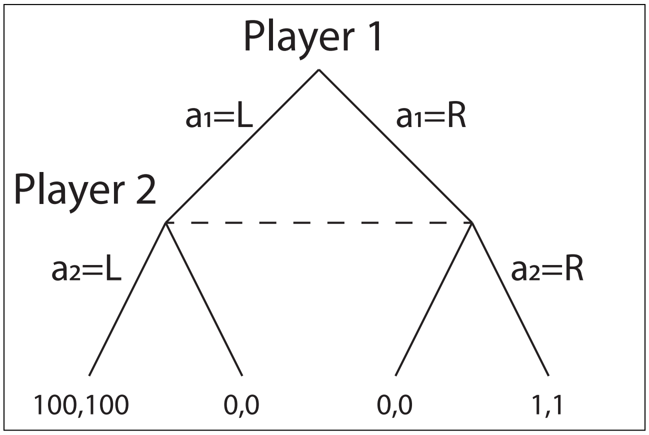

Figure 4: One Extensive Form Representation of Table 2)

6 What is an Interesting Question?

One arguable downside of the generality of simple models is the challenge of choosing what questions one should focus on when analyzing the model.[29] Due to its ubiquity, there are an infinity of ways to frame this challenge. Sadly, I do not have a silver bullet for this, but I do have some general principles.

Is the prediction interesting? Of course, there are several ways for a prediction to be interesting. Have any other scholars considered the predicted phenomenon in question empirically or theoretically before? If so (and especially if it has attracted interest recently), then this implies that there is a non-empty pool of potential reviewers for the model who might find it of interest. If not, then one should ask oneself whether one wants to try and convince readers and reviewers to "care about" the prediction.

Assuming that the phenomenon has attracted scholarly attention recently, the next question is whether the prediction distinguishes your model from one or more other explanations for the phenomenon in question. If not, then one must consider whether this is surprising: it might be, depending on the divergence between the building blocks and/or solution concept used to analyze the model. That's a fairly rare occurrence in my experience, but it is definitely possible to publish important work on how/why two seemingly disparate explanations are actually (even if only partially) *observationally equivalent*.[30]

---

[29] For the purposes of this chapter, I am largely ignoring purely "mathematical" contributions. These are important, but they're often less about modeling itself than about extending the class of models that one can solve/use.

[30] An example of this is provided in Gailmard and Patty (2019*b*), who construct a model of hyper-rational voters and politicians that generates predictions consistent with what might appear to be myopic or irrational voters.

Is the prediction "testable"? Even if the prediction is not particularly interesting, it might serve as a way for scholars to empirically validate or reject the model's assumptions and/or solution concept. Of course, some predictions are testable *in theory* in the sense that the data required has not been collected yet or, in some cases, cannot be collected with current techniques. Note that this folds back onto the question of observational equivalence within the set of explanations that have been offered for the phenomenon under consideration — a gold standard with respect to the importance & testability questions is a prediction that distinguishes the model from at least one other explanation *and* is testable, even if only in theory.

Can the prediction be tested independent of the model's other assumptions? An example of a commonly employed parameter in many models is an actor i's *discount factor*, which I will denote by $\delta_i \in [0,1]$.[31] For example, this is the only parameter in the alternating offers bargaining model presented by Rubinstein (1982), where the unique subgame perfect Nash equilibrium involves the first proposer offering the following division of the unit of transferable utility:

$$x^* = (x_1^*, x_2^*) = \left(\frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right)$$

and the second player accepting it, concluding the bargaining. In order to test this prediction, one must somehow measure individual discount factors (*i.e.*, their "patience levels"). This is a difficult challenge and, indeed, the measurement of it will almost certainly be making the same assumption about exponential discounting. Any such test is simultaneously a test of both the assumption that individuals evaluate the future using an exponential discount and the logic of the model *given this assumption is valid*. Applying casual backward induction to the process of ultimately publishing the model and influencing future work in your field, note that a failure of the data (which includes the estimates/measurements of δ_i) to confirm the prediction of the model is a rejection of at least one of two different assumptions: the first is the structure and solution of the model itself, and the second is the much more foundational assumption about how people evaluate future potential outcomes.[32]

General Tip: Present Your Work. The most important practical advice I can give about how to determine which question(s) are interesting to other scholars is that presenting your work to others is like voting: *do it early and often.* People often worry too much about "being scooped." Every

---

[31] I am thinking of any model that adopts the common (even if empirically dubious) assumption that individuals evaluate intertemporal trade-offs (*i.e.*, "money today versus more money tomorrow") using an exponentially discounted payoff function, where the value of receiving utility u_i after t times periods is equivalent to receiving $\delta_i^{t-1} u_i$ today.

[32] In the context of Rubinstein (1982) and other complete information bargaining models, this is also a test of the assumptions that players have (1) commonly known discount factors, (2) commonly known payoff functions and (3) common knowledge of each others' rationality.

important social science model with which I am familiar has been developed and/or extended by multiple scholars, often independently. The measure of a model is not how original it is as much as "how much impact does the argument the model is used to make have on the literature?" A model unread is tautologically a model with no impact.

7 Finally, When is a Model "Done"?

The final challenge when building model is to know when to stop. Again, I am unaware of any silver bullet in answering this question, but have a few principles.

Do you have any more ideas? When a student or colleague asks me whether the model is ready to submit, my first question is usually, "is there anything you wish the model would include?" And my follow-up question is usually, "is there anything in the model that you don't like?"

If the answer to the first question is "yes," then the question becomes whether and how the absent aspects of the model can be included. A good example of something that people often "want to include" in their model (particularly game theoretic models) is to make the model infinitely repeated. I understand the urge, of course: infinitely repeated games are sometimes very easy to analyze (with the right assumptions, of course) and, to not mince words, "make the model seem fancy/complicated."[33] I almost always urge the author to resist this urge (at least until the article gets a revise and resubmit that asks for such an extension). This is for several reasons.

In addition to the folk theorem discussed above, it is important to remember that the assumptions required to actually solve for an equilibrium in such a setting will make the model much more fragile.[34] Accordingly, one should apply a form of cost-benefit analysis when choosing whether to extend the model in such a way: will the predictions and/or insights drawn from the extended model be sufficiently greater than the current model's to warrant the increase in the extended model's complexity and fragility? While adding such extensions can be thought of as "robustness checks," this is one of the reasons I urge scholars (particularly untenured faculty and graduate students) to not extend the model in this kind of way prior to being directly asked by a reader, reviewer, or editor to do so.[35]

Does the model have a point? Referring back to the discussion of interesting questions above, a second important question to consider prior to declaring the model finished is "what is the main

---

[33] Note that these two motivations are potentially inconsistent with each other.

[34] This is also true of several "fancy" aspects of some game theoretic models, such as global games (Morris and Shin (2003)), Bayesian persuasion (Kamenica and Gentzkow (2011)), and Markov perfect equilibrium (Maskin and Tirole (2001)), to name three.

[35] I offer the same kind of advice to scholars doing empirical work, too.

point of the model?" Or, in other words, "if readers took one message away from your model, what would it be?" This is one of the most important reasons to share your model with colleagues *as you are developing it*. Oftentimes, the best person to evaluate the point of a model is not the model builder. And, finally, an additional value of thinking about this question is pragmatic: *where will you submit the model for publication*?

Does the model try to make too many points? Echoing the advice I gave in Section 3, models should be as simple as possible not only in terms of their structure and assumptions, but also in terms of their message. For reasons of space, I have mostly focused on building a model, rather than the art of writing the model up in an article intended for peer-reviewed publication. That said, the most important advice I can give in writing a model up is: *focus on one idea* in any given paper.[36] Especially when starting out, scholars often have an understandable urge to try and demonstrate "how important" their model is by cramming as many predictions into the article as possible (and I am definitely guilty of doing this, and probably will do it again!). Not only does this stray from the strengths of building a simple model described above, it also will be more likely to confuse your reviewers and, ironically, can increase the odds that your article is rejected because each reviewer asks for you to focus a revision on two or more disparate aspects of the model and its predictions. Keep the reviewers "on point" by choosing one goal in the article. If your model has two or more important points, remember that you can use the same model in multiple papers.[37]

Conclusion Following my own advice, I do not remember any general advice beyond what I wrote above, so I'm done here.

---

[36] Furthermore, I think this is good advice for *any* social science article. Barry Weingast has written a very useful description of this and other advice for writing an academic paper (Weingast (2010)).

[37] If a model makes multiple important points, it is an excellent candidate to serve as the centerpiece of a book length manuscript.

References

Austen-Smith, David and Jeffrey Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review* 82(2):405–422.

Axelrod, Robert. 1970. *Conflict of Interest: A Theory of Divergent Goals with Applications to Politics*. Stamford, CT: Markham Publishing.

Banks, Jeffrey S. 1985. "Sophisticated Voting Outcomes and Agenda Control." *Social Choice and Welfare* 1(4):295–306.

Baron, David and John Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83:1181–1206.

Cho, Seok-Ju and Insun Kang. 2015. "Open Primaries and Crossover Voting." *Journal of Theoretical Politics* 27(3):351–379.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.

De Marchi, Scott and Scott E Page. 2014. "Agent-based Models." *Annual Review of Political Science* 17(1):1–20.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Farquharson, Robin. 1969. *The Theory of Voting*. New Haven, CT: Yale University Press.

Fearon, James D. 1995. "Rationalist Explanations for War." *International organization* 49(3):379–414.

Friedman, James W. 1971. "A Non-Cooperative Equilibrium for Supergames." *The Review of Economic Studies* 38(1):1–12.

Gailmard, Sean and John W. Patty. 2012. *Learning While Governing: Information, Accountability, and Executive Branch Institutions*. Chicago, IL: University of Chicago Press.

Gailmard, Sean and John W. Patty. 2017. "Participation, Process, & Policy: The Informational Value of Politicized Judicial Review." *Journal of Public Policy* 37(3):233–260.

Gailmard, Sean and John W. Patty. 2019*a*. "Giving Advice *vs.* Making Decisions: Transparency, Information, and Delegation." *Political Science Research & Methods* 7(3):471–488.

Gailmard, Sean and John W Patty. 2019*b*. "Preventing Prevention." *American Journal of Political Science* 63(2):342–352.

Gilligan, Thomas and Keith Krehbiel. 1987. "Collective Decision-Making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics, and Organization* 3:287–335.

Groseclose, Tim and Nolan McCarty. 2001. "The Politics of Blame: Bargaining before an Audience." *American Journal of Political Science* 45(1):100–119.

Harsanyi, John C. 1973. "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points." *International journal of game theory* 2(1):1–23.

Harsanyi, John C and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: The MIT Press.

Hotelling, Harold. 1929. "Stability and Competition." *Economic Journal* 39(1):41–57.

Joseph, Michael F. 2021. "A Little Bit of Cheap Talk is a Dangerous Thing: States Can Communicate Intentions Persuasively and Raise the Risk of War." *The Journal of Politics* 83(1):166–181.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101(6):2590–2615.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago, IL: University of Chicago Press.

Maskin, Eric and Jean Tirole. 2001. "Markov Perfect Equilibrium: I. Observable Actions." *Journal of Economic Theory* 100(2):191–219.

McKelvey, Richard D. 1986. "Covering, Dominance, and Institution-Free Properties of Social Choice." *American Journal of Political Science* 30:283–314.

McKelvey, Richard D. and John W. Patty. 2006. "A Theory of Voting in Large Elections." *Games and Economic Behavior* 57(1):155–180.

Miller, Nicholas R. 1980. "A New Solution Set for Tournaments and Majority Voting: Further Graph-Theoretical Approaches to the Theory of Voting." *American Journal of Political Science* 24(1):68–96.

Monroe, Nathan, John W Patty and Elizabeth Maggie Penn. 2018. "What Is Pivotal Politics (and What Else Can It Be)?" *The Journal of Politics* 80(3):1088–1099.

Morris, Stephen and Hyun Song Shin. 2003. Global Games: Theory and Applications. In *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)*, ed. M. Dewatripont, L. Hansen and S. Turnovsky. New York, NY: Cambridge University Press pp. 56–114.

Myerson, Roger and Robert Weber. 1993. “A Theory of Voting Equilibria.” *American Political Science Review* 87(1):102–114.

Patty, John W. 2024. “Designing Deliberation for Decentralized Decisions.” *American Journal of Political Science* 68(2):783–796.

Patty, John W., James M. Snyder and Michael M. Ting. 2009. “Two’s Company, Three’s An Equilibrium: Strategic Voting and Multicandidate Elections.” *Quarterly Journal of Political Science* 4:251–278.

Penn, Elizabeth Maggie. 2006. “Alternate Definitions of the Uncovered Set, and Their Implications.” *Social Choice and Welfare* 27(1):83–87.

Penn, Elizabeth Maggie, John W. Patty and Sean Gailmard. 2011. “Manipulation and Single-Peakedness: A General Result.” *American Journal of Political Science* 55(2):436–449.

Ramsay, Kristopher W. 2011. “Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power.” *International Studies Quarterly* 55(4):1003–1023.

Reiter, Dan. 2003. “Exploring the Bargaining Model of War.” *Perspectives on Politics* 1(1):27–43.

Riker, William and Peter C. Ordeshook. 1968. “A Theory of the Calculus of Voting.” *American Political Science Review* 61(1):25–42.

Rubinstein, Ariel. 1982. “Perfect Equilibrium in a Bargaining Model.” *Econometrica* 50:97–110.

Shepsle, Kenneth A. and Barry R. Weingast. 1984. “Uncovered Sets and Sophisticated Voting Outcomes with Implications for Agenda Institutions.” *American Journal of Political Science* 28(1):49–74.

Weingast, Barry R. 2010. “Caltech Rules for Writing Papers: How to Structure Your Paper and Write an Introduction.” Mimeo, Stanford University.
URL: *https://weingast2023.sites.stanford.edu/caltech-rules-writing*

Zarazaga, Rodrigo. 2016. “Party Machines and Voter-Customized Rewards Strategies.” *Journal of Theoretical Politics* 28(4):678–701.