Endogenous Beliefs in Models of Politics*

William Minozzi†

November 2012

Forthcoming, *American Journal of Political Science*

Abstract

In the standard rational choice model, actors have exogenously given beliefs that perfectly match objective probabilities. As such, these beliefs cannot be optimistic or motivated by preferences, even though substantial empirical evidence indicates that human beliefs routinely satisfy neither of these criteria. I present a tractable, Endogenous Beliefs Model and apply it to three different political environments from across the subfields of political science. In the model, players form beliefs that maximize a utility function that represents preferences over outcomes and the anticipatory experience of uncertainty. Applications include voter turnout, taxation and collective choice, and crisis bargaining. The model captures the empirical evidence about belief formation much better than the standard model. Moreover, these applications show how rigidly insisting on the standard rational choice model rejects otherwise reasonable explanations by fiat, precisely because of its implausible assumptions about beliefs.

*I thank Scott Ashworth, Ethan Bueno de Mesquita, Jim Fearon, Michael Neblo, Philipp Rehm, Dave Siegel, Craig Volden, and Alan Wiseman for helpful comments and conversations.

†Assistant Professor, Department of Political Science, 2137 Derby Hall, Ohio State University, Columbus, OH 43210 Phone: 614-247-7017, Email: minozzi.1@osu.edu

Most contemporary formal models of politics are built on rational choice theory, which assumes that actions are motivated by preferences and beliefs.[1] For example, the classic decision-theoretic model of turnout postulates that a citizen chooses whether to vote based on her preferences over candidates, her costs of voting, and her beliefs about how likely her vote is to be pivotal in the election (Riker and Ordeshook 1968). On the one hand, rational choice models posit exogenously given preferences. In the turnout model, a citizen's preferences over candidates and costs of voting are assumed to be given. On the other hand, in the variant of rational choice that dominates political science, beliefs are assumed to conform to an exogenous, objective standard. Beliefs thus serve as a rigid background against which the implications of preferences play out. In the turnout model, the citizen's beliefs about her probability of being pivotal are required to decrease as the size of a population increases, and one can draw the inference that turnout should be lower in larger electorates.

The standard assumptions about beliefs have been tremendously generative, yielding valuable insights into almost every facet of political science. But political scientists have not explicitly confronted the way the standard assumptions limit the generality and utility of the model. This fact is surprising given the vast amount of empirical evidence that contradicts the standard assumptions. Empirically, beliefs tend to be overly optimistic rather than realistic (Weinstein 1980), to be motivated by desires rather than independent of them (Kunda 1990), to be automatic rather than conscious (Bargh and Chartrand 1999), and to depend on prior information more than logic would allow (Edwards 1982). Put starkly, human belief formation is often very different from that posited in the standard rational choice model, and these differences are bound to alter foundational results, causal mechanisms, empirical implications, and predictions in a wide range of cases.

Recent work at the intersection of psychology and economics has begun to address the gap between standard rational choice assumptions and well-established empirical regularities of decisionmaking. By reexamining the role that beliefs play in decisionmaking, a substantial

[1] Although the terms "formal" and "rational choice" are sometimes conflated, the two are distinct. There are formal models that are not rational choice, and rational choice theories yet to be formalized.

body of formal work has begun to explore how and why individuals form beliefs and the theoretical implications of alternatives to the standard rational choice model (e.g., Akerlof and Dickens 1982; Brunnermeier and Parker 2005; Caplin and Leahy 2001). The result is a nascent theory of endogenous belief formation. Although this theory may be familiar to some, its insights have not yet found their way into the political science toolbox.

In this paper, I move the research agenda forward in three ways. First, I extend the decision-theoretic model of endogenous beliefs into the multiperson, game-theoretic environment that can accommodate strategic settings. Strategic considerations often cause major, unanticipated changes in decisionmaking, and, as I show, the case of beliefs is no exception. Second, I draw out some of the enormous and pervasive implications of endogenous beliefs for politics. With a set of simple models drawn from across the subfields of political science, I illustrate how rigidly applying the standard rational choice model can sometimes lead us to discard simple explanations that are perfectly consonant with a formal theory and that more closely match the empirical evidence about belief formation. Third, this convergence between theory and evidence promises to appeal to many scholars (particularly political psychologists) who reject formal theoretic tools on the reasonable, though incorrect, basis that these tools *require* a set of empirically unsound assumptions about human behavior.

I begin by reviewing the standard rational choice model of beliefs, and then make the case for a model of endogenous beliefs. The balance of the paper is devoted to the presentation of an Endogenous Beliefs Model.[2] The subsequent sections apply the model in a series of simple examples from different political environments: voter turnout, collective choice of tax rates, and crisis bargaining before conflict. Taken together, these examples illustrate that the standard assumptions about beliefs dramatically and unnecessarily limit the range of causal mechanisms that formal theory can adequately explain.

[2] The Online Appendix offers a formal presentation of the Endogenous Beliefs Model for that class of games.

Beliefs in Models of Politics

The fundamental goal of any rational choice model is to connect an actor's choices to her underlying preferences.[3] When a choice features uncertainty, it also depends on beliefs. The dependence is usually embodied in an expected utility representation. Beliefs are assumed to satisfy some set of axioms, and these axioms imply that actors make choices that correspond to expected utility maximization. Because maximization can often be a relatively simple operation to carry out, rational choice models can be fruitfully applied to a gigantic range of political environments.

In the most frequently applied rational choice model (Von Neumann and Morgenstern 1944), which I call "the standard model," beliefs are assumed to correspond perfectly to some objective probabilities of events.[4] This assumption ensures an easily derivable expected utility representation of preferences, but the model is problematic on at least two grounds. First, belief formation is simply excluded from the standard model. The standard model enters the decisionmaking process *in medias res*, after the emergence of perfect accordance between beliefs and objective reality. While all models must take something to be exogenous, a common goal of formal theory is to endogenize previously exogenous features, especially when there is reason to think that those features are the effects of a more fundamental cause. For example, Carrubba (2009) builds an endogenous model of the development of judicial influence, Greif and Laitin (2004) offer an endogenous theory of institutional change, and Baron (1993) endogenizes how legislators join political parties within a parliament. To date, there have been few attempts to endogenize belief formation in political science using standard game-theoretic tools.

A second problem with the standard model is that its assumptions about these exogenous beliefs lack empirical plausibility. Beliefs are simply assumed to match objective reality

[3] For an excellent overview of the theory of choice, both with and without uncertainty, see Kreps (1988).

[4] In particular, this assumption is constituted by the independence and continuity axioms. Many game theory texts present the rudiments of the standard model, so I do not dwell upon them here.

(perhaps subject to informational limits), yet there is a large amount of psychological research that documents gaps between beliefs and objective probabilities. Persistent, consequential differences between the standard model and empirical evidence about beliefs have emerged (e.g., Tversky and Kahneman 1981; Edwards 1982). For example, beliefs in the standard model are completely independent of preferences, yet research on motivated reasoning and hot cognition suggests that, empirically, beliefs can very much be influenced by desires (Kunda 1990; Redlawsk 2002; Taber and Lodge 2006; Braman and Nelson 2007).

Thus, a large gap has opened between the standard model and what we know about beliefs and belief formation. But is this gap really a problem? When developing theoretical models, one ought not add complexity for complexity's sake. The standard model may be more than adequate for a wide variety of phenomena, and a more complex model is not necessarily a better model. The simplifying assumptions of the standard model can usefully focus attention on the most salient aspects of a causal mechanism, and the large body of work based on the standard model helps clarify which features of any particular example drive the equilibrium predictions. Furthermore, one might still rely on the standard model if there is no suitable alternative.

But there can also be good reasons to move away from the standard model in a wide range of circumstances. It is clear that the exogenous beliefs in the standard model are distortions of those we observe empirically. What is not yet clear is whether these distortions actually matter for many important cases, and whether a suitable, tractable alternative model can be constructed. To address the former concern, however, one must first tackle the latter and consider what alternative options might be available to replace the standard model.

One option would be simply to admit exogenously given, *subjective* beliefs to a model of politics. If we define a *rational* actor to be one who makes consistent choices based on an implicit set of preferences and beliefs, then there is nothing intrinsically rational about objective beliefs.[5] Just as there is no accounting for tastes, there may be no accounting for

[5] The term *rational* has many different definitions (Lupia, McCubbins, and Popkin 2000, 4-8). According to an alternative, normative definition of *rational*, rational choice theory can be taken to be a theory of how

beliefs. Savage (1972) and Anscombe and Aumann (1963) develop rational choice models with subjective beliefs, and show that expected utility representations do not depend on objective beliefs. However, moving away from objective beliefs entails costs as well as benefits. Exogenously given subjective beliefs yield an embarrassment of explanatory riches. For example, in Riker and Ordeshook's (1968) model of voter turnout, preferences over candidates and the voting act are assumed to be exogenous, while beliefs about the probability that a voter is pivotal are required to be objective and to correspond to the number of potential voters. If both preferences and beliefs were "free" parameters, one could simply explain variation in turnout as a result of overly optimistic subjective beliefs about the probability of being pivotal. Thus, the most optimistic would vote, and the rest would not. The result is similar to the D term that Riker and Ordeshook use to explain turnout via a large, positive, duty-based payoff for voting. Neither solution offers much additional explanatory value beyond the assumption itself. When constructing an alternative model of beliefs, one must take care not to assume away the interesting questions and explain too much, too easily. And, of course, if we take the exogeneity of beliefs in the standard model to be problematic, then the assumption of exogenous, subjective beliefs offers little improvement over the assumption of exogenous, objective beliefs.

Objective beliefs are incompatible with a large amount of evidence, yet exogenously given subjective beliefs offer too much explanatory freedom and do not constitute a model of belief formation. What is needed is a model of endogenous beliefs in politics. Such a model should simultaneously conform to the systematic biases we observe, yet do more than simply assume that any subjective beliefs are possible. I present a model that does both.

The Endogenous Beliefs Model (EBM) builds on work in the field of behavioral eco-

to make the *best* choices. Under this definition, requiring beliefs to be objective probabilities is appropriate because doing so ensures that better choices are made. Many of the early developers of rational choice theory intended the theory to be used in exactly this way, and any time one evaluates psychological biases or uses welfare analysis, this is the relevant meaning of *rational*. But this definition is not typically used in political science. Instead, most political scientists intend their rational choice models to be taken in a positive manner, in which case requiring beliefs to be objective is much less justifiable, and must appeal implicitly to a desire for parsimony, tractability, and comprehensibility.

nomics, especially Akerlof and Dickens (1982), Brunnermeier and Parker (2005), and Caplin and Leahy (2001). The central premise of the model is that belief formation is based on preferences that can be represented by a utility function. In the model, players form subjective beliefs before the game begins, and then make choices in the game based on those beliefs. Formally, beliefs are formed by an *agent* who has access to objective probabilities. However, beliefs are not required to reflect this objective information. Instead, the beliefs maximize a utility function composed *anticipatory utility* experienced before uncertainty is resolved and *outcome utility* experienced after uncertainty has resolved. Once a player has formed beliefs, she retains only the belief formed by this agent and has no recall of the belief formation process, and thus no access to objective probabilities.[6] In the equilibrium of the game, players know each other's subjective beliefs, but do not use this knowledge to update their own beliefs. Instead, they simply regard their opponents' beliefs as being incorrect.

These assumptions reflect key disjunctures between the standard model and the weight of empirical evidence about beliefs. Players in the EBM form beliefs that are motivated by their preferences, as in motivated reasoning (Kunda 1990; Redlawsk 2002; Taber and Lodge 2006; Braman and Nelson 2007). They form beliefs in light of objective probabilities, but recall neither these probabilities nor the belief formation process once play begins. Thus, belief formation is automatic and unconscious (Bargh and Chartrand 1999). Beliefs in the EBM have a may differ from objective probabilities because they depend on the anticipation of future potential outcomes, in keeping with idea that emotions experienced during the resolution of uncertainty affect beliefs (Bell 1985; Loomes and Sugden 1986; Loewenstein 1987; Mellers et al. 1997; Rottenstreich and Hsee 2001). Because these beliefs affect the experience of anticipation of potential outcomes, they are generally overly optimistic (Weinstein 1980;

[6] The EBM makes heavy use of both the multi-agent framework (Selten 1975) and the long available but little used game-theoretic tool of imperfect recall (Kuhn 1953; Piccione and Rubinstein 1997). While it can sometimes seem natural to think of an agent at any point in a game as if it were identical to the player, this interpretation is not apt in the case of belief formation. In particular, it would be incorrect to interpret beliefs as a choice because the player has no recollection of their formation, even though belief formation is modeled as utility maximization by the belief-forming agent. The model is made more explicit in the next section and formally specified in the Online Appendix.

Svenson 1981).

Formally, the EBM extends the standard model by assuming that each player's beliefs maximize a utility function. To be sure, such maximization is not implied by two-system model of cognition in which belief formation is automatic. But maximization is the dominant formal tool used to model decision-making in politics. Therefore, building a formal model with endogenous beliefs on the basis of maximization is a reasonable first step.

Other scholars have studied non-standard beliefs in models of politics, but have focused on belief *change* rather than belief *formation*. That is, they have considered non-standard models with *exogenous* beliefs. These models enter the decisionmaking process after beliefs are already present, as in the standard model. Beliefs in these models are exogenously given parameters, just like preferences. For example, Schiemann (2007), Hafer and Landa (2007), Lupia and Menning (2009), and Lupia, Levine, and Zharinova (2010) all study incomplete information games in which prior beliefs adhere to the standard model, but beliefs are not updated via Bayes' rule, in contrast to the standard model. Alternatively, Smith and Stam (2004) and Bullock (2009) differ from the standard model by positing exogenously given subjective beliefs and then analyzing the consequences of updating via Bayes' rule in such a setting. The alternative model developed here studies the formation of beliefs in the first place, which must occur prior to change, whether it is governed as in the standard model or not. The EBM begins with a belief-formation stage, after which players play a game just as in the standard model, but with their endogenously formed beliefs.[7]

To illustrate the EBM, I develop three examples. The virtue of these examples is their simplicity; the goal is to make clear exactly what role endogenous beliefs play. The examples illustrate the model in a decision-theoretic problem; a simultaneous, static game; and a sequential, dynamic game. Because the examples refer to different political environments,

[7] Dickson (2006) and Johnson and Fowler (2011) also present models of endogenous beliefs in politics, but they take an evolutionary game theory approach rather than the individual-level approach of this paper. The evolutionary approach is certainly promising but constitutes a significant leap away from the standard model. Instead, I construct a model of endogenous beliefs that conserves as much of the standard model's apparatus as possible, to help convey the value added by the approach. That is, the model in this paper constitutes a substantially smaller step away from the standard model than does the evolutionary approach.

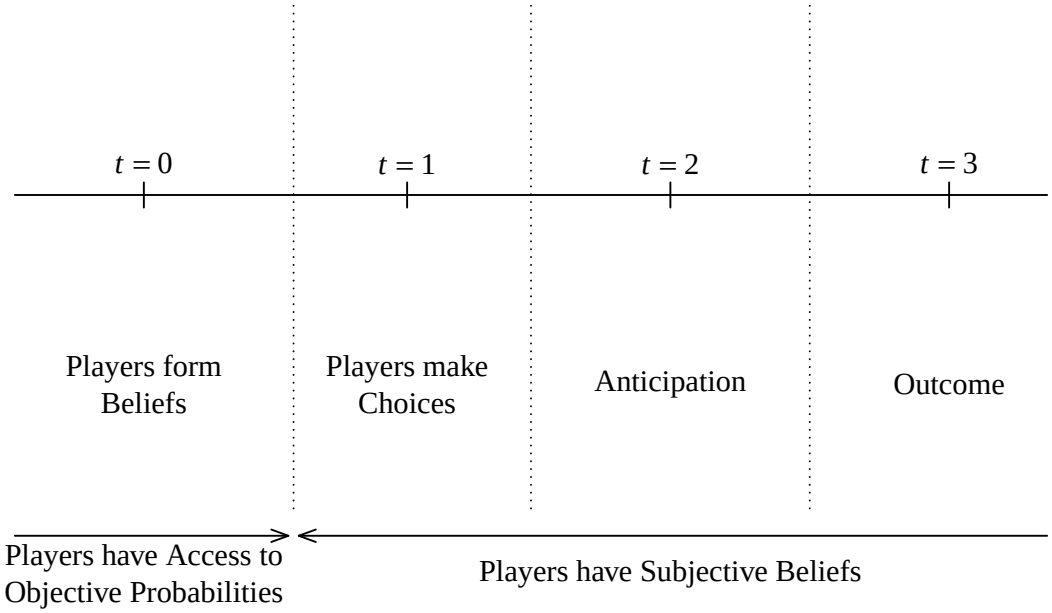

Figure 1: Timing in the Endogenous Beliefs Model

they illustrate that the EBM has important consequences across the fields of political science.

Voter Turnout & the Endogenous Beliefs Model

To illustrate the rudiments of the Endogenous Beliefs Model (EBM), consider a simple, decision-theoretic model of voter turnout under uncertainty, after Riker and Ordeshook (1968).[8] A citizen decides whether to vote in an election. If she does not vote, she pays no cost of voting and the winner of the election is decided without her input. If she votes, she pays the cost of voting and makes a difference in the election outcome with some probability. The cost of voting is small relative to her desire to make a difference in the election. Thus, voting might result in a better or worse outcome than abstaining.

A timeline will help elucidate the EBM and clarify its differences from the standard model (see Figure 1; for additional detail, see the Online Appendix). First, at time $t = 0$, the citizen forms beliefs about whether her vote will make a difference. In the standard model, beliefs would be given exogenously, as objective probabilities. In the EBM, belief

[8] While the EBM would perhaps be better applied to a game like that of Palfrey and Rosenthal (1983), the goal of this section is to take the simplest possible model as a starting point.

formation is modeled as utility-maximization by an agent whose actions will not be recalled by the citizen later in the game. Next, at $t = 1$, the citizen decides whether to vote. Here, the expected utility that she maximizes depends on her beliefs, however she formed them at $t = 0$. This step is the same in the standard model and the EBM, as the citizen has no recall of belief formation in either model. Then, at $t = 2$, the citizen experiences a pause before the outcome of the election becomes clear, during which she anticipates what may or may not happen. Although not typically included in the standard model, this step can be accommodated therein without altering any of the empirical predictions. Finally, at $t = 3$, the outcome of the election is announced. Again, this step is the same in both models.

To describe the citizen's preferences, first consider the common ground between the standard model and the EBM. If the citizen chooses not to vote, her *outcome utility* is normalized to 0. Outcome utility represents preferences over the experience at $t = 3$, which is familiar from the standard model. Voting might increase or decrease outcome utility based on what happens in the election. For simplicity, suppose that the citizen's outcome utility is 1 if she is the pivotal voter and that the cost of voting is $c > 0$.

According to the standard model, belief formation at $t = 0$ is exogenous, and so the citizen's experience during $t = 2$ has no bearing on her beliefs and no effect on her decision. Thus, outcome utility constitutes the total utility for the citizen, and this decision problem can be solved straightforwardly. Because the citizen does not explicitly form beliefs in the standard model, one must assume that she believes that she will be pivotal with some objective probability p. Thus, p is an exogenously given parameter in the standard model. The expected utility from voting is $p - c$, and so the citizen votes if and only if $p \geq c$.[9] Of course, p is typically tiny.

In the EBM, the citizen forms beliefs at $t = 0$. Here, the citizen's beliefs are given by $\hat{p} \in [0, 1]$, the subjective probability that she will be pivotal. The citizen's beliefs are not required to match the objective probability, so $\hat{p}$ need not equal p. Instead, the citizen's

[9] For simplicity, assume that the citizen votes when she is indifferent. This assumption does not affect the qualitative results.

beliefs are formed by an agent who has access to p, which establishes a useful point of contact between the two models. In the EBM, any difference between $\hat{p}$ and p represents motivated reasoning, denial, aspiration, or delusion on the part of the citizen.

Given her beliefs $\hat{p}$, the citizen votes only if it will maximize her subjective expected utility based on $\hat{p}$. That is, when choosing whether to vote at $t = 1$, the citizen no longer has access to the original objective probability p, nor does she recall how she formed the belief $\hat{p}$, in keeping with the automaticity with which people form beliefs (Bargh and Chartrand 1999). Instead, she evaluates the expected utility of voting as $\hat{p} - c$. Thus, in the EBM, the citizen chooses to vote if and only if $\hat{p} \geq c$. This condition is very similar to that from the standard model, in which the citizen chooses to vote if and only $p \geq c$. The explanatory difference between the models therefore hinges on the beliefs that emerge from the EBM.

The beliefs that emerge from the EBM are subjective yet also may be influenced by the objective probabilities familiar from the standard model. Formally, the EBM maintains the distinction between objective probabilities and subjective beliefs via two modeling assumptions. First, the citizen is modeled as a sequence of *multiple selves* (Elster 1987). This is sometimes called the agent-normal form (Selten 1975), in which a player is represented as a set of independent agents who make choices at each stage of the game. This formalism undergirds the familiar logic of credible commitment: an agent at time 1 has no control over what a different agent at time 2 may do, yet both are interpreted as the same player. In the EBM, p and $\hat{p}$ are represented mathematically as heterogeneous beliefs held by different agents at different points in the game. Second, the agent who makes a choice at $t = 1$ has *imperfect recall* of belief formation (Kuhn 1953; Piccione and Rubinstein 1997).[10] Thus, the citizen is modeled as forming subjective beliefs by an agent who has access to objective probabilities, yet the citizen recalls only the beliefs that were formed and neither the process that generated them nor the objective probabilities that influenced them.

Belief formation in the EBM is modeled as the action of an agent at $t = 0$ who maximizes

[10]This explains why I use different verbs ("form" and "choose") for the actions at $t = 0$ and $t = 1$, respectively.

total utility. Total utility depends on how beliefs affect future choices as well as potential experiences both before and after the resolution of uncertainty.[11] Importantly, total utility only plays a role in the EBM at $t = 0$. Specifically, total utility depends on how subjective beliefs affect choices at $t = 1$, the experience of uncertainty resolution at $t = 2$, and outcomes at $t = 3$. Utility that will be experienced at $t = 2$ is called *anticipatory* utility, and it represents the experience of anxiety, dread, enthusiasm, or other anticipatory emotions. Beliefs that lead the citizen not to vote at $t = 1$ (i.e., $\hat{p} < c$) result in no possibility of being pivotal, and hence anticipatory and outcome utilities both equal 0. Beliefs that lead the the citizen to vote (i.e., $\hat{p} \geq c$), result in some chance of being pivotal, and thus potentially nonzero utilities. At $t = 2$, the citizen will not yet know what the outcome of the election will be, and so anticipatory utility will be $\hat{p} - c$. At time $t = 3$, the election's outcome will be revealed, and thus outcome utility will be $p - c$, as in the standard model.

Total utility at $t = 0$ trades off anticipatory utility and outcome utility, as a weighted average of these two components. Specifically, assume that outcome utility has weight 1 and anticipatory utility has weight $\delta \geq 0$.[12] I refer to δ as the citizen's *taste for anticipation*. Thus, the total utility associated with subjective beliefs that lead to voting ($\hat{p} \geq c$) is $\delta(\hat{p} - c) + (p - c)$, and total utility associated with subjective beliefs that lead to not voting ($\hat{p} < c$) is 0. Of especial interest will be any similarities between the standard model and the EBM as δ goes to 0 because such similarities indicate the explanatory role that endogenous beliefs, absent anticipatory utility, may play. Importantly, the mathematical apparatus that constitutes the belief formation stage—including total utility, anticipatory utility, the objective probability, the taste for anticipation, and the process itself—will not be recalled by the future agent who decides whether to vote at $t = 1$. All that will be recalled by that agent is the subjective belief $\hat{p}$.

[11] Here, I mean "uncertainty" in the general sense that any random variable has an uncertain outcome, rather than in the technical meaning of, e.g., Knightian uncertainty.

[12] Preferences over beliefs and outcomes are therefore represented by a psychological expected utility function (Caplin and Leahy 2001).

Tastes for anticipation in this model are analogous to the risk preferences that have been familiar elements of the standard model for half a century. Both are components of a psychological model of human decisionmaking. That is, tastes for anticipation and for risk are both exogenously given preferences that explain why different individuals faced with objectively identical choices under uncertainty make different decisions. But these two types of preferences also have significant differences. Risk preferences enter into the decision problem *after* beliefs have been formed, while tastes for anticipation play a role in determining *how* those beliefs form in the first place. In principle, a rational actor can be assumed to have both types of preferences. For the purposes of this paper, preferences are assumed to be risk neutral in order to isolate the consequences of tastes for anticipation.

At $t = 0$, the citizen forms beliefs that influence her later turnout decision. The solution concept is subgame perfect equilibrium. In this example, endogenous beliefs are related to turnout as follows:

Voter Turnout with Endogenous Beliefs.[13] *If $\delta > 0$, the citizen forms the belief $\hat{p}^* = 1$ and votes if and only if $p > c - \delta(1-c)$. Otherwise, the citizen forms beliefs $\hat{p}^* \in [0, c + \frac{1}{\delta}(c-p))$ and does not vote.*

Endogenous beliefs maximize total utility at $t = 0$. Formally, the belief-forming agent compares the total utility from subjective beliefs that lead to voting with that from beliefs that will lead to abstaining. This agent should not be thought of as the player, but is a formalization of the belief formation process. The citizen forms beliefs that lead to voting if and only if $\delta(\hat{p} - c) + p - c \geq 0$, or $\hat{p} \geq c + \frac{1}{\delta}(c - p)$. The right side of this inequality is a lower bound on subjective beliefs that lead to voting. Because these beliefs are probabilities, the largest $\hat{p}$ can be is 1. Therefore, if the right side is larger than 1, no subjective beliefs ever lead to voting. The lower bound is large when the probability of being pivotal is small ($p < c$) and the citizen has very low taste for anticipation (as $\delta \to 0$). On the other hand, if the citizen's taste for anticipation is very large (as $\delta \to \infty$), then the choice does not depend

[13] All proofs are in the appendix.

much on the objective probability that she will be pivotal. In that case, the citizen will vote if $p \geq c - \delta(1-c)$. If she does vote, she forms the most optimistic beliefs possible ($\hat{p} = 1$), which maximizes anticipatory utility. Substitution of these beliefs into the inequality provides the formal conditions under which the voter turns out.

When the citizen does not value anticipation much, outcomes from the EBM match those from the standard model, although in general, endogenous beliefs need not. In fact, for any $\delta \leq \frac{c-p}{1-c}$, no subjective beliefs can compel turnout. However, given that no beliefs cause turnout, any endogenous beliefs are consistent with the model. Thus, $\hat{p}^*$ could be bigger, smaller, or equal to p, meaning that beliefs could be optimistic, pessimistic, or realistic. Thus, the EBM yields the same prediction about turnout but allows for much more variation in possible beliefs.

Voters in the EBM have beliefs about being pivotal that are overly optimistic, yet are rational in the sense that their choices maximize expected utility based on those beliefs. In fact, if we take the common measures of external efficacy—that is, a citizen's belief that she has a say in government—to be a measure of such optimism, then there is a well-documented empirical linkage between overoptimism and turnout (e.g. Abramson and Aldrich 1982). Thus, the EBM provides a theoretical explanation for an existing empirical regularity that would be difficult and awkward to explain within the standard model. This (unanticipated) convergence is exactly the sort of "postdiction" that characterizes theoretical progression in science (Lakatos 1970). Moreover, the EBM makes novel predictions as well. For example, voters who do turn out should also be optimistic when it comes to other political decisions.

Both the EBM and the D term in Riker and Ordeshook's (1968) model explain turnout. Some might suspect that tastes for anticipation are playing a similar role to Riker and Ordeshook's (1968) D term, and that the implications of anticipatory utility might not add up to much beyond the observation that a taste for anticipation leads to optimism. But the two explanations are very different. First, the EBM is not ad hoc. It applies to a very large set of games. Second, the model reconciles intuitive explanations that are hard to derive

from the standard model, and generates empirically valid postdictions in addition to clearly testable novel predictions. The next example verifies both these differences and pushes the EBM into a strategic, game-theoretic context.

Taxation & Endogenous Prospects of Upward Mobility

The voter turnout example is decision-theoretic, meaning that it requires the citizen to ignore other citizens, even those whose choices may influence the outcome she cares about. The next example extends the EBM to a game-theoretic model of collective choice under uncertainty. To illustrate the flexibility and applicability of the EBM, this example is drawn from the extensive literature on income inequality, redistribution, and tax policy (e.g. Iversen and Soskice 2001; Moene and Wallerstein 2001). The example also demonstrates how the standard model can lead us to reject explanations that are easily intelligible in the context of the EBM.

Why would a person who is poor—and who is very likely to remain poor in the future—favor taxation policies that benefit the rich? In other words, when the poor are in the majority, why do they not always expropriate the rich? One answer is that poor citizens choose not to expropriate the rich because they believe they will be rich someday (Okun 1975). Indeed, Alesina and La Ferrara (2005) find that individuals who believe that the opportunities for upward mobility are equally distributed are less likely to prefer income redistribution, even controlling for objective measures of mobility.[14] Although this answer is simple and compelling, it is difficult to fit into a standard model because of the constraint that beliefs must be objective probabilities. Two options emerge: one could either add significant complexity to the standard model in the form of auxiliary assumptions, or toss out the prospects of upward mobility explanation entirely.[15]

[14]Moreover, Alesina and La Ferrara (2005) find evidence that *subjective* beliefs about future prospects are associated with preferences about redistribution. Elsewhere, Alesina, Glaeser, and Sacerdote (2001) show that much of the difference between the American and European welfare states can be explained by differences in subjective beliefs.

[15]There are other explanations of under-expropriation. One could explain under-expropriation as a result of tax inefficiencies (Meltzer and Richard 1981), the risk of coups (Acemoglu and Robinson 2001), or biases

Many scholars study redistribution using the standard model and the median voter theorem for collective choice in one dimension (Black 1958). In the simplest such model, if a majority of voters are poor and likely to remain so, the poor are predicted to expropriate the rich. It is possible to explain the observed under-expropriation with prospects of upward mobility by significantly amending this standard model. But such an explanation requires a long list of auxiliary assumptions. Benabou and Ok (2001) show that if (1) tax policies are sufficiently long lasting, (2) taxes are not too nonlinear (i.e., they are not graduated by income), and (3) objective, expected, future income is a concave function of present income (i.e., increases in one's future income are likely to be larger for those with lower incomes than those with higher incomes), then prospects of upward mobility emerge from the standard model as an explanation of under-expropriation. However, each of these assumptions is strong and empirically implausible. Tax policies change frequently, and progressive taxation is common (see e.g., Steinmo 1996). The concavity assumption requires that the poor must *objectively* expect to see larger increases in future income than the rich, which is at least a strong requirement. And, to the extent that tax rates are impermanent or nonlinear, the objective, expected, future income function must be even more concave for the result to hold.

These auxiliary assumptions can be interpreted in two ways. On the one hand, the assumptions indicate that one can fit prospects of upward mobilities and under-expropriation into the standard model. On the other hand, these assumptions highlight how the standard model must be altered to account for straightforward causal mechanisms. If any of these assumptions is violated, then either we must abandon prospects of upward mobility as an explanation of under-expropriation, or we must amend the standard model to include this explanation. The latter is the goal of the following example.

Consider a society of n (odd) citizens who collectively choose a tax rate $\tau \in [0, 1]$. Based

in political systems (Putterman, Roemer, and Silvestre 1998). Alternately, preferences for redistribution may be more accurately viewed as preferences for social insurance (Iversen and Soskice 2001; Moene and Wallerstein 2001). Although each explanation has merit, none captures the idea that prospects of upward mobility drive under-expropriation. The goal here is to see how one might do this in the standard model and in the EBM.

on average future income $\bar{y}$, tax revenues will generate a lump-sum payment in services to each citizen. Lump-sum transfers are defined by the budget constraint to be $T = \tau\bar{y}$. Citizens collectively choose the tax rate τ according to an open rule.[16]

There is uncertainty about future income in this example. Objectively, each citizen i belongs to one of two types: the likely rich and the likely poor. Citizens who are likely rich can objectively expect to earn income y_r, and the citizens who are likely poor can objectively expect to earn income $y_p < y_r$.[17] Assume there are more likely poor citizens than likely rich, so the number of likely poor citizens is $m > \frac{n}{2}$. Play in the game proceeds as follows: all citizens form beliefs about future income (at $t = 0$), collectively choose tax policy under an open rule (at $t = 1$), anticipate the resolution of uncertainty about future income (at $t = 2$), and eventually earn income, pay taxes, and garner transfers (at $t = 3$).

First, consider the tax policy preferences of citizens at $t = 1$, based on their (subjective) expectations about future income. In both the standard model and the EBM, each citizen i has beliefs over future income represented by her subjective expectations $\hat{y}_i$. In the standard model, these subjective expectations must match objective expectations ($\hat{y}_i = y_i$); in the EBM, any $\hat{y}_i \in [y_p, y_r]$ is possible. Regardless of how she formed those beliefs, i uses $\hat{y}_i$ to inform her preferences over tax rates. That is, given $\hat{y}_i$, her subjective expected utility over τ is $(1 - \tau)\hat{y}_i + \tau\bar{y}$, which is a linear combination of her own income (minus taxes) plus the benefits from average tax revenues. This utility function is standard in the literature (see e.g., Iversen and Soskice 2001). The utility is linear in τ, which means that each citizen's preferred tax rate will be a corner solution. That is, i's most-preferred tax rate is either $\tau = 0$ if she expects to be sufficiently rich ($\hat{y}_i > \bar{y}$) or $\tau = 1$ if she expects to be sufficiently

[16] When beliefs conform to the standard model, Iversen and Soskice (2001) refer to this example as, "the $t = 1$ model," where t refers to tax rates. To clarify the explanatory purchase of endogenous beliefs, I assume that there are no inefficiencies (e.g., deadweight losses) associated with taxation and that individuals cannot choose their levels of labor endogenously. Each of these additional features could certainly be added once analysis of the simpler example is complete, but they are not needed for the results established here.

[17] In this example, expectations of future income are sufficient statistics for beliefs. Therefore, I leave objective and subjective probabilities implicit.

poor ($\hat{y}_i \leq \bar{y}$).[18] Of course, in reality, other considerations limit taxes to be between these two extremes. For the present purposes, $\tau = 0$ can be thought of as a low tax, under-expropriation regime and $\tau = 1$ can be thought of as a high tax, high-expropriation regime.

In the standard model, all citizens have objective expectations about their future incomes. In this case, there is a unique equilibrium tax rate of $\tau^* = 1$, meaning that the likely poor always expropriate the likely rich. This result follows because there are more likely poor than likely rich, because the median voter is likely to be poor, and because the median voter believes she will be poor.

Now consider the example in the context of the EBM. All citizens now form subjective expectations $\hat{y}_i$ about income at $t = 0$. Each citizen's beliefs are modeled as the action of a belief-forming agent who maximizes total utility. For citizen i, if the tax rate is τ, then anticipatory utility is $(1 - \tau)\hat{y}_i + \tau\bar{y}$ and outcome utility is $(1 - \tau)y_i + \tau\bar{y}$. Total utility at $t = 0$ is therefore given by

$$\delta((1 - \tau)\hat{y}_i + \tau\bar{y}) + (1 - \tau)y_i + \tau\bar{y}.$$

Unlike the standard model, endogenous prospects of upward mobility emerge in the EBM under certain conditions.

Conditions for Endogenous Prospects of Upward Mobility. *If* $\delta > \delta^* := \dfrac{\bar{y} - y_p}{y_r - \bar{y}}$, *then there is an equilibrium in which the tax rate is* $\tau^* = 0$. *Otherwise,* $\tau^* = 1$.

In the low-tax, under-expropriation case ($\delta > \delta^*$), the benefits represented by anticipatory utility outweigh the adverse consequences represented by outcome utility, and every citizen i forms the most optimistic expectations possible ($\hat{y}_i = y_r$). Alternatively, in the high-tax, high-expropriation case ($\delta \leq \delta^*$), a sufficient number of the likely poor form realistic beliefs and the result is a high tax rate. More specifically, enough poor citizens to constitute a bare majority of ($\frac{n+1}{2}$) form relatively low expectations (any $\hat{y}_i \leq \bar{y}$) about their future incomes.

[18] Assume that if a citizen is indifferent, she chooses the expropriating tax rate $\tau = 1$.

The rest of the likely poor ($m - \frac{n+1}{2}$ citizens) form overly optimistic expectations ($\hat{y}_i = y_r$), experience the higher anticipatory utility as well as the likely benefits of high taxes. The likely rich form realistic beliefs ($\hat{y}_i = y_r$) because their anticipatory and outcome utilities yield no tradeoff. Thus, if any of the likely poor citizens with low expectations instead forms beliefs that her future income will be large, the tax rate would be lower.

To understand these two cases, consider each citizen's beliefs in turn. First, the likely rich citizens always form objective expectations. For the likely rich, the incentives from anticipatory utility and outcome utility are aligned. If they were to lower their subjective expectations, they would lose anticipatory utility *and* vote for tax rates that could lead to a loss of income and outcome utility. Thus, $\hat{y}_i = y_r$ for all of the likely rich regardless of δ.

Next, consider the likely poor. Each of the likely poor forms one of two possible beliefs. First, she may form the belief that she is likely to be so poor as to prefer a high tax rate at $t = 1$. In this case, she will expect low anticipatory utility, but high outcome utility, and her total utility will be $\delta\bar{y} + \bar{y}$. This belief leads to high taxation and high income redistribution, so that everyone earns the average income level $\bar{y}$. Alternatively, a likely poor citizen may believe that she will be so upwardly mobile ($\hat{y}_i > \bar{y}$) as to prefer a low tax rate at $t = 0$. The total utility for this more optimistic belief is $\delta y_r + y_p$, as it leads to low taxes, under-expropriation, and no income redistribution. The relevant comparison is between the total utilities associated with each belief. A bit of algebra yields the condition for endogenous prospects of upward mobility.

The conditions revolve around the boundary δ^*, which is always greater than 0. If citizens do not value anticipation much (δ near 0), the outcome is the high tax, high expropriation, just as it is in the standard model. As in the decision-theoretic voter turnout example, actors in this example form endogenous beliefs that match objective probabilities when anticipation is not a factor.

The relationship between endogenous prospects of upward mobility and under-expropriation depends on the objective expected income levels and average income. As δ^* increases, the

conditions for under-expropriation become more stringent. Therefore, we can examine the partial derivatives of the boundary δ^* with respect to each variable to determine whether an increase in that variable has a direct effect on endogenous prospects of upward mobility and under-expropriation.

Interestingly, the conditions for under-expropriation become less stringent as either y_p or y_r increases and more stringent as $\bar{y}$ increases. Holding average income constant, the boundary δ^* decreases in both y_r and y_p.[19] Therefore, if either y_r or y_p increases, under-expropriation should become more common. If y_r increases, then the anticipatory utility from optimistic expectations increases, thus making under-expropriation more preferable. If y_p increases, the outcome utility from optimistic expectations increases, also making under-expropriation more preferable. Conversely, if both y_r and y_p are held constant, δ^* increases in average income $\bar{y}$.[20] Hence, as $\bar{y}$ increases, under-expropriation should become less common. Increases in $\bar{y}$ lead to increases in both the anticipatory and outcome utility of objective beliefs, making expropriation preferable for smaller values of δ.

In the standard model, a set of very strong auxiliary assumptions is required for prospects of upward mobility to drive the under-expropriation we observe empirically. One might reason that such prospects cannot be the real causal force at work, that instead some other reason must be driving under-expropriation. However, in a simple example with endogenous beliefs, prospects of upward mobility emerge and explain under-expropriation. Thus, one should not be so quick to jettison an explanation like prospects of upward mobility simply because it is difficult to reconcile with the standard model. Moreover, this example yields another postdiction: all else equal, increases in average income should lead to more redistribution via taxes. Indeed, in their longitudinal study of tax rates in OECD countries, Swank and Steinmo (2002) find that tax rates of capital (as opposed to labor) increase on average in periods that follow economic growth. This example also makes a novel prediction: holding

[19] Holding $\bar{y}$ constant, the partial derivative of δ^* with respect to y_r is $\frac{\partial \delta^*}{\partial y_r} = -\frac{\bar{y}-y_p}{(y_r-\bar{y})^2} < 0$. Similarly, the partial derivative of δ^* with respect to y_p is $\frac{\partial \delta^*}{\partial y_p} = -\frac{1}{y_r-\bar{y}} < 0$.

[20] Holding both y_p and y_r constant, the partial derivative of δ^* with respect to $\bar{y}$ is $\frac{\partial \delta^*}{\partial \bar{y}} = \frac{y_r-y_p}{(y_r-\bar{y})^2} > 0$.

average income constant, increases in either the low or high end of the income distribution should lead to lower taxes and less expropriation.

Crisis Bargaining & Endogenous Mutual Optimism

In the voter turnout example, strategy plays no role in the formation of endogenous beliefs, and, in the taxation example, the role of strategy in belief formation is similar to the role of strategy of votes in the median voter theorem. As a consequence, the effects of tastes for anticipation have been simple and straightforward. In both examples, greater tastes for anticipation lead to more optimistic beliefs. In other settings, the strategic consequences of belief formation can lead to subtle relationships between objective probabilities and endogenous beliefs, even in the absence of tastes for anticipation. The next example presents such a case to address the role of optimism in conflict and crisis bargaining. This example also emphasizes how the strictures of the standard model can lead us to disregard parsimonious explanations that depend on optimistic beliefs.

Why do some states engage in costly wars rather than resolve disagreements through peaceful settlement? Blainey (1988) argues that a crucial clue to understanding conflict "is the optimism with which most wars [are] commenced" (p. 35). Other scholars who write detailed case studies on the origins of wars have also recognized and diagnosed the prevalence of optimism in international affairs. Van Evera (1999) describes a plethora of cases in which false hope and illusions of victory lead to conflict. Snyder (1991) documents how powerful nations can hurt themselves by falling prey to self deception by cultivating rosy scenarios. And Johnson (2004) documents the prevalence of overconfidence in conflict and identifies the potential benefits of overly positive thinking.

But Fearon (1995) questions the origins of such inconsistent expectations in rational accounts of conflict, writing"[H]ow can rationally led states have conflicting expectations about the likely outcomes of military conflict? In the extreme case, how could both sides rationally expect to win?" (p. 391). Fearon correctly observes that, within the standard

model, the only possible explanation for inconsistent expectations about the outcome of conflict is private information. If both states have sufficiently inaccurate private information, their expectations can lead to war—given an additional auxiliary assumption. The game Fearon analyzes satisfies this assumption, which Fey and Ramsay (2007) go on to isolate. The required assumption is that the aggressive party in the crisis must be able to *commit* to starting a conflict for conflict to be an equilibrium outcome. If instead both states must prefer conflict in order for a conflict to be an equilibrium outcome, states who are mutually optimistic because of private information cannot both want to start a conflict. In the standard model, inconsistent beliefs about the probabilities of victory cannot survive the self-evident, public, common-knowledge-producing event of conflict. If a player perceives that her opponent is willing to engage her in conflict and updates logically, she must infer that her opponent knows something that she does not. Fey and Ramsay conclude that we should look elsewhere for rational explanations of conflict.

In contrast to the standard model, differing expectations about the probabilities of victory need not be informative in the EBM. For example, the "above average effect" (e.g., Svenson 1981), so-named because most people profess to be of above average ability, is very well known. Awareness of the above average effect may then lead one to doubt self-professions of high ability. That is, even if one person knows that another believes himself to be above average, she need not revise down her beliefs about her own relative ability. If beliefs are subjective, a person cannot necessarily infer that another's actions reveal anything vital about private information.

Consider a simple crisis bargaining game, in which an aggressive state A demands some $x \geq 0$ from a second defending state B.[21] The latter chooses whether to pay. If B pays A the amount x, B keeps $1 - x$, and the game ends. If B chooses not to pay, A decides whether to initiate conflict against B. If A does not do so, the game ends. If A does initiate conflict, the outcome is determined probabilistically. With objective probability p_A, A wins the conflict.

[21] I use female pronouns for A and male pronouns for B. I also assume that if a state is indifferent between conflict and settlement, he or she chooses settlement.

Otherwise, with objective probability $p_B = 1 - p_A$, B wins. The winner receives payoff 1, and the loser receives payoff 0. In any case, conflict costs both states $c > 0$. Strategies in this game are given by the amount x that A demands, the largest acceptable demand $\bar{x}$ that B is willing to pay, and a choice a by A about whether to initiate conflict if the demand is rejected, where $a = 1$ if A initiates conflict and $a = 0$ otherwise. In the framework of the EBM, states form subjective beliefs about their probabilities of victory, $\hat{p}_i$, at $t = 0$, play the above game during $t = 1$, experience anticipation at $t = 2$, and experience outcomes at $t = 3$.

As in the previous examples, I focus first on choices based on any possible, given subjective beliefs at $t = 1$. Each state $i \in \{A, B\}$ has formed beliefs given by $\hat{p}_i$, the subjective probability that i will win an eventual conflict. Because these probabilities may match the objective probabilities, this reasoning applies to both the standard model and the EBM. States rely on their beliefs when they choose their actions, and backward induction can be used to find the subgame perfect equilibrium given subjective beliefs. Because this example is a bit more complicated than the previous two, it helps to explicitly analyze actions in the crisis bargaining game for any given pair of subjective beliefs $(\hat{p}_A, \hat{p}_B)$.

Subgame Perfect Equilibrium of the Crisis Bargaining Game. *Given subjective beliefs, every equilibrium of the crisis bargaining game is characterized by*
(i) a non-credible threat $(\text{any } x^*, \bar{x}^* = 0, a^* = 0)$ *if* $\hat{p}_A \leq c$,
(ii) settlement $(x^* = \bar{x}^* = 1 - \hat{p}_B + c, a^* = 1)$ *if* $\hat{p}_A > c$ *and* $\hat{p}_A + \hat{p}_B \leq 1 + 2c$, *and*
(iii) conflict $(x^* > \bar{x}^* = 1 - \hat{p}_B + c, a^* = 1)$ *if* $\hat{p}_A + \hat{p}_B > 1 + 2c$.

This equilibrium characterization describes what states will do given any particular pair of beliefs, not the beliefs that will be formed. Given their beliefs, A and B can find themselves with three different equilibrium outcomes. Figure 2 depicts these outcomes for the cost of conflict $c = 0.2$. If A believes that she will win with very low probability ($\hat{p}_A \leq c$), no demand she issues will be backed by a credible threat, and B will therefore refuse to pay any demand. Instead, if A believes she will win with high enough probability ($\hat{p}_A > c$),

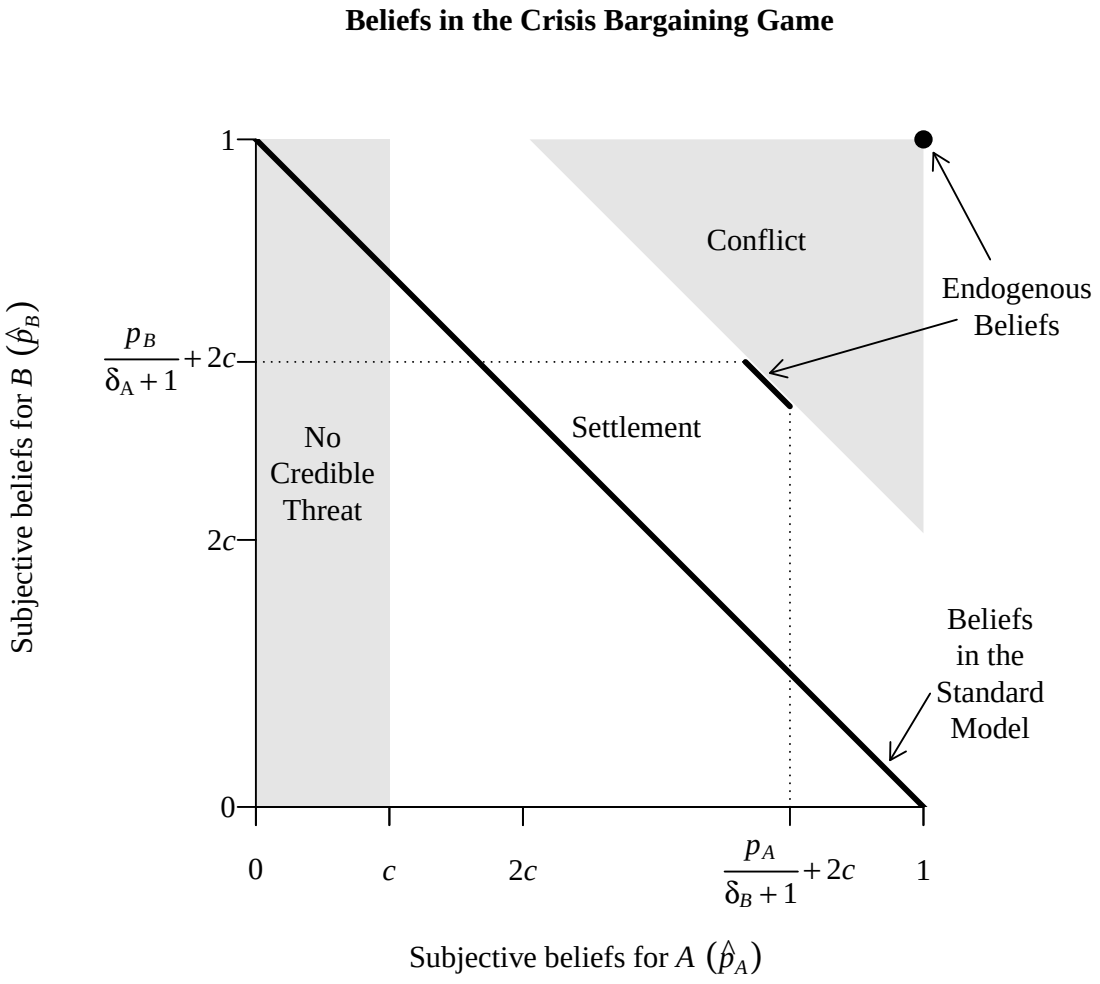

Figure 2: **Beliefs in the Crisis Bargaining Game.** For any pair of subjective beliefs, the figure displays the equilibrium outcome. In the region on the left, A's demands are backed by a non-credible threat; in the central region the two states settle; and in the top right corner, conflict ensues in equilibrium. In the standard model, beliefs must be objective and therefore lie on the bold line running down the diagonal. Thus, conflict does not occur in the standard model. In the EBM, the two states always have optimistic beliefs marked by the bold line on the boundary between the Settlement and Conflict regions and the point in the upper right corner, leading to conflict only in the latter case.

the game ends either in settlement or in conflict. The determining factor is the level of mutual optimism between A and B. If the sum of their subjective beliefs is large enough ($\hat{p}_A + \hat{p}_B > 1 + 2c$), the two will be unable to agree on any demand, and the result will be conflict.

In the standard model, beliefs are exogenous and match objective probabilities ($\hat{p_A} = p_A, \hat{p_B} = p_B$, and so $\hat{p_A} + \hat{p_B} = 1$). In Figure 2, these beliefs are found on the diagonal line that stretches from the top left corner to the lower right corner. The possible equilibrium outcomes in the standard model are a non-credible threat (if $p_A \leq c$) or settlement (otherwise). Conflict cannot occur because it requires a significant threshold of mutual optimism ($\hat{p}_A + \hat{p}_B > 1 + 2c$), which is strictly prohibited by the standard model (since $\hat{p}_A + \hat{p}_B = 1$). Thus, this example illustrates how mutual optimism is inconsistent with conflict in the standard model. In fact, conflict simply cannot occur in this game within the standard model.

In contrast, in the EBM, A and B form beliefs (at $t = 0$) about what will happen during crisis bargaining (at $t = 1$). At $t = 0$, each belief-forming agent takes into account how beliefs will affect both on their anticipated probability of victory (at $t = 2$) in the case that conflict occurs, and the effect that beliefs have on the actual outcome (at $t = 3$). If the subjective beliefs would lead to settlement, total utilities would be $(\delta_A + 1)x^*$ for A and $(\delta_A + 1)(1 - x^*)$ for B. If the subjective beliefs lead to conflict, total utilities would be $\delta_i(\hat{p}_i - c) + p_i - c$ for both $i \in \{A, B\}$. Again using backward induction, one can find conditions in which mutual optimism occurs endogenously and leads to conflict.

Conditions for Endogenous Mutual Optimism. *If* $1 - 2c \leq \frac{p_A}{\delta_B + 1} + \frac{p_B}{\delta_A + 1}$, *the outcome will be settlement* ($\hat{p}_A^* \in [1 - \frac{p_B}{\delta_A + 1}, \frac{p_A}{\delta_B + 1} + 2c], \hat{p}_B^* = 1 - \hat{p}_A^* + 2c$). *Otherwise, the outcome will be conflict* ($\hat{p}_A^* = \hat{p}_B^* = 1$).

In addition to the equilibria for all possible pairs of subjective beliefs, Figure 2 also presents some examples of endogenous beliefs.[22] If costs are very high or if tastes for anticipation are

[22] In the figure, $\delta_A = \delta_B = \frac{1}{2}$ and p_A. Therefore, conflict results in equilibrium and endogenous beliefs are $\hat{p}_A = \hat{p}_B = 1$. The other endogenous beliefs depicted are for illustrative purposes.

very low ($1 - 2c \leq \frac{p_A}{\delta_B+1} + \frac{p_B}{\delta_A+1}$), settlement is the outcome of the game, and this outcome is shown in the figure by the bolded line segment on the border between settlement and conflict. In this case, states maintain mutually optimistic beliefs. If costs are very low and if tastes for anticipation are high ($1 - 2c \leq \frac{p_A}{\delta_B+1} + \frac{p_B}{\delta_A+1}$), conflict is the outcome of the game, and this outcome is shown in the figure by the dot in the upper right corner. Here, both states form the most optimistic beliefs possible because conflict is inevitable and such beliefs maximize anticipatory utility.[23]

Unlike in the previous two examples, the absence of tastes for anticipation ($\delta_A = \delta_B = 0$) does not necessarily lead to objective beliefs in the crisis bargaining game. Here, the states form mutually optimistic beliefs ($\hat{p}_A^* \in [p_A, p_A + 2c]$ and $\hat{p}_B^* = 1 - \hat{p}_A^* + 2c$), although these beliefs will not be sufficiently optimistic to engender conflict. These beliefs emerge because of their strategic value. In particular, the maximum demand that the defending state B is willing to pay depends on his beliefs. He is willing to pay less if his beliefs about his probability of victory are greater. Since state A demands the maximum that B is willing to pay, B gains outcome utility from optimistic beliefs. Thus, B's endogenous beliefs are inflated (by $2c$), which lowers what A will demand, but not so much that it will lead to conflict. Similarly, A's optimistic beliefs undergird credible threats, even if her probability of victory is quite low. In this example, anticipatory utility is not necessary to induce optimism.

The example also implies a complex relationship between the balance of power and the probability of conflict. On the one hand, when the two states are evenly matched in the example ($p_A = p_B = \frac{1}{2}$), an increase in the taste for anticipation of either state makes conflict more likely. On the other hand, when the two states are unbalanced, conflict depends much more on the taste for anticipation of the weaker state than that of the stronger state. For example, suppose that p_A is near 1 and p_B is near 0. In this case, the defending state B's

[23] Again, as a simplification, a corner solution is found. In reality, this equilibrium might best reflect circumstances under which mutual optimism makes conflict more likely rather than certain. The point of the example is to show that the intuition can be more easily expressed with the EBM than with the standard model.

taste for anticipation δ_B is in the denominator of the $\hat{p}_A^*$ term, and thus an increase in δ_B does more to increase $\hat{p}_A^*$ than an increase in δ_A would. The reason is that higher tastes for anticipation make the weaker state more likely to forego settlement. In this example, both states must prefer conflict for conflict to occur in equilibrium. When the balance of power tips in one direction, the weaker state must have a higher taste for anticipation to offset the relative difference in objective probabilities.

The crisis bargaining game shows how mutual optimism can lead to conflict. If beliefs are not required to be objective, then states cannot learn about their objective chances of victory by observing their opponents. Instead, conflict depends on the factors that lead states to form their beliefs. Conflict occurs if tastes for anticipation are large and if the probabilities of victory are sufficiently balanced. Moreover, there is a tradeoff between these tastes for anticipation and the balance of power. To the extent that the probabilities of victory are unbalanced, conflict depends on the weaker state having a higher taste for anticipation; and vice versa.

But the most notable feature of this example is that it provides a formal explanation for how mutual optimism can emerge endogenously and foster conflict. The example demonstrates how the standard model can us lead to reject a familiar and highly intuitive idea because of the strong and empirically unsupported assumptions in the standard model. The standard assumptions of exogenous, objective beliefs and Bayesian updating quash the mutual optimism that is allowed in the EBM, despite our everyday experience with cognitive biases like the above average effect. Because mutual optimism is inconsistent with conflict according to the standard model, one must infer *either* that this explanation of conflict is simply wrong *or* that the standard model cannot adequately represent the explanation.

General Lessons about Endogenous Beliefs

These three examples illustrate several ways in which the EBM can alter models of politics. From these examples, we can distill several general lessons about the relationship between the

EBM and the standard model. First, consider what players must know about the game and each other's beliefs to play equilibrium in the two models—that is, consider their epistemic foundations. In both models, once beliefs have been formed, those beliefs must be assumed to be mutually known for players to play equilibrium.[24] That is, players are assumed to know each other's beliefs, even though players in the EBM may disagree with each other about the likely consequences of their actions. Essentially, the fact that players may agree to disagree about those consequences does not affect their ability to play equilibrium. They may believe each other to be mistaken, yet still have mutually known (and correct) conjectures about how others will play.

Just as the standard model and the EBM both require mutual knowledge of beliefs, equilibrium predictions based on the strongest conceptions of rationality are unaffected by the EBM. Consider the role of endogenous beliefs in the face of a strictly dominated strategy. For example, in the turnout example if the cost of voting is prohibitively high, voting is strictly dominated by not voting, even when a citizen is pivotal. There is no outcome in which any beliefs would compel voting in this case, and so the equilibria of the standard model and of the EBM agree. In general, if there exists a strictly dominated strategy in the standard model, then it must be strictly dominated in the EBM, and it will never be a best response in either model. Moreover, any game that has a unique strictly dominance solvable equilibrium in the standard model will have the same equilibrium in the EBM.[25] However, the same cannot be said for weakly dominated strategies. Indeed, weakly dominated strategies are primarily differentiable from strictly dominated strategies because they are best responses to *some* beliefs. In fact, some games that have a unique weakly dominance solvable equilibrium

[24] Mutual knowledge differs from the stronger more frequently encountered notion of common knowledge. The former merely demands that each player knows each other player's beliefs, and not the meta-requirements that each player knows that every other player knows those beliefs, etc. Aumann and Brandenburger (1995) show that, in the context of a common prior (i.e., in the standard model) mutual knowledge of payoffs, rationality, and conjectures about how others will play is sufficient for equilibrium play. VandenSteen (2001) shows that these requirements also suffice for equilibrium play when players have heterogeneous beliefs about payoffs, as is the case in the EBM.

[25] I gratefully acknowledge an anonymous reviewer for pointing this out.

within the context of the standard model have multiple equilibria in the EBM.[26]

This is not to say that the EBM necessarily multiplies the number of equilibria found in the standard model. If so, then the EBM would necessarily exacerbate the multiple equilibria problem familiar from the standard model. Instead, the EBM can sometimes change the set of equilibria entirely, at times even resulting in a different unique equilibrium predictions in the two models. For example, consider the crisis bargaining game. In the standard model, no equilibrium involves conflict, regardless of parameter values. However, in the EBM with large tastes for anticipation, the unique equilibrium prediction is conflict. Moreover, there are cases in which a unique equilibrium prediction based on the standard model is at odds with a unique equilibrium prediction based on the EBM regardless of tastes for anticipation.[27]

Although the EBM adds value in several ways, these benefits hardly mean that the standard model should simply be replaced. A wiser course of action would be to start with the standard model and then move to the EBM. Any predictions that do not change with the move may be said to be "robust" to endogenous beliefs. And those that do change should prompt more careful thought about the potential for self-serving beliefs and tastes for anticipation in the decisionmakers who are being modeled, which will then inform which model is more appropriate. Moreover, the two models then offer competing predictions that can be tested against each other.

Conclusion

I have presented a flexible, tractable, model of endogenous beliefs and applied it to several different political environments. My goal has not been to add complexity for complexity's sake. Indeed, in the taxation and crisis bargaining examples, the standard model cannot adequately capture its intended phenomenon without even more complexity, either by adding auxiliary assumptions or amending assumptions about beliefs. Rather, what I have shown is that changing one's assumptions about beliefs—in this case, bringing those assumptions

[26] See the Online Appendix for an example.

[27] See the Online Appendix for an example.

in line with psychological and economic theory, experimental evidence, and even everyday common sense—can lead to dramatic differences in the inferences that one can draw. In keeping with the desire to avoid undue complexity, the above examples are all as simple as possible. Yet, they show that the standard model and its implausible assumptions can lead us to reject *by fiat* otherwise simple, intuitive, valid arguments.

In addition to making belief formation endogenous, the Endogenous Beliefs Model (EBM) introduces a new kind of taste that to formal models in political science: the taste for anticipation. These tastes are empirically evident in human behavior, theoretically antecedent to the familiar phenomenon of motivated reasoning, and analytically analogous to risk preferences, and they constitute an incentive to form optimistic beliefs. In each example, tastes for anticipation capture the tradeoff between the experience of uncertainty resolution and the consequences of unrealistic beliefs. In each example, these tastes interact with the objective probabilities to explain when and why phenomena like turnout, prospects of upward mobility, or mutual optimism occur. Future work should focus on how actors who have different risk preferences and tastes for anticipation might behave in political environments.

The examples presented here showcase how the EBM can even be more parsimonious than the standard model. In the tax policy example, the standard model requires several very strong auxiliary assumptions for prospects of upward mobility to explain why the poor do not expropriate the wealthy. Thus, either this explanation is worthless (despite evidence to the contrary) and should be discarded, or the standard model's assumptions about beliefs are not sufficiently flexible to capture the logic of the explanation. Similarly, in the case of crisis bargaining, the standard model suggests that mutual optimism is inconsistent with conflict. Again, either this explanation of conflict is not useful, or the standard model is not adequate to represent the logic of the argument.

Not only do these examples show the applicability of the alternative model to different areas of political science, but they also show that the model can be applied to decision-theoretic environments, static games with simultaneous moves, and dynamic games with sequential

moves. The move from decision problems to static games and then to dynamic games shows that allowing for endogenous beliefs can yield interesting and subtle implications as beliefs begin to play a strategic role. For example, in the crisis bargaining game, endogenous beliefs never match objective probabilities because both players can benefit strategically by maintaining unrealistic optimism.

The EBM has many desirable features. It is tractable, flexible, and can be applied to a large class of interesting cases other than the three examples illustrated here. The model adds complexity in a limited, disciplined way that captures an extremely general, well-replicated, and important set of stylized facts. Moreover, the EBM also makes comparison with the standard model easy, which means that it is straightforward to ascertain whether the amended assumptions add value or merely a superfluous layer of complexity. In fact, the EBM preempts the need for ad hoc complexity in the form of the additional strong assumptions required to make phenomena like optimism intelligible within the standard model. Finally, the EBM does a better job than does the standard model at conforming to the empirical regularities that psychologists and scholars of judgment and decisionmaking have documented.

Moreover, it is easy to see how this setup can be extended to include different mechanisms. These examples show how the model can applied to complete information environments, leaving incomplete information and learning to be incorporated in the future. One way to do so would be to follow a similar timeline to that of the EBM. An actor can act as if she first updates her beliefs objectively according to Bayes' rule and then form new endogenous beliefs based on that information and her incentives represented by anticipatory and outcome utility. Similarly, future work can and should extend this approach to include other emotional incentives. An actor might form beliefs with an eye to averting future disappointment. Fear of poor outcomes could encourage investment in tasks that promise tangible benefits in the future. And anxiety is known to lead actors to spend more time investigating and deliberating over their alternatives (MacKuen et al. 2007).

Many scholars remain skeptical of the value that rational choice and formal theory can add to the study of politics. For some, this skepticism may not owe to the method itself, but to the assumptions made in the standard model that can leave it brittle and unnecessarily inflexible. Here, I have presented an alternative model that maintains many of the standard model's assumptions, utilizes many of the same tools, and preserves the precision and clarity that makes formal theory so valuable. The alternative is based on assumptions that better incorporate much of the empirical evidence about belief formation. It allows one to see how the strong and under-motivated assumptions of the standard model lead directly to conclusions that contradict much existing scholarly work, never mind everyday experience. And it is tractable. When systematic evidence contradicts the foundational assumptions of a theory, when those assumptions lead us to reject seemingly reasonable conclusions, and when there is hope for the development of a simple, tractable, and useful alternative, it is time to add that alternative to our formal toolkit.

References

Abramson, Paul R., and John H. Aldrich. 1982. "The Decline of Electoral Participation in America." *American Political Science Review* 76(3): 502–521.

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

Akerlof, George A., and William T. Dickens. 1982. "Economic Consequences of Cognitive Dissonance." *American Economic Review* 72(3): 307–319.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Preferences for redistribution in the land of opportunities." *Journal of Public Economics* 89(5-6): 897 – 931.

Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote. 2001. "Why Doesnt the United States Have a European-Style Welfare State?" *Brookings Paper on Economics Activity* 2: 187–278.

Anscombe, F. J., and R. J. Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics* 34(1): 199–205.

Aumann, Robert, and Adam Brandenburger. 1995. "Epistemic Conditions for Nash Equilibrium." *Econometrica* 63(5): pp. 1161–1180.

Bargh, John A., and Tanya L. Chartrand. 1999. "The Unbearable Automaticity of Being." *American Psychologist* 54(7): 462–479.

Baron, David P. 1993. "Government Formation and Endogenous Parties." *American Political Science Review* 87(1): 34–47.

Bell, David E. 1985. "Disappointment in Decision Making under Uncertainty." *Operations Research* 33(1): 1–27.

Benabou, Roland, and Efe A. Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *Quarterly Journal of Economics* 116(2): 447–487.

Black, Duncan. 1958. *The Theory of Committees and Elections.* Cambridge: Cambridge University Press.

Blainey, Geoffrey. 1988. *The Causes of War.* 3rd ed. New York: The Free Press.

Braman, Eileen, and Thomas E. Nelson. 2007. "Mechanism of Motivated Reasoning? Analogical Perception in Discrimination Disputes." *American Journal of Political Science* 51(4): pp. 940–956.

Brunnermeier, Markus K., and Jonathan A. Parker. 2005. "Optimal Expectations." *American Economic Review* 95(4): 1092–1118.

Bullock, John G. 2009. "Partisan Bias and the Bayesian Ideal in the Study of Public Opinion." *Journal of Politics* 71(July): 1109–24.

Caplin, Andrew, and John Leahy. 2001. "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics* 116(1): 55–79.

Carrubba, Clifford J. 2009. "A Model of the Endogenous Development of Judicial Institutions in Federal and International Systems." *Journal of Politics* 71(1): 55–69.

Dickson, Eric S. 2006. "Rational Choice Epistemology and Belief Formation in Mass Politics." *Journal of Theoretical Politics* 18(4): 454–497.

Edwards, Ward. 1982. "Conservatism in Human Information Processing." In *Judgment under Uncertainty*, ed. Pail Slovic Daniel Kahneman, and Amos Tversky. Cambridge University Press.

Elster, Jon, ed. 1987. *The multiple self.* Cambridge University Press.

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

Fey, Mark, and Kristopher W. Ramsay. 2007. "Mutual Optimism and War." *American Journal of Political Science* 51(4): 738–754.

Greif, Avner, and David D. Laitin. 2004. "A Theory of Endogenous Institutional Change." *American Political Science Review* 98(4): 633–652.

Hafer, Catherine, and Dimitri Landa. 2007. "Deliberation as Self-Discovery and Institutions for Political Speech." *Journal of Theoretical Politics* 19(3): 329–360.

Iversen, Torben, and David Soskice. 2001. "An Asset Theory of Social Policy Preferences." *American Political Science Review* 95(4): 875–893.

Johnson, Dominic D. P., and James H. Fowler. 2011. "The evolution of overconfidence." *Nature* 477(September): 317–320.

Johnson, Dominic D.P. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press.

Kreps, David M. 1988. *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

Kuhn, H. W. 1953. "Extensive Games and the Problem of Information." *Contributions to the Theory of Gamess II* pp. 307–317.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3): 480–498.

Lakatos, Imre. 1970. *Criticism and the Growth of Knowledge*. Cambridge University Press chapter Falsification and the Methodology of Scientific Research Programs, pp. 91–196.

Loewenstein, George. 1987. "Anticipation and the Valuation of Delayed Consumption." *The Economic Journal* 97(387): pp. 666–684.

Loomes, Graham, and Robert Sugden. 1986. "Disappointment and Dynamic Consistency in Choice under Uncertainty." *Review of Economic Studies* 53(2): 271–282.

Lupia, Arthur, Adam Seth Levine, and Natasha Zharinova. 2010. "When Should Political Scientists Use the Self-Confirming Equilibrium Concept? Benefits, Costs, and an Application to Jury Theorems." *Political Analysis* 18(1): 103–123.

Lupia, Arthur, and Jesse O. Menning. 2009. "When Can Politicians Scare Citizens Into Supporting Bad Policies?" *American Journal of Political Science* 53(1): 90–106.

Lupia, Arthur., Mathew D. McCubbins, and Samuel L. Popkin. 2000. *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*. Cambridge University Press.

MacKuen, Michael, George E. Marcus, W. Russell Neuman, and Luke Keele. 2007. *The Affect Effect*. University of Chicago chapter The Third Way: The Theory of Affective Intelligence and American Democracy, pp. 124–151.

Mellers, Barbara A., Alan Schwartz, Katty Ho, and Ilana Ritov. 1997. "Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options." *Psychological Science* 8(6): 423–429.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(3): 914–927.

Moene, Karl Ove, and Michael Wallerstein. 2001. "Inequality, Social Insurance, and Redistribution." *American Political Science Review* 95(4): 859–874.

Okun, Arthur M. 1975. *Equality and Efficiency, the Big Tradeoff*. Brookings Institution Press.

Palfrey, Thomas R., and Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice* 41: 7–53.

Piccione, Michele, and Ariel Rubinstein. 1997. "On the Interpretation of Decision Problems with Imperfect Recall* 1." *Games and Economic Behavior* 20(1): 3–24.

Putterman, Louis, John E. Roemer, and Joaquim Silvestre. 1998. "Does Egalitarianism Have a Future?" *Journal of Economic Literature* 36(2): 861–902.

Redlawsk, David P. 2002. "Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making." *Journal of Politics* 64(04): 1021–1044.

Riker, William H., and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62(1): 25–42.

Rottenstreich, Yuval, and Christopher K. Hsee. 2001. "Money, Kisses, and Electric Shocks: On the Affective Psychology of Risk." *Psychological Science* 12(3): 185–190.

Savage, Leonard J. 1972. *The Foundations of Statistics*. New York: Dover Publications.

Schiemann, John W. 2007. "Bizarre Beliefs and Rational Choices: A Behavioral Approach to Analytic Narratives." *Journal of Politics* 69(2): 511–524.

Selten, R. 1975. "Reexamination of the perfectness concept for equilibrium points in extensive games." *International journal of game theory* 4(1): 25–55.

Smith, Alistair, and Allan C. Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6): 783–813.

Snyder, Jack. 1991. *Myths of Empire: Domestic Politics and International Ambition*. Ithaca, New York: Cornell University Press.

Steinmo, Sven. 1996. *Taxation and democracy: Swedish, British and American approaches to financing the modern state*. Yale University Press.

Svenson, Ola. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica* 47(2): 143 – 148.

Swank, Duane, and Sven Steinmo. 2002. "The New Political Economy of Taxation in Advanced Capitalist Democracies." *American Journal of Political Science* 46(3): pp. 642–655.

Taber, Charles S., and Milton Lodge. 2006. "Motivated skepticism in the evaluation of political beliefs." *American Journal of Political Science* pp. 755–769.

Tversky, Amos, and Daniel Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science* 211(4481): 453–458.

Van Evera, Steven. 1999. *Causes of War: Power and the Roots of Conflict*. Ithaca, New York: Cornell University Press.

VandenSteen, Eric J. 2001. Essays on the Managerial Implications of Differing Priors PhD thesis Stanford University Graduate School of Business.

Von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Weinstein, Neil D. 1980. "Unrealistic Optimism about Future Life Events." *Journal of Personality and Social Psychology* 39(5): 806–820.