Skepticism and credulity: a model and applications to political spin, belief formation, and decision weights

James D. Campbell*
Providence College and University of California, Berkeley

March 2, 2021

Abstract

In this paper I model a decision maker who forms beliefs and opinions using a dialectic heuristic that depends on their degree of skepticism or credulity. In an application to political spin, two competing parties choose how to frame commonly observed evidence. If the receiver is sufficiently credulous, equilibrium spin is maximally extreme and generates short, superficial news cycles. When receivers vary in their skepticism, there is partisan sorting by skepticism parameter: the more credulous group systematically favors one party and displays hostility to evidence and a media they see as biased. In behavioral applications in which the frames arise from the decision maker's internal deliberation, a decision maker with the same credulous nature would display known behavioral anomalies in forming beliefs and forming decision weights from stated probabilities. The dialectic model therefore captures a simple psychological mechanism and matches closely some stylized facts across these three disparate applications.

Keywords: political spin, persuasion, belief formation, subjective probability, skepticism, credulity

JEL codes: D91, D81, D83, D72

*Email: jamescampbell@berkeley.edu. I am grateful to Bernhard Ganglmair and Stephanie Tilden for extremely helpful discussions and to two anonymous referees for their invaluable comments. Thanks to seminar participants at Providence College and the 2020 Eastern Economic Association Meetings and to the Berkeley Teach-Net community for their advice and feedback.

1 Introduction

In his dissent in *Abrams v. United States*, Oliver Wendell Holmes wrote that "the best test of truth is the power of the thought to get itself accepted in the competition of the market"[1]. The 'marketplace of ideas' is now a familiar analogy, for example in discussing the market structure of the internet (Nunziato, 2018; Lombardi, 2019), the regulation of speech on college campuses (Roth, 2017), or the speech of public school teachers (Davis, 2004). Advocates of the marketplace suggest that we should permit and welcome the expression of all thoughts—even falsehood, exaggeration, or conspiracy—since the truth will win out.

This argument relies on two assumptions. First, that the marketplace of ideas will resolve around some broadly accepted or dominant idea. Second, that the winning idea, when it is crowned, will be the truth. But in the marketplace as it really is, will we see consensus? Does competition among ideas discover the likeliest story, or is the truth something different that might be overlooked?

In this paper I model a setting in which adversaries compete to influence the outcome of the marketplace of ideas: political spin, the art of selling interpretations of reality to suit one's own purposes. The key feature of the model is that receivers who hear the competing interpretations weigh them up according to a behavioral heuristic that depends on their degree of skepticism or credulity. I show how and why the outcome of the 'marketplace of ideas' can be systematically different from the truly likeliest explanation, in a way that depends on the credulity of the audience.

To provide supporting evidence for this behavioral model (which is primarily inspired by Froeb et al., 2016) and to demonstrate its applicability to internal dilemmas as well as external persuasion, I apply it to the the formation of beliefs and to the formation of decision weights by a decision maker. In these contexts an excessively credulous decision maker displays behavioral anomalies consistent with stylized facts as described in, for example, Benjamin (2018) and Tversky and Kahneman (1992) respectively.

This paper therefore contributes to two strands of literature: persuasion by the construction of narratives (Schwartzstein and Sunderam, 2019; Eliaz and Spiegler, 2020; Bénabou et al., 2018)

[1] Abrams v. United States (1919).

and attempts to find unified psychological motivations for behavioral phenomena in economic settings (for example, Enke and Graeber, 2021 on cognitive uncertainty and Ambuehl and Li (2018) on responsiveness to information as a common factor explaining a set of empirical behavioral regularities, and Rustichini et al., 2016; Stango et al., 2017 and Dean and Ortoleva, 2019 exploring experimentally observed correlations in behaviors). In sum, I propose a tractable model that captures a person's degree of skepticism versus credulity, and suggest that this aspect of a person's psychology can rationalize a set of behaviors across the three applications.

In the political spin application, two political parties choose a frame for some commonly observed evidence about the world. Their goal is to try to influence the opinion of a receiver to be closer to their preferred endpoint on a one-dimensional political spectrum. The frame they select is a uniform probability distribution from which the party claims the evidence was drawn. Each frame must cover the evidence, but can stretch beyond it. For example, one party may try to spin a rise in the unemployment rate as a result of poor policymaking, while their opponent may try to spin the same data as the result of factors beyond their control. One party may spin a high-profile crime as indicative of an epidemic, while the other may spin the same crime as a fluke.

The receiver of political spin in the model processes the frames of the two parties in a particular way. They use a heuristic that takes a weighted average of the means of the two frames, where the weights carry a penalty for implausibility. That means that if it is very unlikely that one party's explanation would have generated the observed evidence, then the receiver discounts that explanation quite heavily. For example, say that a particular model of car is involved in a large number of accidents that are consistent with brake failure. The accidents certainly *could* be a coincidence, but explaining them as coincidence is less plausible than explaining them as the result of a common defect. The receiver would entertain both explanations, but lean toward the one that was more likely to have generated the observed evidence.

To capture potential differences in how credulous or skeptical people are, the degree to which they penalize the more implausible frame is summarized in a single behavioral parameter called *skepticism*. The more skeptical the receiver, the more they penalize the relatively implausible

frame; in the opposite direction, the more credulous the receiver, the less they penalize the relatively implausible frame. In sum, we can see the receiver as a boundedly rational decision maker with a particular behavioral type whose opinions are influenced by the political messages they hear. They are a citizen with limited time or attention to spend thinking about politics who interprets the competing narratives without thinking too hard about the fundamentals.

The political spin game gives us something quite different from Holmes's ideal for the marketplace of ideas. The model predicts political spin from both sides that is—rationally—as extreme as possible in equilibrium. When this is heard and processed by the receiver, we should not expect the 'truth' (in the sense of the maximally likely explanation) to emerge. For a receiver with a behaviorally plausible degree of credulity, we will see them hold opinions that are systematically biased in the direction of the more far-fetched spin.

Further, if the spin is processed by different receivers with different skepticism parameters, then we should not expect consensus after the fact. If the evidence on different issues consisently skews in one direction or another, a predictable partisan sorting emerges in attitudes to evidence. One party's supporters will be excessively credulous of spin and hostile to evidence and the media that reports it, while other party's supporters will be excessively attached to precise explanations and consider themselves to have the facts entirely on their side.

Next I consider some extensions of the main model to allow for ad hoc partisanship and motivated reasoning, and to analyze the incentives for the senders to produce or suppress new evidence. It is always beneficial to a party to release evidence that forces their opponent to expand their frame. The incentive to do so is highest when the existing evidence is narrow in scope and most favorable to the opponent, and the incentive is lower for less contentious issues and for issues with more existing evidence. This incentive structure is consistent with a world of short, superficial news cycles that jump from conflict to conflict.

In short, the political spin model predicts that the 'marketplace of ideas' rewards extremes and exaggeration, does not lead to consensus, and does not lead to truth. However, since the model involves a particular opinion-forming process by the receiver, the question remains of whether this particular approach is a reasonable way to model the evaluation of competing explanations.

Therefore, in order to provide some supporting evidence for the behavioral model I use for the receiver, after considering the political spin game I go on in Section 6 analyze in a bit more detail the model of the credulous or skeptical decision maker in the context of individual choice. In this context, one interpretation of the model is of a decision maker who forms beliefs by thinking dialectically about a best case and a worst case scenario. This is similar to the spin game in the case in which both parties choose the most extreme possible frames, except that the frames here come from introspection rather than an outside sender. The notion of this kind of dialectic introspection has a long history in philosophy and culture, as I briefly summarize below in Section 6.

It turns out that some well-known behavioral anomalies from prior literature are consistent with the predictions of this model of dialectic belief formation. In particular, the skepticism parameter that is associated with maximal spin in the political spin game is similar to the skepticism parameter that best calibrates the dialectic model to those behavioral anomalies: excessive credulity to outlandish explanations. In this sense I argue that while persistent disagreement in the spin game may be viewed as merely a consequence of any old form of heterogeneous, non-Bayesian receivers, the particular form that their information processing takes is a plausible and empirically relevant one.

The two examples I focus on to make this case are as follows:

Example 1: forming beliefs about the world

A person observes noisy signals about the state of the world. An optimistic view would hold that unfavorable signals are just bad luck, while a pessimistic view would hold that unfavorable signals reflect the true, unhappy state of affairs. How does an adjudication between those competing explanations differ from a Bayesian view of the evidence?

In this example, the dialectic model provides a psychological explanation for a person's systematic bias away from Bayesian posteriors. The model implies excessive stubbornness of beliefs and the irrelevance of new evidence to a person's beliefs. I also discuss how a symmetrically distributed skepticism parameter in the population leads to a distribution of beliefs that is asymmetrically skewed away from the Bayesian standard. The model's predictions in this application are consistent with our best understanding of evidence on belief updating, as surveyed in, for

example, Benjamin (2018).

Example 2: forming decision weights from stated probabilities

A decision maker is faced with a choice that depends on a stated probability. They are not good at interpreting probabilities, so form an idea of the relative likelihood of the outcomes by considering a lucky version of the future with an unlucky version of the future. How will their interpretation be difference from the true probability?

In this second example, the individual takes only a single piece of information—the probability of an outcome versus an alternative—as the input to the model. The model can in this case explain the hypothesized and empirically estimated shape of the decision weights function in Cumulative Prospect Theory (Tversky and Kahneman, 1992). Experimental evidence on the shape of subjective assessments of probability matches well the shape of the dialectic interpretation of probabilities, where the skepticism parameter governs the skewness of the subjective assessment away from the truth.

Political messaging is just one example of a strategic application of this dialectic framework for how a person processes messages. In Section 7 I suggest some other applications. In particular, I discuss how this decision framework could help us to better understand behavior in the context of effort provision and investment decisions in risky projects, strategic decisions in an oligopoly, patterns of life cycle consumption and saving, and labor supply decisions.

2 Related literature

The primary inspiration for this paper is Froeb et al. (2016). The idea of that paper is to model adversarial decision making as weighing competing interpretations of evidence. A court adjudicates between a plaintiff and defendant but cannot perfectly judge whose interpretation was likelier to have generated the evidence. In the equilibrium of the interpretation selection game, the decision of the court is biased in favor of the unlikelier, more extreme interpretation.

A key difference in this paper is that the receiver here makes mistakes due to their parameterized behavioral characteristics rather than randomly distributed noise. The version of the adversarial framework that I use is less general than in Froeb et al. (2016) since I restrict the

frames to be uniform distributions. This lets me parameterize the model in a parsimonious but meaningful way to highlight the role of what I will call skepticism or credulity.

As well as the nature of the model, this paper also differs in the nature of the applications in that I consider both non-strategic, behavioral processes and strategic outside influencers as possible sources for frames that the receiver is weighing. The political spin application is closest in spirit since it focuses on how the plausibility penalty affects the optimal choices of adversarial political messengers. In the behavioral applications, however, I combine the idea of the adversarial process with the idea of multiple selves (Ambrus and Rozen, 2013) or multiple rationales (Kalai et al., 2002; De Clippel and Eliaz, 2012) in a decision making process.

The model I propose contributes to a literature on decision makers who use narratives to interpret data. Schwartzstein and Sunderam (2019) studies single and competing persuaders who propose a model in the form of a likelihood function to a receiver to explain a commonly observed history of past outcomes. The proposer seeks to induce the receiver to take an action that they prefer, which occurs when the persuader proposes a model that better fits the history than the receiver's default model. Eliaz and Spiegler (2020) models a representative agent who evaluates a supply of narrative-policy pairs: a policy recommendation and a narrative in the form of a directed acyclic graph that suggests a causal model that generated the values of commonly observed variables. The agent evaluates these according to their anticipatory utility, capturing the agent's affinity for hopeful narratives. In Bénabou et al. (2018), image-conscious individuals within a network structure choose whether to communicate narratives to their successors. In a setting in which moral actions in the face of temptation produce positive externalities, the narratives emphasize excuses or responsibilities that change the relative weight on personal costs versus the positive externality.

This paper also features a receiver who interprets frames over observed data, but with a focus on a particular application to political spin and with an empirically plausible behavioral heuristic. I focus on the incentives that lead to extreme partisan polarization in political messaging and on how those incentives can be connected to the characteristics of a boundedly rational receiver. Since the receiver's heuristic has a single behavioral parameter for skepticism or credulity, I explore patterns of behavior across settings according to the receiver's type, without assuming

either that the decision maker is either unboundedly rational or idiosyncratically biased in a particular direction.

There are several prior strands of literature that consider intrapersonal conflict as a motivator for beliefs and decisions. Most notably there is the literature on tension between present and future versions of oneself, for example from Peleg and Yaari (1973) on consumption paths with changing tastes, O'Donoghue and Rabin (1999) on sophisticated and naive anticipation of future preferences, Kahneman (2011) on impulsive and deliberative selves, and so on. In another vein, Ding (2007) considers a class of intraperson games (motivated by the angel and devil on one's shoulders) between with competing selves that care about efficiency and equity. Jamison and Wegener (2010) discusses neuroscientific evidence for the multiple selves approach. Another strand is the 'multiple utilities' approach (an excellent review of this concept in philosophy and economics can be found in White, 2006). An example of this is the idea of a selfish self and ethical self that are mediated by a third, overarching self.

My approach in this paper has affinity with multiple selves models, but it is a little distinct. There is indeed a tension between two frames, but these could be the result of personal reflection or of outside influence, not competing selves. The decision maker is not wrestling with another version of themselves who wants different things, but instead is a single flawed adjudicator who is muddling through the evidence using an imperfect heuristic.

Benjamin (2018) has a thorough overview of the evidence on the type of anomalies in belief formation and inference that I have in mind in this paper. Specific examples of the two anomalies of individual choice I focus on are the inability to accurately interpret and compare stated probabilities (Kahneman and Tversky, 1979) and the inability to accurately update beliefs (Ouwersloot et al., 1998), but I will discuss this in more detail below after deriving the implications of the model in the context of the behavioral applications.

3 Political spin game

In this section I present the basic model that captures some key features of political spin. Later we will consider extensions to the basic model to address other aspects of political messaging and

also consider alternative applications of the basic model to different settings other than political messaging.

A line segment from 0 to 1 represents possible states of the world. There is at least one piece of exogenous evidence that is drawn from a contiguous non-degenerate uniform distribution over a subset of the line segment whose mean is the true state. As an example, Figure 1 shows a situation in which there are two pieces of evidence at 0.6 and 0.8.

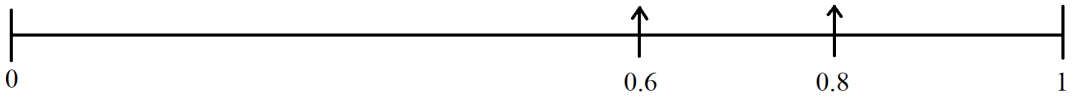

Figure 1: Two pieces of evidence, 0.6 and 0.8

There are two strategic players, political parties left (L) and right (R),[2] and a non-strategic representative citizen. Call the parties the senders and the citizen the receiver. The evidence and the way it was generated are common knowledge to the senders and receiver, but the true state and the endpoints of the specific uniform distribution that generated the evidence are not known to anyone.

The senders simultaneously choose a frame for the observed evidence. Each frame must itself be a contiguous non-degenerate uniform distribution over a subset of the line segment, which represents that sender's claim about the underlying distribution from which the evidence was drawn. Each sender's frame must include the scope of the observed evidence.[3][4]

In the example from Figure 1, the uniform distribution chosen by each sender must cover at least $[0.6, 0.8]$ and can be wider if they choose. A uniform distribution over $[0.55, 0.9]$ is a valid

[2]While there is no obstacle in theory to allowing for more than two players, we may view the two player case as either capturing a true two-party system or a situation in which only these two parties are relevant to the decision-maker.

[3]Senders may not simply ignore evidence when they construct their frame. This means that the evidence represents whatever is uncontroversially agreed upon by all, thereby abstracting from the problem of what counts as a piece of evidence to different people.

[4]Another implication of the specific assumption made about the data-generating process is that the true state cannot be 'too far' from the most extreme piece of evidence. For example, in the example above, since there is a piece of evidence at 0.8, the mean of the uniform distribution that generated the data cannot be less than 0.4. In this sense the commonly agreed-upon evidence quickly narrows down both the range of possible strategies each sender can use and the range of possible true states. The most wildly extreme but conceivable true states, near the endpoints of the $[0, 1]$ line segment, are immediately ruled out by evidence in the interior.

choice of frame, since it includes the available evidence and has no gaps. A uniform distribution over $[0.65, 0.9]$ is not valid, since the evidence 0.6 is not included in the frame and so could not possibly have come from the suggested distribution.

The receiver interprets the frames, and their interpretation is what determines the payoff to each sender. Denote the mean of sender i's frame by μ_i and the width of sender i's frame by w_i. Denote for later reference the upper and lower bounds chosen by sender L as $\bar{L}$ and $\underline{L}$ respectively, and similarly the upper and lower bounds chosen by sender R $\bar{R}$ and $\underline{R}$ respectively.

Since each frame must include the observed evidence, μ_i and w_i depend on both the evidence and the senders' choices of how to frame it. The receiver forms an opinion about the state of the world by taking a weighted average of the mean of the two senders' frames, as in the following equation, in which $s > 0$ is a parameter that enters as an exponent and measures the *skepticism* of the receiver:

$$\hat{\mu} = \frac{w_R^s}{w_L^s + w_R^s}\mu_L + \frac{w_L^s}{w_L^s + w_R^s}\mu_R. \tag{3.1}$$

The payoff to sender R is $\hat{\mu}$ and the payoff to sender L is $1 - \hat{\mu}$. That is, each sender prefers the receiver's assessment to be closer to their preferred endpoint of the line 0 to 1. Sender L prefers the assessment of the receiver to be closer to 0, and sender R prefers it to be closer to 1.

3.1 Interpreting the model and the skepticism parameter

The receiver places more relative weight on the mean of the frame that is more precise and plausible, and places less relative weight on the mean of the frame that is more vague and implausible. This is captured in the model as the width of the uniform distributions chosen by the senders. Since the frames must at least cover the observed evidence, this could alternatively be interpreted as a penalty for speculatively extending the frame beyond the available evidence. If we were to generalize the model to not restrict the senders to uniform distributions, we may interpret 'plausibility' as related to the likelihood of the suggested distribution (as in the approach in Froeb et al., 2016).

As we can see from Equation 3.1, the extent of the relative weighting of the two frames

depends on the receiver's skepticism s. This parameter appears as an exponent on the width terms in the receiver's assessment $\hat{\mu}$: when s is small, the receiver weighs more equally the means of the two frames and pays less attention to their relative plausibility; when s is large, the receiver heavily discounts the frame that is relatively less plausible. This model of the receiver's assessment of the frames is similar to that used in Skaperdas and Vaidya (2012) and Hirshleifer and Osborne (2001).

The skepticism parameter is capturing the extent to which the receiver is either credulous or skeptical of far-fetched suggestions about where observed data came from. The lowest value of the parameter $s = 0$ is equivalent to a receiver who is so credulous that they simply weigh 50-50 the competing explanations they here without considering at all which is more far-fetched. An extremely high skepticism parameter would represent a receiver who favors almost exclusively the narrower frame. Importantly, 'errors', in the sense of yielding an incorrect judgement relative to the true maximum likelihood, are manifest in both the too-credulous and the too-skeptical directions: a receiver can be 'too credulous' or 'too skeptical'. Further, there is no single value of this parameter such that the receiver would make assessments that match the maximum likelihood in all situations, a point I return to in Section 6.1.[5]

In the model evidence is drawn from a uniform distribution centered on the true state, which means that as the evidence accumulates in the model it will eventually span the support of that underlying distribution. The receiver, though, forms a belief using the dialectic heuristic over the senders' frames and not the picture painted by the evidence. This means either that the receiver is sufficiently inattentive that they use the heuristic even when there is a large volume of evidence, or that there are no more than a few pieces of evidence available. It further means that if we imagine a situation in which evidence accumulates and eventually comes to span the true distribution from which it was drawn, then the prediction of the dialectic model will eventually arrive at a long-run belief governed by their heuristic.

In sum what I have in mind is a receiver who is boundedly rational and engages the world by thinking dialectically about competing views. As is well known, people are quite fallible

[5]I should note here that I am not totally comfortable with calling this parameter skepticism. What I have in mind could also be thought of as an inverse measure of credulity, naivety, or gullibility. To avoid being confusing I have picked a name that matches the direction in which the parameter is increasing: bigger s, more skepticism. Nevertheless I want to keep an open mind about what exactly is being measured.

when interpreting evidence or probabilities. The receiver here could well be thought of as a constrained optimizer with limited cognition, time, or attention who embodies a behavioral type—the skepticism parameter—that governs their flawed judgements.

In Sections 6.1 and 6.2, I will make a case for this model of decision making based on some comporting motivation from individual decision theory applications. The applications are a person updating their belief in response to evidence, and a person who makes subjective assessments of stated probabilities. As part of that exercise I will suggest a calibration of the skepticism parameter in light of empirical evidence on how people behave in those applications. If we consider a receiver who naively considers a dialectic between the best and worst case interpretations of evidence about the world, we will see that the skepticism parameter that best calibrates the model for that data also yields appealing predictions when adopted for the receiver in the political spin game.

4 Equilibrium spin and analysis

We can immediately rule out that either sender enlarges their frame beyond the evidence in the direction of their opponent's preferred pole. That is, $\underline{R}$ and $\bar{L}$ are in equilibrium certainly at the lower and upper bounds of the evidence respectively. The reason is that for either sender to go beyond the bound in their less-preferred direction is unambiguously worse for them: it increases the width of their frame and it moves the mean of their frame further from their preferred point.

However, if we think about the incentive for each sender to increase the width of their frame toward their preferred pole, there is a tradeoff. The benefit of expanding the bounds of their frame is that it moves the mean of their frame, which is what will be weighed by the receiver, closer to their preferred point. But the cost of widening their frame is that it decreases the relative weight that the receiver will put on their their explanation. The relative importance of these considerations depends on skepticism—how much the frame widths matter in the receiver's weighting. If the receiver has low skepticism, there is little reason for a sender to improve the plausibility of their explanation.

Result 1. *For a sufficiently credulous receiver, with a skepticism parameter of 1 or lower, the*

unique Nash equilibrium of the political spin game has both senders choose the most extreme possible frame.

Appendix A has the derivation of this result. A credulous receiver—one with a low enough value of the skepticism parameter s—does not penalize implausible frames very heavily. No matter what the underlying evidence, there is always a sufficiently small s that each sender prefers to choose a frame that extends all the way to their preferred endpoint of the spectrum. This s does not have to be so low that the receiver does not consider plausibility at all ($s = 0$). The threshold is certainly no lower than $s = 1$, and is higher than that in cases when the evidence is more favorable to the sender's side. The more extreme the evidence, the more credulous the receiver must be in order for maximum spin by both parties to be an equilibrium, but we can say that $s = 1$ is sufficiently credulous to yield this for any evidence.

Is it plausible that a receiver using a boundedly rational rule of thumb would be excessively credulous? In a setting close in spirit to the context of political messaging—'fake news'—Pennycook and Rand (2019) finds that "susceptibility to fake news is driven more by lazy thinking than it is by partisan bias per se". The authors find that insufficient analytical reasoning rather than motivated, partisan reasoning explains a person's inability to discern fake news from real. The same insufficiently analytical type has a hard time distinguishing fake news regardless of whether it conforms to their political ideology. Similarly, Lantian et al. (2020) finds that people with lower critical thinking ability are more likely to believe in conspiracy theories.[6]

We can further compare this result to the prior literature on adversarial decision making. Like Froeb et al. (2016), Result 1 has both senders choose a frame that exaggerates in their preferred direction, and has the sender for whom the evidence is more unfavorable exaggerate more. The behavioral interpretation of the receiver's response is quite different, being moderated by the skepticism parameter rather than capturing fallibility of judgement. The skepticism parameter also echoes the 'handicap' used in Skaperdas and Vaidya (2012), albeit in a different way—there, the handicap idiosyncratically benefits or hinders each of the two competing persuaders in a way that is unrelated to which one the evidence favors, whereas here the penalty the receiver applies

[6]In this interpretation, the very credulous decision maker is one who puts more weight on conspiracy theories, in the sense that the wider frame is 'more conspiratorial'. It goes further beyond the scope of the evidence, being precisely tailored to fit the evidence into a convoluted and more far-fetched narrative.

depends on how the persuaders' frames correspond to the evidence. Hirshleifer and Osborne (2001) considers the sensitivity of the outcome of the adversarial process to the effort put forth by each party; the 'force exponent' parameter governs how much effort matters relative to the truth of the matter and appears in a broadly similar fashion to the skepticism parameter. The distortion of the receiver's belief here and the distortion of the adversarial outcome there depend in a similar way on the two respective parameters: a more credulous receiver, as with more importance on effort in Hirshleifer and Osborne (2001), induces a greater response by the persuaders.

To visualize what the equilibrium frames look like if skepticism is sufficiently low, Figure 2 continues the example in which there are two observed data points, 0.6 and 0.8. Sender L's frame claims that these have been generated from a uniform distribution over $[0, 0.8]$ and sender R's frame claims that they came from a uniform distribution over $[0.6, 1]$.[7] The two frames are the most extreme frames in the direction of each sender's preferred end of the spectrum that still include the observed data.

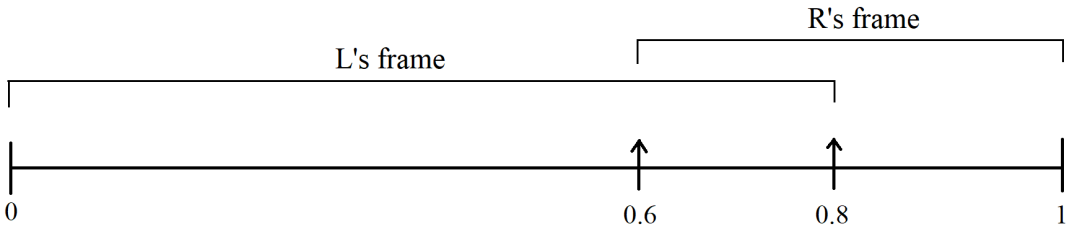

Figure 2: The most extreme frames with two observations

The credulous receiver does not penalize wide, imprecise frames heavily, and so both parties are free to make their frames stretch all the way to their preferred pole. When would sender L prefer a frame other than $[0, 0.8]$ and sender R prefer a frame other than $[0.6, 1]$? Say that R selects the most extreme frame, $[0.6, 1]$. It turns out that the threshold for s below which the best response of L is to select $[0, 0.8]$ lies between $s = 1$ and $s = 2$. For s below that threshold, the best response of L is to be maximally extreme, since the receiver of the message is very

[7] Notice that, following from Footnote 4, this is equivalent to each sender claiming that the true state is as close to their preferred pole as it can possibly be given the data-generating process and the commonly agreed-upon evidence.

credulous. On the other hand, for s above that threshold there is an incentive for a sender to use a narrower frame that does not stretch all the way to the endpoint—such as L using the frame $[0.5, 0.8]$ in the example. The reason is that a more skeptical decision maker discounts the mean of the wider frame more heavily, and so reducing the width of the frame becomes more important relative to moving the mean of the frame in the sender's preferred direction.[8] In short, the maximally extreme frame is a best response as long as the receiver is credulous enough. As I will argue below in Section 6, the degree of credulity that is required for these maximally extreme frames to be an equilibrium is plausible since a similar parameterization makes the receiver's deliberative model consistent with known anomalies in individual choice.

On the other hand, if the receiver was sufficiently skeptical, then there is an incentive for a sender to use a narrower frame that does not stretch all the way to the endpoint—such as L using the frame $[0.5, 0.8]$ in the example. The reason is that a more skeptical decision maker discounts the mean of the wider frame more heavily, and so reducing the width of the frame becomes more important relative to moving the mean of the frame in the sender's preferred direction.[9]

The skepticism parameter here works to discount the relatively implausible frame, so it works a little counterintuitively. An extremely skeptical person will form a posterior belief that is *more* skewed toward an extreme endpoint than Bayes' Rule implies. They are excessively convinced by a frame with tight bounds and excessively dismissive of the frame with looser bounds. One may imagine a skeptical person having a belief closer to the middle of the line, since they are difficult to convince of anything. In this model, the parameter I have called skepticism is quite the opposite, because it relates to skepticism about *explanations*, not skepticism about *evidence*.

4.1 Susceptibility to spin

Next let us consider what would happen in the model if there were different receivers with different skepticism parameters. First, if the senders could tailor their frames to each audience,

[8] Although it is not present in this model, another reason why a sender may prefer a narrower frame may be that they care not just about the receiver's point belief but also their confidence in the state.

[9] Although it is not present in this model, another reason why a sender may prefer a narrower frame may be that they care not just about the receiver's point belief but also their confidence in the state.

then they may prefer not to use the most extreme possible frames when courting a sufficiently skeptical audience. For example, a political candidate may be much more strident and extreme when talking to an audience at a rally, but more equivocal and sober when talking to wealthy donors behind closed doors. This could be rationalized here if the closed-door audience is more skeptical, since then they would discount an extreme frame more than the rally attendees.

But what about a situation in which the parties have to present a single frame to a range of individuals with different degrees of skepticism? Which receivers will be susceptible or seemingly immune to spin? If the senders are selecting the most extreme frames, the aggregate interpretation of them will depend on the distribution of the skepticism in the population. But ex post, it will seem that the spin 'worked' on some people but not others—while those with low s are successfully pulled in the direction of the far-fetched spin, those with high s are pushed toward the more sober spin. This is consistent with evidence that there exist types of people who are susceptible to conspiracy theories. For example, Miller et al. (2016) identifies high-knowledge but low-trust individuals as being particularly susceptible, in line with the receiver here who is fully aware of the evidence but whose credulity means that they are not at all convinced by narrow, parsimonious frames.

Figure 3 again uses the example of Figure 2. It shows how the posterior belief depends on the value of s in the example with evidence 0.6 and 0.8. The maximum likelihood for the bounds of a uniform distribution that may have generated this evidence is $[0.6, 0.8]$, but in general the receiver's posterior belief is not the mean 0.7 of that distribution.

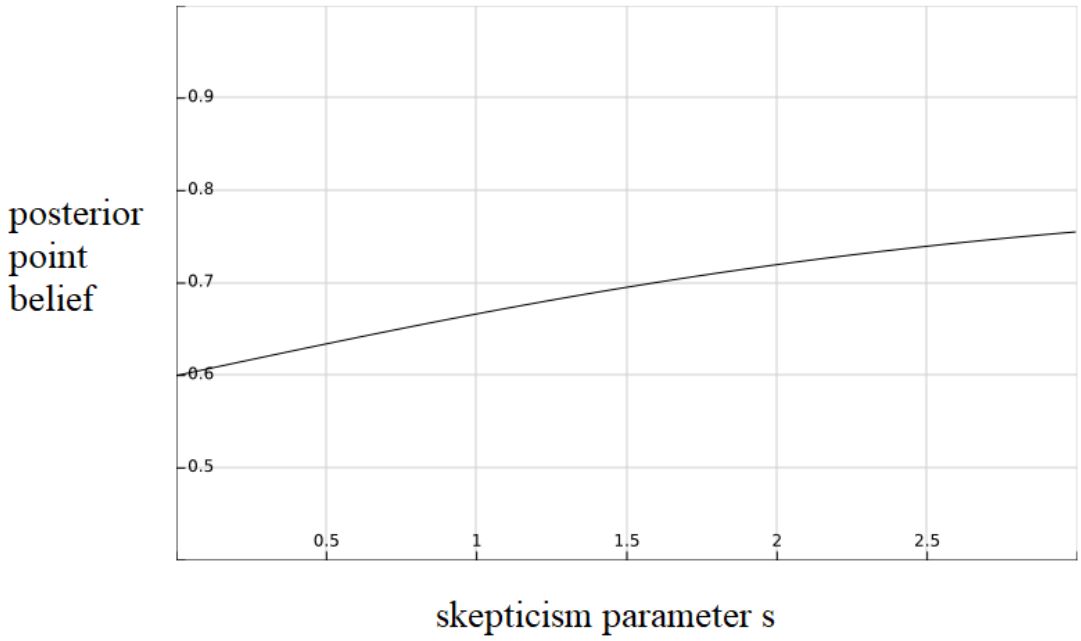

Figure 3: Posterior point belief as a function of skepticism s

As we can see, a very skeptical receiver will form a belief that is higher than 0.7, favoring to excess the narrower frame, and a very credulous receiver will form a belief that is lower than 0.7, favoring to excess the wider frame. If $s = 0$ then the receiver is so credulous that they impose no penalty for wider frames, and so the posterior belief simply reflects the average of the means of the two frames, 0.6. If receivers of various skepticism parameters were exposed to the same maximally extreme frames, we would see receivers with low skepticism parameters hold an opinion skewed down from the maximum likelihood estimate of 0.7, and receivers with high skepticism parameters hold an opinion skewed up.

Of course, the sender's optimal strategy may change depending on the distribution of the skepticism parameter in the target population. Consider an illustrative example of how the distribution of s matters in the case with evidence at 0.6 and 0.8. Say that a proportion c of receivers are very credulous, with $s = 0$, and the remaining proportion $1 - c$ of receivers are very skeptical, with $s = 3$. Fix R's strategy at $[0.6, 1]$ and consider the problem for L. Compared to choosing $[0, 0.8]$ (that is, $\underline{L} = 0$) if sender L makes their frame narrower they face a tradeoff between improving their standing with the $s = 3$ receivers and the harming their standing with

the $s = 0$ receivers. The narrower frame only hurts them with the credulous receiver, who does not consider the width of the frames at all, but it (at first) helps them with the skeptical receiver, who penalizes wide frames quite heavily.

The optimal strategy for L depends on their goal. In the baseline model above, L's payoff was simply the distance between the receiver's belief and their preferred pole. With a distribution of receivers there are different ways to think of the sender's incentives: they may care about all of the receiver types versus just some, about shifting public opinion versus winning an election, or about hitting some threshold opinion versus driving opinion as far as possible in their direction. The optimal choice for L will be sensitive to what blend of motivations they have.

To give one simple example, say that L seeks to make the average opinion in the population as low as possible. L's payoff then depends on c. Figure 4 shows how the average opinion depends on $\underline{L}$ for three values of c.

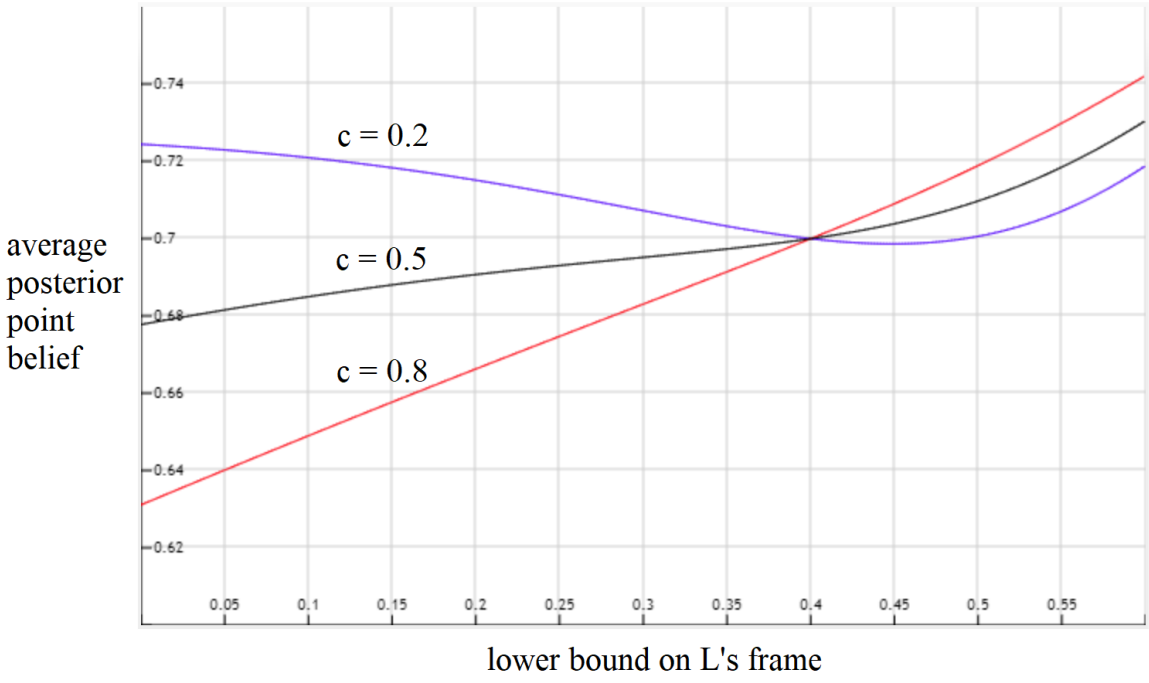

Figure 4: Average posterior point belief as a function of $\underline{L}$

When the proportion of credulous receivers c is high, the widest possible frame is optimal—that is, $\underline{L} = 0$. As c decreases, the value to L of narrowing the frame increases, so that eventually the optimal frame is narrower. In sum, how much to spin the evidence therefore depends on the

distribution of s, which determines how susceptible the population is to extreme spin.

4.2 Partisan polarization and sorting

Consider the example of climate change. The weight of scientific evidence supports the position that climate change is a serious and pressing concern. Say that, for reasons that would be exogenous to the model, one political party would prefer it if people formed an opinion that climate change was not a serious or pressing concern, while the other would prefer it if people formed just the opposite opinion.

Result 1 showed that in world with low skepticism on average, the two parties optimally and effectively use maximally extreme frames, even in the face of extremely skewed evidence. Since the evidence is gathered predominantly at one end of the spectrum, the party that denies the problem of climate change will have adopted a frame that is much wider—less plausible—than their opponent. To square such evidence with the platform requires a tortured explanation. How will receivers respond?

Maximally extreme asymmetric frames together with a varied distribution of the skepticism parameter implies polarization in the response to spin. If we were to take stock of what receivers held ex post opinions—after the spin game—closer to one end of the political spectrum or the other, we would observe systematic differences in skepticism among adherents of different party messages. A correlation between the behavioral parameter skepticism and political inclination would emerge.

Result 2. *As long as the midpoint of the evidence is not precisely at the midpoint of the spectrum, receivers' ex post positions will be distributed according to their skepticism parameter. That is, one party will enjoy relatively more support from receivers who are very credulous and the other party will enjoy relatively more support from receivers who are very skeptical.*

Since in the model the receiver has no partisanship or ideology, the result does not depend on which party the 'conspiracies' favor.[10] Pennycook et al. (2015) and Hart and Graether (2018) find

[10]In a world in which there is idiosyncratic partisanship, this result would imply that credulous receivers will move towards the wider frame and skeptical receivers will move towards the narrower frame, *relative* to where their inherent partisanship started them off.

that personality factors including "bullshit receptivity" (similar in spirit to credulity here in the model) are associated with belief in generic, non-partisan conspiratorial statements. This suggests that there may indeed exist types of individual whose credulity renders them suggestible to conspiracy theories in general. Similarly, Douglas et al. (2016) finds that "hypersensitive agency detection" predicts belief in conspiracy theories, even after accounting for political ideology.

Ignoring for the moment ex ante partisanship, the relationship that emerges here between a receiver and the evidence, or the media that reports it, is consistent with some prior evidence about how partisans process information. To an excessively credulous receiver, the evidence and the media seems unaccountably biased towards the explanation and the preferred pole of their less preferred party. The partisan sorting here is thus consistent, for example, with evidence that the media's supposed liberal bias is the result of conservative elites' claims of liberal bias (Watts et al., 1999; Lee, 2005). Similarly, Fessler et al. (2017) provides suggestive evidence that a person's political orientation is associated with their credulity with respect to negative claims about the world. A hypothetical disinterested media—one that reported the midpoint of the evidence, for example—does not reach the same conclusions as the dialectic model, especially for very low skepticism. This may well color the credulous receiver's feelings toward the messengers.

To the credulous naively dialectic receiver, it is the media itself, or indeed the evidence itself, that seems biased away from the supposedly balanced conclusion drawn from the framing game. Therefore and by contrast, 'both sides' reporting will appear fairer to this receiver than reporting that ignores the frames and reports the evidence. This asymmetric reaction to the evidence is consistent with the finding that, for example, Republicans appear to discount new scientific evidence on climate change more than Democrats (Campbell and Kay, 2014).

The interaction between partisanship and the receiver's relationship to the evidence here is in line with Stephen Colbert's famous quip that "[i]t is a well known fact that reality has a liberal bias" (Colbert, 2006). In his telling, of the two major political parties in the United States, one is routinely confronted by expert advice and opinion that is quite far away from its preferred platform, for example on climate change or the effect of tax cuts, in keeping with the pattern of partisan sorting we see in the model.

4.3 The arrival of new evidence

The model as presented above is static and so does not explicitly consider the dynamics of what happens when new evidence arrives. However, if we were to extend the model in that direction, another facet of the receiver's process in the model is that reinforcing evidence does not in this model change the opinion of any receiver:

Result 3. *Reinforcing evidence, in the sense that it falls within the scope of the existing evidence, has no effect on the receiver's opinion.*

An increasing volume of evidence doesn't change $\hat{\mu}$.[11] Returning to the example of climate change, a receiver with low skepticism will believe the problem to be less significant than the evidence suggests, *and* this position will not change at all when they encounter reinforcing information. The absorption of new, reinforcing arguments without any effect on an individual's ex post position is in the style of the 'disconfirmation bias' of Taber and Lodge (2006).

A different possibility that is not present in the model is that receivers may themselves have a bias toward one end of the spectrum or another. This may manifest as motivated information gathering, for example the second effect of 'confirmation bias' in Taber and Lodge (2006), so that new evidence is selected by the receiver. But note that we do not necessarily need this idiosyncratic bias to exist for a receiver's ex post opinions to appear systematically skewed towards one party or another. If the evidence congregates more toward one side of the spectrum, then the more far-fetched explanations will be consistently observed from the partisan framer of the opposite side. Even without idiosyncratic preferences, the ex post opinion of a receiver with low skepticism will be skewed to the far-fetched. This means that it is possible that an individual's degree of skepticism can influence their (seemingly partisan) dialectic response to frames and in turn their political position via the mechanism in this model, as well as the more familiar opposite direction in which the individual's political position influences their response to information.

[11] In the model, this result depends on fact that the receiver is forming (and the senders care about) their best guess about the state rather than a full belief distribution over the state. If, for example, the receiver's level of confidence in their belief about the state matters, then as more evidence accumulates their confidence may change, which in turn may inform the senders' strategies.

Finally, in the model there would be no negative consequence for a sender if they propose a frame that is later proved false by the arrival of new evidence outside the scope of the frame. This may reflect, for example, a world in which both the senders and the receiver are in the dark about the true state, and so the receiver does not attribute the newly contradictory evidence to mean that the sender wilfully attempted to mislead with the original frame. An alternative could be that the receiver would lose faith in a sender in this situation and so would discount their subsequent frames more in the future.

5 Extensions to the political spin game

5.1 Incentives to generate or suppress new evidence

Next, we can consider the incentives of the parties to release or suppress evidence, in light of its implications for the political spin game. This relates the model to prior work on motivated release of evidence, for example Daughety and Reinganum (2000), Gentzkow and Kamenica (2016), and Bull and Watson (2018). Say that a sender is considering releasing (or fabricating) evidence that is unfavorable to their opponent, in the sense that it would force the opponent to expand their frame in an unfavorable direction.[12] To address this question, we will consider the effect on the receiver's assessment $\hat{\mu}$ of raising $\bar{L}$ (since the problem is symmetric for both parties, this also will tell us about what happens to $\hat{\mu}$ when lowering $\underline{H}$).

Result 4. *Say that the receiver is sufficiently credulous that the senders use maximally extreme frames and that a sender could release evidence that forces their opponent to expand their frame. The marginal value of releasing the evidence is higher when the opposing sender's existing frame is narrower in scope.*

The derivation of this result can be found in Appendix C. It is unambiguously good to force the other sender to expand their frame because it reduces the weight the receiver places on their frame while moving the midpoint of their frame in their less preferred direction. However, the amount by which such actions shift the receiver's assessment $\hat{\mu}$ in the sender's favor depends on

[12] Recall that the model does not allow for the possibility of simply ignoring evidence by leaving it outside of the frame, a point discussed in more detail in Section 5.3.

the evidence. These 'dirty tricks' are most effective at shifting $\hat{\mu}$ when the scope of the current evidence is narrowest in favor of the opposing party. In sum, forcing the other party to expand their frame is always beneficial, but more beneficial when the current evidence is more precise and more favorable to the opponent.

This is a somewhat counterintuitive effect of the receiver's credulity: the sender is expanding their opponent's frame, so why would this be more valuable when the receiver is credulous of wide frames? The reason is that the receiver's credulity means that their weight on the newly expanded frame is still high. This is ideal for the 'dirty tricks' sender, who has shifted the mean of their opponent's frame toward their preferred point and so is happy that the receiver still weighs it highly.

As an application of this result, consider a situation in which the two parties compete with spin on several issues, not just one issue as in the model. For example, political competitors try to persuade voters on taxes, health care, foreign policy, and so on, while potentially responding also to new and unexpected issues as they arise. If the party seeks to use resources to discover new evidence unfavorable to their opponent, on what issues would this do the most to change the mind of the receiver? According to Result 4, the best bang for the buck is to allocate these resources towards issues for which the current scope of evidence is relatively tight in favor of the other party, and away from issues for which the current scope of evidence is wider.

It follows that the second piece of evidence on fresh issues is the most valuable to a party for whom the initial evidence is worse, since this after all is the evidence that arrives when the frame is narrowest. There are therefore diminishing returns to efforts by either party to seek or suppress evidence on a given issue. If we think about a situation in which parties compete over many issues over time, then this is consistent with short and ever-changing news cycles: the competing senders (i) race to 'control the narrative' on an issue, (ii) seek to obfuscate by sowing doubt and uncertainty, and (iii) quickly change the subject to the next issue of the moment where the returns are still high. Continuing to devote resources to a heavily mined issue, pedantically trying to reinforce one's point, would be suboptimal: once the other side has sown doubt, the prudent course is to attack the fresh territory of a new issue.[13] In that sense

[13]If fabricating evidence is in play, this approach is quite similar to patterns of disinformation and trolling online (Marwick and Lewis, 2017).

this result is complementary to Result 3, which states the ineffectiveness of reinforcing evidence. Combined, they highlight aspects of the model that match well a world of short, superficial news cycles and a discourse that prioritizes airing 'both sides' over the weight of evidence.

5.2 Partisanship and asymmetric response to spin

In the model the receiver is non-partisan—they treat the two senders identically. As we have just seen, the receiver's apparent ex post partisan leaning arises in the model from the conjunction of their degree of skepticism plus a systematic skew of evidence toward one side of the spectrum. This is consistent with Van Prooijen et al. (2015) in that political extremism and openness to far-fetched conspiratorial thinking are linked, on both the left and right sides of the political spectrum. Of course, partisanship is in reality a much more complex force.

Another channel that is well-founded in our best understanding of partisanship is that the receiver's partisan leaning influences their response to each party's frames in an asymmetric way. For example, some pre-existing partisan preference may make the receiver interpret new evidence in favor of their preferred party. This is consistent with the evidence on confirmation bias from Taber and Lodge (2006), as mentioned earlier in Section 3, and evidence from, for example, Nisbet et al. (2015) that both conservatives and liberals react negatively to dissonant scientific communication. While this is beyond the scope of this paper, it could be incorporated in the model in a couple of ways. One would be as an idiosyncratic weight on the receiver's evaluation of the frame of their preferred party. A more extreme approach to capture dogmatic insistence on extreme partisan beliefs could be to add to the model an idiosyncratic weight on the receiver's preferred endpoint, over and above the receiver's relationship to the frames. This would be equivalent to allowing the receiver to believe in a different data-generating process than the model's uniform distribution, and in practice allow the receiver to hold beliefs that are literally impossible given the evidence.

In both cases, whether this would change the incentives of the parties to use extreme spin would depend in part on whether there is systematic variation in skepticism by party orientation (as discussed in Section 4.2) and on the strength of partisanship versus the frames. For example, in a world with biased partisans, if party A has systematically more skeptical partisans than

party B, then B has a greater incentive to 'court' A's partisans with a less extreme, narrower frame. Using the narrower frame has a lower cost to party B in terms of the beliefs held by their own partisans, who are credulous but very attached to party B, than in the baseline model.

Another sense in which we could introduce asymmetry to the model is in the incentives of the parties themselves. The parties in the spin game are interested purely in influencing the receiver of the message in their preferred direction. For that reason, shamelessness is an asset for a framer here and integrity to the truth a liability. A framer who was idiosyncratically incapable of stretching the truth—perhaps one who used a frame that is simply the scope of the evidence—would certainly be at a disadvantage relative to a shameless opponent. This is already directly implied by Result 1 on the incentive to use extreme frames, and so we can easily see how integrity will backfire on its bearer if it was included in the model.

5.3 Evidence unexplained by the frames

In the model, the senders are restricted to select frames that must include all of the evidence. One possible way to justify this is that the receiver will discount to zero any frame that does not include the evidence. The set of evidence, in this conception, is that which is undisputed and must be explained. The question of what exactly counts as undisputed is one that is assumed settled prior to the model.

This is a notion that may also naturally interact with the partisanship of the receiver discussed in the previous section. For example, Redlawsk (2002) provides evidence that a person's disposition toward a political candidate can lead them to avoid new information about them. In such a situation the party would be freed to leave unfavorable evidence out of their frame without fear of penalty, at least from their supporters.

It would be possible to have a less severe treatment of unexplained evidence by the receiver, perhaps an additional discounting of a frame that leaves evidence unexplained. The penalty for unexplained evidence may plausibly be related to the skepticism parameter within the model, or to the idiosyncratic partisanship mentioned in the previous section. If this had an effect it would be to weakly increase the extremeness of the parties' frames, since narrowing a frame by leaving some evidence outside it would be subject to the same incentives as narrowing a frame

in general. All else equal, the sender would always prefer to narrow the frame toward their own preferred endpoint than away from it.

It is notable, however, that for big, politically contentious questions conspiracy theories are rife. Conspiracies are the quintessential example of outlandish explanations of inconvenient evidence. Evidence in Van Prooijen et al. (2015) suggests that stronger political ideology—perhaps captured by a more extreme assessment by the receiver in the dialectic model—is associated with greater belief in conspiracy theories. In the classic example of climate change, Douglas and Sutton (2015) contends that "[p]erhaps to an unparalleled extent, people on both sides of the issue champion climate change conspiracy theories" and points out that there are numerous prominent peddlers of climate conspiracies in the worlds of politics and punditry.

It may be surprising that such people feel the need to explain rather than ignore evidence, but this would be consistent with a tough penalty imposed by receivers for leaving information unaccounted for, even if the alternative is a far-fetched story. Another less conspiratorial example from a different area is found in Bisgaard (2015). Survey data from Britain from 2004-2010 suggests that partisans by and large agreed about the basic facts of worsening economic conditions, but nevertheless disagreed about whether the governing party was responsible. This is another instance of accepting but selectively explaining away inconvenient evidence.

Nevertheless, incorporating the possibility of unexplained evidence would allow the model to blend with other approaches to belief formation based on motivated reasoning. For example, Bénabou and Tirole (2016) analyze motivated reasoning using a model of beliefs as economic goods. This generates non-Bayesian behaviors that are complementary to those generated by the dialectic model in this paper.

5.4 Multidimensional evidence

Finally, another possible extension of the model would be to admit multidimensional evidence. The model has only a one-dimensional spectrum for the parties' frames and the receiver's assessments—better for one party towards one end, and better for the other party towards the other end. This could be extended to a multidimensional spectrum. For example, it is typical to see ideology analyzed on the two (supposedly separable) dimensions of economic and

social considerations. Contemporary analysis of the public's policy attitudes, for example in Treier and Hillygus (2009) and Klar (2014), seeks to locate ideology on multidimensional maps of this kind.

Evidence, frames, and assessments would then be located in a unit square. The receiver's penalty would be for the area of the frame rather than its width. Putting this together with the kind of idiosyncratic biases discussed in Section 5.2, matching the model to different distributions of ideology and different party configurations (beyond a simple two party left-right configuration) would become possible. Although the basic incentives of the model would quickly translate, the precise effect on the receiver's belief and the types of frames that the parties would use would be more complicated to predict. The intensity of the receiver's preference for each dimension, their preference for each party, and the possibly asymmetric credulity across the two dimensions of ideology would all influence the exact nature of the problem faced by each party.

6 Psychological examples of the dialectic model

In this section I will consider in more detail the model of how the receiver interpreted the competing frames in the political spin game. This model can be applied to individual decision-making situations, where the frames that the decision-maker deliberates over come from introspection rather than an outside influencer. To provide supporting evidence for this model of the receiver, I want to show that when applied in this way it predicts some well-known behavioral anomalies in individual choice problems.

The dialectic model can be viewed in this context as capturing the idea that there are optimistic and pessimistic ways to see the world. The same situation can look quite different depending on your perspective. That piece of bad news you just got was just bad luck: you're doing great, and you'll get them next time for sure. Or was it? That piece of bad news means that *you're* bad and you should quit now. Both stories are consistent with the evidence. So which is it? Isn't the truth somewhere in between?

Dialectics like this are a common thing. We have already considered the possibility that they are manipulated by others: the political spin game presents competing explanations of

the world; the 'good cop, bad cop' routine presents the accused with a reassuring story and a scary one. But sometimes these dialectics might come from within: the glass could be half-full, or it could be half-empty; you dream about winning the lottery, but maybe it'll be a waste of two bucks. What these all have in common is that in a complicated world, thinking about the subtleties of the evidence or the mathematics of probability is hard, while thinking about competing narratives is easy.

So dialectics of the type we consider here may take a few different forms. They could be internal moral struggles, like the angel and the devil on your shoulders.[14] They could be representative of different psychological states, such as the confident and anxious versions of oneself. They might be be induced by strategic actors, for example by a good cop/bad cop routine or the propaganda of political spin. But what they have in common is that a decision maker is setting up or is presented with a dilemma of interpretation. The tendency toward this kind of dialectic is a familiar way to deal with a complicated world, and so the idea here is to suggest a common behavioral framework for the notion of the dialectic and derive its implications in different decision making problems.

6.1 Forming a belief based on evidence

First consider the case of belief formation in the face of evidence. A decision maker would like to make their best guess about the state of the world. Just as in the political spin game, the domain of possible data and states is on the unit interval $[0, 1]$. The true state of the world is a point in the unit interval but it is observed only with noise. The decision maker observes events that are drawn from an underlying uniform distribution over a subset of the $[0, 1]$ interval whose mean is the true state.[15]

The decision maker entertains two competing conjectures about what the underlying distribution that generated the evidence might be; these are the two frames on the data that the decision maker will adjudicate. The frames are two continuous uniform distributions: one stretches from

[14] See Black (2014) for an excellent, varied history of the idea of this concept, from Plato's *Phaedrus* to *The Simpsons*.

[15] I do not consider the case in which the decision maker themselves is responsible for seeking out evidence. Also, the frames in this application must include all the evidence that the decision maker is aware of, but it is possible that there could be evidence that exists but of which the decision maker is unaware.

0 to the upper bound on the evidence; the other stretches from the lower bound on the evidence to 1. These frames are the most extreme underlying distributions that could have contained the data. One way to interpret the frames is as the most optimistic and pessimistic readings of the observed data: the best and worst case scenarios given what we can see.

The decision maker forms a posterior point belief about the true state of the world, in the same manner as in the political spin game (Equation 3.1). We may consider this belief as something the decision maker will act on or that will influence their state of mind. They form the point belief by taking a weighted average of the means of the two extreme frames. The decision maker's prior belief about the true state is assumed to be flat over $[0, 1]$.[16]

The weights are the relative widths of the other frame's uniform distribution. This skepticism means that the more implausible the frame—in the sense that the observed data is more skewed to one side within the distribution or, equivalently, that is has a higher variance and lower likelihood of having generated the data—the less weight that frame will get in the decision maker's evaluation. How the decision maker weighs the two frames depends on the parameter s.

Some examples of applications that can be captured by this model are the three following:

Example 1: noisy signals about the quality of a project

The decision maker has an idea or invention but they don't know its quality. Endpoint 1 represents the best possible quality that the idea could have and endpoint 0 represents total failure. Evidence is generated when critics, peers, or colleagues give their assessment of the idea.

Example 2: deviation from a target

Say that the decision maker has a target they are aiming for, for example an ideal target for their blood sugar level. The target level can be represented as endpoint 1, and specific blood sugar readings can be represented as points of evidence. A reading that is further away from the target, above or below, is a piece of evidence closer to 0 and a reading that is closer to the target is a piece of evidence closer to 1.

Example 3: failure rates

[16]This excludes the possibility that the decision maker may evaluate the plausibility of a frame according to its relationship to their prior belief, for example by placing less weight on a narrow frame that excludes a state that they had believed was very probable. See Schwartzstein and Sunderam (2019), for example, for analysis of a model in which the receiver uses a procedure of that kind.

A company is testing new software. The unit interval represents the success rate of the tool: endpoint 1 is 100% accuracy and endpoint 0 is 0% accuracy. Each test of the software on a test file has a success rate which is a point of evidence.

The decision maker in this model is good at imagining extreme, deterministic, or certain scenarios but not good at applying Bayes's Rule or thinking probabilistically. They therefore weigh up the two scenarios that are easiest for them to imagine, attempting to balance those competing explanations using an ad hoc heuristic.

In Figure 3 above we considered a case with two pieces of evidence, at 0.6 and 0.8. As another example, say that there had been two pieces of evidence at 0.2 and 0.3. Figure 5 shows the posterior belief in that case, as derived from Equation 3.1.

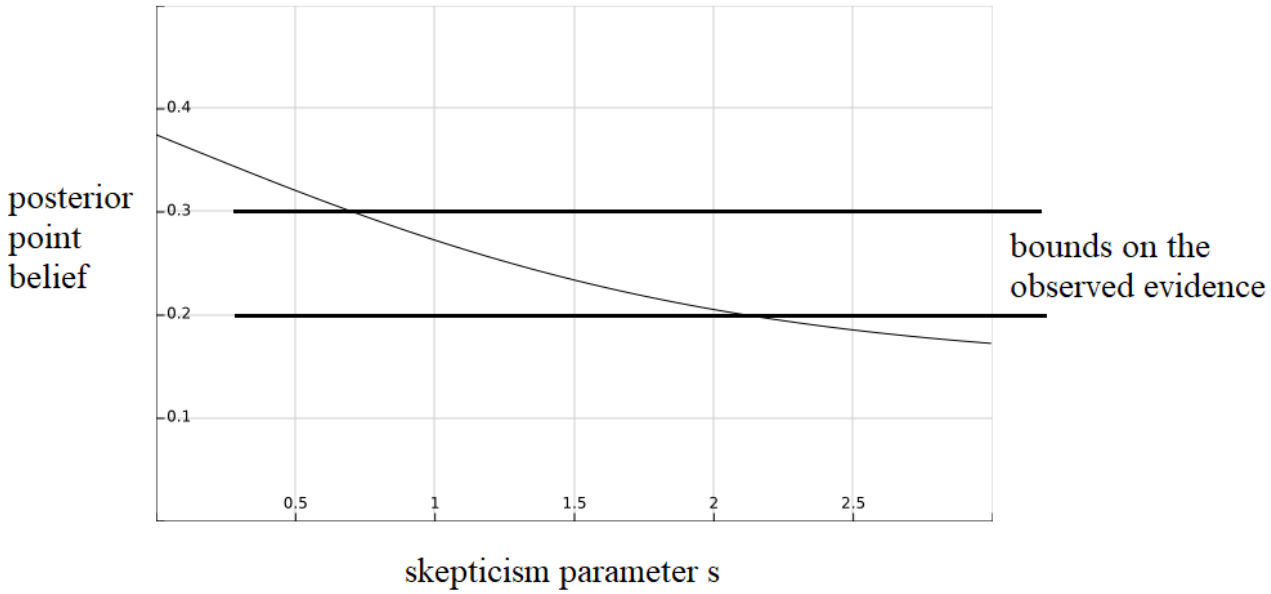

Figure 5: Posterior point belief as a function of skepticism s

From this case we see that for extreme values of skepticism, the decision maker's posterior belief after adjudicating the frames can be outside the scope of the available evidence. This is true in both directions around the mean of the evidence, consistent with experimental evidence that some people underreact and some overreact to new information relative to Bayes (Epstein et al., 2010). Since the frames need not center around a point within the observed evidence, the

dialectic that produces a weighted average of them need not do so either.[17]

If decision makers forms beliefs according to the dialectic model, some implications include:

1. *Irrelevance of new evidence*

   The arrival of new evidence is irrelevant to the decision maker's posterior belief if it is within the scope of the original evidence. This is because new evidence of that type does not change the width and mean of either frame. People's beliefs do not move toward an increasing volume of similar evidence, but do move toward new, dissimilar evidence. Although this follows from the fact that the frames in the model are restricted to be uniform distributions, it is consistent with evidence on how people process new information. For example, Guess and Coppock (2018) that there is no 'backlash' effect to counter-attitudinal messages such as fact-checking. In the model, fact-checking could lead to either a null response or a move toward the evidence, depending on whether the message is similar to information that the person has been exposed to before.

2. *Everyone makes mistakes*

   The 'right' skepticism for one problem, in the sense that it generates a posterior belief that agrees with the maximally likely state, is wrong for another. If s is a fixed behavioral parameter carried by an individual, this means that no-one will get the right posterior belief every time.

3. *Aggregate belief is systematically far-fetched*

   Say that the distribution of s in the population was symmetric around the value of s that would yield the maximum likelihood estimate. In that case beliefs are asymmetrically skewed toward the more outlandish explanation. Mistakes are not symmetrically distributed, as we can see from Figures 3 and 5: the 'misses' are bigger in the direction of less skepticism.

[17] If we consider a case in which evidence continues to accumulate over time, then the evidence will eventually come to span the true distribution from which it was drawn. In that case, the decision maker's belief would eventually come to be fixed at the dialectic model's interpretation of the true uniform distribution that generated the data. For example, if the evidence was truly drawn from a distribution over $[0.2, 0.3]$, then the decision maker's belief would eventually conform to that in Figure 5.

4. *Conservatism*

   For low values of the skepticism s, the decision maker's posterior assessment after deliberation over the two frames is systematically closer to 0.5 than the maximum likelihood estimate of the mean considering the observed data. This is consistent with conservatism in belief formation that is excessively equivocal and underreacts to new evidence. The decision maker's conservatism will be more pronounced in situations with more extreme evidence, consistent with experimental evidence from psychology (Peterson and Miller, 1965; Phillips and Edwards, 1966; Slovic and Lichtenstein, 1971).

The last of these implications and the arguments below in Section 6.2 and Section 3 lead me to conjecture that an appropriate value for the skepticism parameter in calibrating the model is less than 1.[18] Conservatism in belief formation, overweighing of small probabilities, and maximally extreme political messaging are all consistent with the implications of this model for $s < 1$. On the other hand, for high values of the skepticism parameter, the decision maker's assessment is further from 0.5. A very skeptical person, after weighing the two frames, may therefore end up holding a more extreme belief than is justified by the data.[19]

To expand on the conservatism implication a little, in a comprehensive overview of the evidence on belief updating, Benjamin (2018) identifies several stylized facts from the literature. These include persistent underinference in the face of new information that takes hold after only a single signal and is more severe when there is more evidence. This is precisely what my model generates when the decision maker is insufficiently skeptical (or excessively credulous, depending on your preferred interpretation). This is because an accumulation of evidence certainly changes the Bayesian posterior, but does not do anything here unless it changes the shape of the two frames. The low-skepticism person in this model is quick to jump to systematically inaccurate conclusions and hard to budge from them. Two known possible channels for this type of evidence on belief updating are extreme belief aversion—a distortion away from less extreme posterior

[18] A credulous decision maker has a similar idea to a decision maker that undervalues high-quality information as in Ambuehl and Li (2018).

[19] Afrouzi et al. (2020) presents experimental evidence of subjects overreacting to recent observations of a data generating process, suggesting a role for costly, imperfect memory. Although the context of imperfect recall is beyond the scope of the dialectic model, the presence of both overreaction and underreaction in the experimental literature on belief updating suggests a possible avenue to explore the role of credulity and inference from narratives in this context.

beliefs—or underinference from data. The naively dialectic process with low skepticism is another possible channel.

The notion of conservatism can also be an alternative explanation for the well-studied 'disposition effect' in finance, that investors have a "tendency to sell assets that have gained value ('winners') and keep assets that have lost value ('losers')" (Weber and Camerer, 1998). Also in a finance application, the representative heuristic and conservatism identified in Barberis et al. (1998) as possible psychological explanations for underreaction to news and overreaction to a *series* of news are also simultaneously reconcilable here.

6.2 Interpreting probabilities

Consider a second behavioral application of the model. A person is interpreting a stated probability and will form a decision weight (following the terminology of Kahneman and Tversky, 1979) from it.

The evidence, in the language of the previous application, is in this case a single point, simply the stated number, between 0 and 1. The two frames are the decision maker's dialectic interpretation of the stated probability. For example, say that the probability is of some favorable outcome happening. Then the frames may be interpreted as representing the decision maker's optimistic and pessimistic views of that stated probability. The decision maker's decision weight is then given by the dialectic model's Equation 3.1. This application is equivalent to the belief formation application with a single piece of evidence.

The key feature of the model in this application is that the skepticism parameter influences the decision maker's decision weight in a clear and intuitive way. The effect of a low skepticism parameter in this application will be to boost the weight on the relatively unlikely outcome in the mind of the individual. A low probability event is less likely to occur; a person's credulity determines how readily they can envision this event occurring relative to more likely events. If the value of the skepticism parameter is very low, the decision maker is credulous: small probabilities will be disproportionately overweighted by the decision maker. Conversely, if the value of the skepticism parameter is very high, the decision maker is skeptical: small probabilities will be disproportionately underweighted by the decision maker.

To see how the model works in this case let us consider what happens for stated probabilities of precisely zero and small but positive. For concreteness, continue to say that the probability in question is of some favorable outcome happening, so that we can continue to use the language of the optimistic and pessimistic frames.

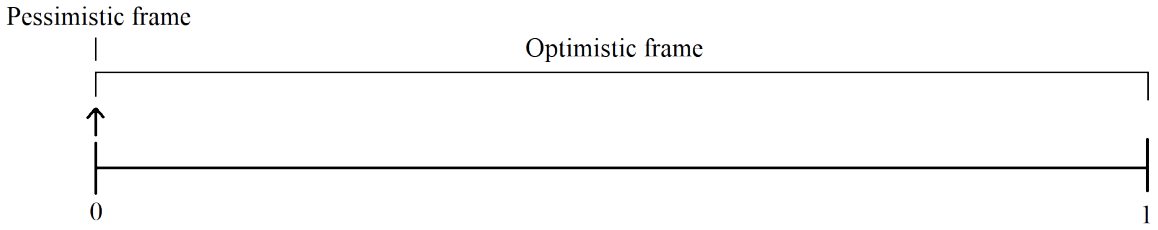

Figure 6: For stated probability 0 or 1, one frame is a point

Figure 6 shows the interpretation of a stated probability of zero. The pessimistic frame here is a zero-width point. In the computation of the perceived probability, the optimistic frame therefore gets zero weight, and the interpreted probability is precisely zero.

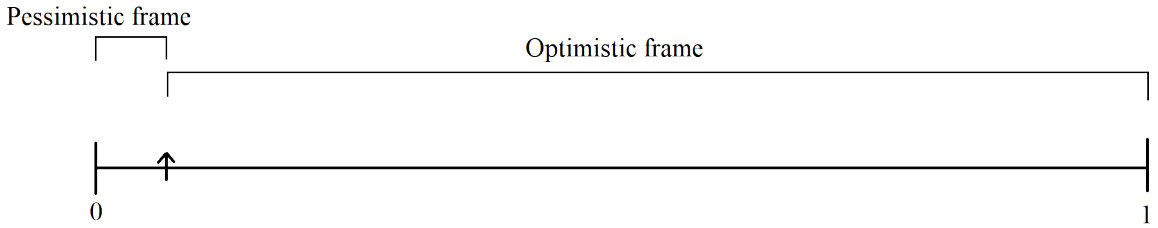

Figure 7: For stated probability close to an endpoint, both frames have width

If instead the stated probability is small and close to zero, the pessimistic frame has width—it may be slim, but there is a chance that things will not fall in the decision maker's favor. The computation of the perceived probability therefore puts positive weight on both frames and so the interpreted probability is positive. From these two examples we get the overweighing of small probabilities, a key behavioral regularity in the decision weights literature.

Across the full range of stated probabilities, the decision maker's decision weights depend on the skepticism parameter s. Figure 8 illustrates this with the shape of the interpreted probability

function for a value of skepticism $s = 0.25$.

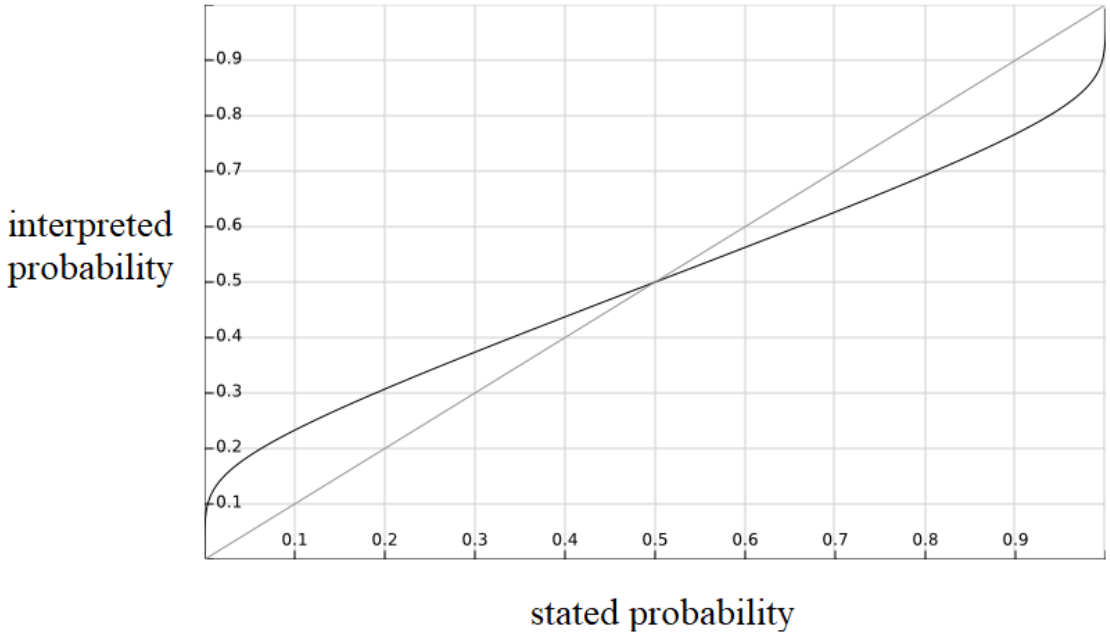

Figure 8: Interpreted probabilities for $s = 0.25$

If $s < 1$, the decision weight function has an inverse-S shape. Decision weights skewed towards $\frac{1}{2}$ relative to the stated probability, meaning that the credulous decision maker perceives either/or probabilities to be too similar to 50-50 shots. There is more sensitivity to changes in probabilities towards the endpoints than there is for middling probabilities.

In this context, the dialectic model of Equation 3.1 and its predictions are similar but not identical to the one-parameter functional form for the decision weight function proposed in Tversky and Kahneman (1992) and estimated there and in Camerer and Ho (1994) and Wu and Gonzalez (1996). The inverse-S shape for the decision weight function is well established empirically and is predicted by both frameworks. In both cases the single parameter (s here and γ in Tversky and Kahneman, 1992) is such that a value of 1 means 'correct' decision weights, although here of course I am interpreting the single parameter as capturing a specific psychological mechanism.

A key difference, however, between the two approaches is that the dialectic model is restricted to return a decision weight 0.5 for a stated probability of 0.5, and for its inverse-S shape to have

rotational symmetry of order 2 with respect to that point. We can see this in Figure 8; the decision maker's underweighing of stated probabilities below 0.5 is identical to their overweighing of stated probabilities above 0.5. By comparison, the Tversky and Kahneman (1992) specification allows for a more general inverse-S shape without either of those characteristics. Further emphasizing this distinction, Gonzalez and Wu (1999) estimates a two parameter functional form for the decision weight function in which one parameter controls the curvature of the inverse-S shape and the other its height; in the dialectic model s governs the curvature but the height is fixed to pass through the point $(0.5, 0.5)$.

On the other hand, a benefit of using Equation 3.1 is that its decision weights are less sensitive to s for very low or very high values of s. That means that we can use Equation 3.1 to model a greater range of decision makers: decision makers with very low and very high skepticism would have unrealistically extreme decision weights in the Tversky and Kahneman (1992) model but can be more plausibly captured by the dialectic model. Having said that, it is worth noting that the behaviorally compelling $s < 1$ I have argued for in the preceding two applications is quite similar to estimates of γ in the 0.5-0.8 range from the decision weight literature, a range in which the two models are in quite close accord.

Recall from Section 6.1, in the interpretation of the model as belief formation with more than one piece of evidence, that $s < 1$ allows that a decision maker's posterior belief after observing some signals can be outside of the range of those signals. Comparing this to the case here, with the stated probability as the lone piece of 'evidence', we see that a decision maker with $s < 1$ would display a high degree of both misinterpretation of probabilities *and* a stubbornness in the face of evidence. They place 'too much' weight on implausible interpretations of data relative to standard mathematical tools, in both the case with a single piece of information and the case with many pieces of information.

By contrast, a decision maker with skepticism of precisely 1 perceives probabilities perfectly, yet will still make mistakes when adjudicating a set of more than one piece of evidence. They form decision weights that are identical to the true, stated probabilities, but in the belief formation interpretation of the model $s = 1$ is still so low as to overweigh the more implausible frame. Loosely speaking, this means that pathologies of belief formation (applying the dialectic model

to frames over more than one piece of evidence) are 'easier' mistakes to make than errors in the interpretation of probabilities ((applying the dialectic model to a single stated probability), since more extreme values of s are required to make mistakes in the latter case. A decision maker with a fixed s of around 1 could display discrepancies from maximum likelihood when adjudicating more than one piece of evidence, but perceive probabilities quite accurately.

To visualize the effect of s in the decision weight application, Figure 9 illustrates the relationship between the skepticism parameter and the decision maker's interpretation of a stated probability of 0.01.

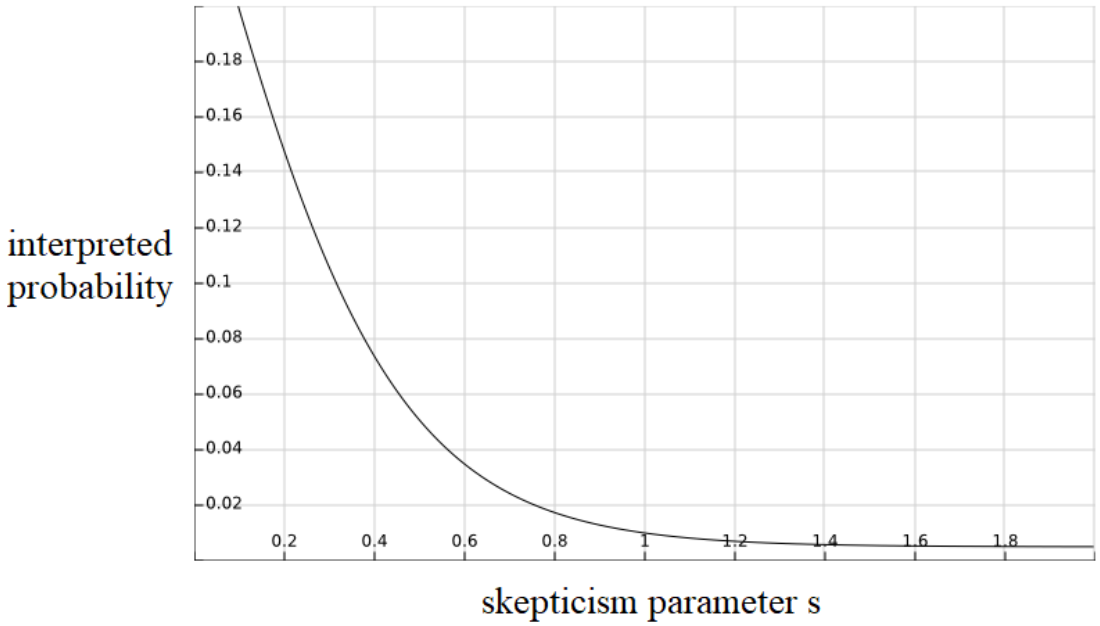

Figure 9: Interpreted probability for a stated probability of 0.01, by plausibility penalty p

A distribution of shapes for the probability weighting function is consistent with available experimental evidence on the best fit for such a function for different decision makers. Gonzalez and Wu (1999) presents evidence for different decision makers having weighting functions with different curvature. This is exactly the type of variation that is consistent with different values of the skepticism in the model in this paper, if we were to repeat the exercise of Figure 8 for different values of s.

The behavioral parameter skepticism in the dialectic model can therefore serve as a ready

interpretation of a simple functional form capturing the differences among decision-makers who perceive probabilities differently. It is a single parameter in the dialectic model that generates empirically relevant decision weight functions from a quite general model that, as we have already seen, is also empirically relevant in other, quite different applications. The dialectic model I have proposed here may then perhaps offer a possible unified behavioral explanation for various phenomena in one stroke.

7 Common elements from the political spin game and behavioral examples, and further applications

We have discussed three cases in which an individual forms beliefs according to the dialectic model. In all three examples, empirically plausible behaviors are consistent with the same feature of the model of dialectic belief formation: a weighing of competing explanations that is too naive, too credulous, insufficiently skeptical. Belief formation that is more stubborn than Bayes' rule, the overweighing of small probabilities and the certainty effect, and the fruitful manipulation of voters via extreme partisan framing are all consistent with the same, low skepticism in deliberation over competing frames. For maximally extreme frames to be a bad choice in the spin game, the model's parameter s must be quite a bit larger than what seems plausible based on the behavioral examples. Therefore, for a reasonable parameterization of the model, extremely skewed political spin is optimal for the senders, and the more skewed, the more effective at influencing the receiver.

While there are surely many ways to think about the idea of multiple selves or non-Bayesian reasoning, I think the idea of the dialectic with skepticism is an interesting model to consider. The idea of dialectic deliberation has a long history in philosophy, psychology, and literature, and I have argued that the model based on this idea is consistent with our understanding of some ways in which the world contrasts with a world of Bayesians. According to this line of thinking, a little more skepticism, a little less credulity, in the population would be a valuable thing.

In conclusion I would like to briefly mention some other settings in which a dialectic decision

maker would behave in an empirically interesting way. What would we predict if we took an excessively credulous dialectic decision maker to other settings?

Consider an effort provision problem. An entrepreneur receives signals about the viability of their current project or idea. They must then decide whether to invest in or devote effort to developing the idea. In this context, an excessively credulous entrepreneur would be hubristic with respect to the viability of their ideas. Similarly, in financial contracting, say that an outside investor seeks to evaluate a project's quality from signals. Low skepticism here would manifest as systematic overinvestment in low quality projects and underinvestment in high quality projects.

If we apply the model to life cycle consumption and savings, we may view the evidence that the decision maker observes as noisy information about their own future prospects. An excessively credulous decision maker who saw bad signals would save less than is optimal since they would be insufficiently swayed by the possibility that their future prospects are low. On the other hand, if they saw good signals they would save more than is optimal since they would be insufficiently swayed by the possibility that they are on a good trajectory.

Since the model of an excessively credulous dialectic decision maker matches quite well some stylized facts about how we see people behave and decide in real situations, developing these and other applications could yield some interesting insight into how we might better design institutions to account for this possible avenue by which people's decisions are imperfect.

References

Abrams v. United States (1919). 250 U.S. 616.

Afrouzi, H., Kwon, S. Y., Landier, A., Ma, Y., and Thesmar, D. (2020). Overreaction and working memory. *NBER Working Papers*, (27947).

Ambrus, A. and Rozen, K. (2013). Rationalising Choice with Multi-self Models. *The Economic Journal*, 125(585):1136–1156.

Ambuehl, S. and Li, S. (2018). Belief updating and the demand for information. *Games and Economic Behavior*, 109:21 – 39.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3):307 – 343.

Bénabou, R., Falk, A., and Tirole, J. (2018). Narratives, imperatives, and moral reasoning. Working Paper 24798, National Bureau of Economic Research.

Bénabou, R. and Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–64.

Benjamin, D. J. (2018). Errors in probabilistic reasoning and judgment biases. NBER Working Paper No. 25200.

Bisgaard, M. (2015). Bias will find a way: Economic perceptions, attributions of blame, and partisan-motivated reasoning during crisis. *The Journal of Politics*, 77(3):849–860.

Black, R. E. (2014). From Charioteer Myth to Shoulder Angel: A Rhetorical Look at Our Divided Soul. *Colloquy*, 10:36–49.

Bull, J. and Watson, J. (2018). Statistical evidence and the problem of robust litigation. http://faculty.fiu.edu/ bullj/BullWatsonSEPRL.PDF.

Camerer, C. F. and Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2):167–196.

Campbell, T. H. and Kay, A. C. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology*, 107(5):809.

Colbert, S. (2006). Correspondents' Dinner Speech. Presented at the 2006 White House Correspondents' Association Dinner, Washington, D.C.

Daughety, A. F. and Reinganum, J. F. (2000). On the economics of trials: adversarial process, evidence, and equilibrium bias. *Journal of Law, Economics, and Organization*, 16(2):365–394.

Davis, E. H. (2004). Protecting the marketplace of ideas: The first amendment and public school teachers' classroom speech. *First Amend. L. Rev.*, 3:335.

De Clippel, G. and Eliaz, K. (2012). Reason-based choice: A bargaining rationale for the attraction and compromise effects. *Theoretical Economics*, 7(1):125–162.

Dean, M. and Ortoleva, P. (2019). The empirical relationship between nonstandard economic behaviors. *Proceedings of the National Academy of Sciences*, 116(33):16262–16267.

Ding, M. (2007). A Theory of Intraperson Games. *Journal of Marketing*, 71:2.

Douglas, K. M. and Sutton, R. M. (2015). Climate change: Why the conspiracy theories are dangerous. *Bulletin of the Atomic Scientists*, 71(2):98–106.

Douglas, K. M., Sutton, R. M., Callan, M. J., Dawtry, R. J., and Harvey, A. J. (2016). Someone is pulling the strings: Hypersensitive agency detection and belief in conspiracy theories. *Thinking & Reasoning*, 22(1):57–77.

Eliaz, K. and Spiegler, R. (2020). A model of competing narratives. *American Economic Review*, 110(12):3786–3816.

Enke, B. and Graeber, T. (2021). Cognitive uncertainty. *NBER Working Papers*, (26518).

Epstein, L. G., Noor, J., and Sandroni, A. (2010). Non-Bayesian Learning. *The B.E. Journal of Theoretical Economics*, 10(1).

Fessler, D. M. T., Pisor, A. C., and Holbrook, C. (2017). Political orientation predicts credulity regarding putative hazards. *Psychological Science*, 28(5):651–660. PMID: 28362568.

Froeb, L. M., Ganglmair, B., and Tschantz, S. (2016). Adversarial Decision Making: Choosing between Models Constructed by Interested Parties. *The Journal of Law and Economics*, 59(3):527–548.

Gentzkow, M. and Kamenica, E. (2016). Competition in persuasion. *The Review of Economic Studies*, 84(1):300–322.

Gonzalez, R. and Wu, G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, 38(1):129–166.

Guess, A. and Coppock, A. (2018). Does Counter-Attitudinal Information Cause Backlash? Results from Three Large Survey Experiments. *British Journal of Political Science*.

Hart, J. and Graether, M. (2018). Something's going on here. *Journal of Individual Differences*.

Hirshleifer, J. and Osborne, E. (2001). Truth, effort, and the legal battle. *Public Choice*, 108(1-2):169–195.

Jamison, J. and Wegener, J. (2010). Multiple selves in intertemporal choice. *Journal of Economic Psychology*, 31(5):832 – 839. Special Issue on Decision Neuroscience.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291.

Kalai, G., Rubinstein, A., and Spiegler, R. (2002). Rationalizing Choice Functions By Multiple Rationales. *Econometrica*, 70(6):2481–2488.

Klar, S. (2014). A Multidimensional Study of Ideological Preferences and Priorities among the American Public. *Public Opinion Quarterly*, 78(S1):344–359.

Lantian, A., Bagneux, V., Delouvée, S., and Gauvrit, N. (2020). Maybe a free thinker but not a critical one: High conspiracy belief is associated with low critical thinking ability. *Applied Cognitive Psychology*.

Lee, T.-T. (2005). The liberal media myth revisited: An examination of factors influencing perceptions of media bias. *Journal of Broadcasting & Electronic Media*, 49(1):43–64.

Lombardi, C. (2019). The Illusion of a "Marketplace of Ideas" and the Right to Truth. *American Affairs*, 3(1):198–209.

Marwick, A. and Lewis, R. (2017). Media manipulation and disinformation online. *Data & Society Research Institute*.

Miller, J. M., Saunders, K. L., and Farhart, C. E. (2016). Conspiracy endorsement as motivated reasoning: The moderating roles of political knowledge and trust. *American Journal of Political Science*, 60(4):824–844.

Nisbet, E. C., Cooper, K. E., and Garrett, R. K. (2015). The partisan brain: How dissonant science messages lead conservatives and liberals to (dis)trust science. *The ANNALS of the American Academy of Political and Social Science*, 658(1):36–66.

Nunziato, D. C. (2018). The marketplace of ideas online. *Notre Dame L. Rev.*, 94:1519.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1):103–124.

Ouwersloot, H., Nijkamp, P., and Rietveld, P. (1998). Errors in probability updating behaviour: Measurement and impact analysis. *Journal of Economic Psychology*, 19(5):535 – 563.

Peleg, B. and Yaari, M. E. (1973). On the existence of a consistent course of action when tastes are changing. *The Review of Economic Studies*, 40(3):391–401.

Pennycook, G., Cheyne, J. A., Barr, N., Fugelsang, J. A., et al. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision making*, 10(6):549–563.

Pennycook, G. and Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188:39–50.

Peterson, C. R. and Miller, A. J. (1965). Sensitivity of subjective probability revision. *Journal of experimental psychology*, 70(1):117.

Phillips, L. D. and Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of experimental psychology*, 72(3):346.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, 64(4):1021–1044.

Roth, M. S. (2017). How free should free speech be on campus? *Washington Post*, Sept. 22, 2017.

Rustichini, A., DeYoung, C. G., Anderson, J. E., and Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, 64:122–137.

Schwartzstein, J. and Sunderam, A. (2019). Using models to persuade. Working Paper 26109, National Bureau of Economic Research.

Skaperdas, S. and Vaidya, S. (2012). Persuasion as a contest. *Economic Theory*, 51(2):465–486.

Slovic, P. and Lichtenstein, S. (1971). Comparison of bayesian and regression approaches to the study of information processing in judgment. *Organizational behavior and human performance*, 6(6):649–744.

Stango, V., Yoong, J., and Zinman, J. (2017). The quest for parsimony in behavioral economics: New methods and evidence on three fronts. NBER Working Paper No. 23057.

Taber, C. S. and Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3):755–769.

Treier, S. and Hillygus, D. S. (2009). The Nature of Political Ideology in the Contemporary Electorate. *Public Opinion Quarterly*, 73(4):679–703.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Van Prooijen, J.-W., Krouwel, A. P., and Pollet, T. V. (2015). Political extremism predicts belief in conspiracy theories. *Social Psychological and Personality Science*, 6(5):570–578.

Watts, M. D., Domke, D., Shah, D. V., and Fan, D. P. (1999). Elite cues and media bias in presidential campaigns: Explaining public perceptions of a liberal press. *Communication Research*, 26(2):144–175.

Weber, M. and Camerer, C. F. (1998). The disposition effect in securities trading: an experimental analysis. *Journal of Economic Behavior & Organization*, 33(2):167 – 184.

White, M. D. (2006). Multiple utilities and weakness of will: A Kantian perspective. *Review of Social Economy*, 64(1):1–20.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management science*, 42(12):1676–1690.

A Result 1

Let us call the upper and lower bounds chosen by sender L $\bar{L}$ and $\underline{L}$ respectively, and similarly the upper and lower bounds chosen by sender R $\bar{R}$ and $\underline{R}$ respectively. The receiver's assessment of the frames is given by:

$$\hat{\mu} = \frac{w_H^s}{w_L^s + w_H^s}\mu_L + \frac{w_L^s}{w_L^s + w_H^s}\mu_H \tag{A.1}$$

$$= \frac{1}{2}(\underline{L} + \bar{L})\left[\frac{(\bar{H} - \underline{H})^s}{(\bar{L} - \underline{L})^s + (\bar{H} - \underline{H})^s}\right] + \frac{1}{2}(\underline{H} + \bar{H})\left[\frac{(\bar{L} - \underline{L})^s}{(\bar{L} - \underline{L})^s + (\bar{H} - \underline{H})^s}\right]. \tag{A.2}$$

Recall that each sender must use a frame with strictly positive width (this is by assumption, but it is equivalent to saying that the distribution of the evidence itself is not degenerate). Consider the perspective of sender R deciding where to set the upper bound of their frame. The derivative of the receiver's assessment with respect to $\bar{H}$ is

$$\frac{0.5(\bar{L} - \underline{L})^s((\bar{L} - \underline{L})^s + (\bar{H} - \underline{H})^{s-1}(s(\bar{L} + \underline{L} - \bar{H} - \underline{H}) + \bar{H} - \underline{H}))}{((\bar{L} - \underline{L})^s + (\bar{H} - \underline{H})^s)^2} \tag{A.3}$$

The denominator here is certainly positive, as is the first term, $0.5(\bar{L} - \underline{L})^s$, in the numerator. We may therefore say that the whole expression is positive if

$$((\bar{L} - \underline{L})^s + (\bar{H} - \underline{H})^{s-1}(s(\bar{L} + \underline{L} - \bar{H} - \underline{H}) + \bar{H} - \underline{H})) > 0. \tag{A.4}$$

When evaluated at the $s = 0$, this expression is certainly positive for any values of the bounds on the two frames. When evaluated at $s = 1$, this expression is equivalent to

$$\bar{L}^2 - \bar{L}\underline{L} - \bar{L}\underline{H} + \underline{L}\underline{H} \tag{A.5}$$

$$= \bar{L} - \underline{L} + \left(\frac{\underline{L}}{\bar{L}}\right)\underline{H} - \underline{H} \tag{A.6}$$

which is certainly larger than zero since $\bar{L} > \underline{L}$ and $\bar{L} < 1$.

This means that sender R can always increase the receiver's assessment $\hat{\mu}$ in their preferred direction (toward the endpoint 1 of the spectrum) by increasing $\bar{H}$ so long as s is sufficiently

small. Since the problem for sender L is symmetric, for both senders to extend their frames as much as possible toward their preferred endpoint forms a Nash equilibrium pair.

B Results 2 and 3

These results follow directly from the characteristics of Result 1 and the properties of the uniform distribution.

C Result 4

$\frac{d\hat{\mu}}{d\underline{L}} > 0$, so that the marginal value to a party of releasing evidence that forces their opponent to expand their frame is always positive. In the case in which the parties use maximally extreme frames, $\frac{d^2\hat{\mu}}{d\underline{L}^2} < 0$, so that the marginal value of these actions is higher when their opponent's existing frame is narrower in scope.

Consider the sender who prefers the receiver to make a high assessment, who we may call the high type sender for convenience. Let us consider the marginal effect of raising $\bar{L}$ on the receiver's assessment $\hat{\mu}$. This captures the effect of the high type sender releasing evidence unfavorable to their opponent.

From Equation A.2, the derivative of $\hat{\mu}$ with respect to $\bar{L}$ is

$$\frac{(\bar{H}-\underline{H})^s((\bar{L}-\underline{L})(\bar{H}-\underline{H})^s+(\bar{L}-\underline{L})^s(s(\bar{H}+\underline{H}-\bar{L}-\underline{L})+\bar{L}-\underline{L}))}{(\bar{L}-\underline{L})((\bar{H}-\underline{H})^s+(\bar{L}-\underline{L})^s)^2}. \tag{C.1}$$

This expression is always positive, since $\bar{L} > \underline{L}$, $\bar{H} > \underline{H}$, and $\bar{H}+\underline{H} > \bar{L}+\underline{L}$. This means that the receiver's assessment $\hat{\mu}$ always moves in the direction favorable to the high type sender when they are able to force $\bar{L}$ to be higher.

Next consider the case in which the senders use maximally extreme frames, so that $\bar{H} = 1$ and $\underline{L} = 0$. The second derivative $\frac{d^2\hat{\mu}}{d\bar{L}^2}$ is given by

$$\frac{0.5s(1-\underline{H})^s\bar{L}^s(\underline{H}(s-1)-s(\bar{L}-1)-\bar{L}-1)-\bar{L}^{2s}(\underline{H}(s+1)-s(\bar{L}-1)+\bar{L}+1)}{\bar{L}^2((1-\underline{H})^s+\bar{L}^s)^3}. \tag{C.2}$$

The denominator and the first term $0.5s(1-\underline{H})^s\bar{L}^s$ in the numerator are certainly positive, since $\bar{L}$ and $\underline{H}$ are between 0 and 1. This leaves the term

$$(\underline{H}(s-1)-s(\bar{L}-1)-\bar{L}-1)-\bar{L}^{2s}(\underline{H}(s+1)-s(\bar{L}-1)+\bar{L}+1). \tag{C.3}$$

The first part of this expression is certainly negative, and the second part being subtracted is certainly positive. This means that the second derivative as a whole is negative. This means that an increase in $\bar{L}$ has a smaller positive effect on $\hat{\mu}$ when $\bar{L}$ is higher. The largest positive effect comes when the upper bound of the existing evidence is least favorable to the high type sender.