Strategic communication of narratives

Gerrit Bauch[*] and Manuel Foerster[†]

February 5, 2026

Abstract

We model the communication of narratives as a cheap-talk game under model uncertainty. The sender has private information about the true data-generating process of public data. The receiver is uncertain about how to interpret the data, but aware of the sender's incentives to strategically provide interpretations (*"narratives"*). We introduce a general class of ambiguity rules resolving the receiver's residual uncertainty. Equilibria are characterized by a positive integer N: an algorithm yields an equilibrium with n different actions for each $1 \leq n \leq N$. The persuasive power of the sender is weaker in the sense of state-wise dominance than with a naïve receiver.

Keywords: Narratives, model uncertainty, ambiguity, strategic communication, cheap-talk game

JEL: C72, D81, D82, D83

1 Introduction

Convincing others often involves the interpretation of past data. When a politician runs for office, she will try to cast her past achievements in the most favorable light. She

[*]Center for Mathematical Economics, Bielefeld University, PO Box 10 01 31, 33501 Bielefeld, Germany. Email: gerrit.bauch@uni-bielefeld.de.

[†]Center for Mathematical Economics, Bielefeld University, PO Box 10 01 31, 33501 Bielefeld, Germany. Email: manuel.foerster@uni-bielefeld.de.

may do so by interpreting past events, claiming a low unemployment rate for herself while blaming economic recessions on external factors such as decisions made by the predecessor or the global economy. By crafting a *narrative* that carefully arranges the arguments *seemingly* reasonably, she hopes to convince voters of her ability. However, rational voters understand that the politician makes these claims not least in order to get elected. Consequently, we expect them not to take her layout of arguments at face value. Rather, they should correct for the politician's own bias, disciplining her ability to manipulate opinions in her favor.

This work develops a novel cheap-talk game under model uncertainty that captures the strategic communication of narratives between a biased sender (she) and an unbiased receiver (he). In contrast to Schwartzstein and Sunderam (2021, henceforth S&S), we assume that the receiver is aware of the sender's strategic incentives to manipulate him, and we introduce a general class of ambiguity rules to resolve any remaining uncertainty. This consequently limits the sender's persuasive power, effectively restricting the set of actions that she can induce.

In our framework, both agents have a common prior belief about the true state of nature θ, e.g., the sender's ability. While the sender also knows the exact underlying process that has generated observable data h from θ, the receiver faces *model uncertainty* (or *ambiguity*) in the sense that he does not know how to interpret the data. For instance, it may be unclear whether a low unemployment rate is linked to the politician's own decisions or to external factors. Any such interpretation is called a narrative or a *model* represented by a likelihood function, providing a probabilistic way to make sense out of observed data. The finite set M of feasible models contains the true data-generating process.

The biased sender knows the true model and submits a cheap-talk message to the receiver, aiming at making him interpret the data h in her favor. Any such message provides a reasoning of how to interpret the data, e.g., claiming a low unemployment rate for oneself while blaming a recession on external factors, formalizing the strategic use of narratives. Upon observing the sender's report, the receiver narrows down the set of feasible models, resolves any remaining uncertainty by an *ambiguity rule*, and then takes an action. The novel notion of ambiguity rules enables us to treat different preferences in the face of uncertainty in a unified way, including max-min expected utility (Gilboa and Schmeidler, 1989), the smooth model of decision making (Klibanoff et al., 2005) and

2

its special case, the classical Bayesian approach. In addition, our framework also lifts the maximum likelihood expected utility, the most prominent preference in the recent literature on narratives (S&S; Jain, 2023; Aina, 2024), to an equilibrium setting.

Communication of narratives thus becomes a game in which both agents strategically seek to maximize their own utility. Observe that we can treat every observable data h as indexing a different game, since h is commonly known ex ante. Given observable data h, an equilibrium is then a stable strategy profile: The sender maps models to messages in a way that maximizes her expected payoff given the true model and the receiver's response. Upon observing a message, the receiver first updates the considered set of models and then applies an ambiguity rule to resolve any remaining uncertainty and determine his best response.

Our analysis characterizes equilibria as partitions of models into intervals, where models are ordered by the associated receiver's bliss points. Any such interval $\tilde{M}$ corresponds to a cheap-talk message, the indicative meaning of which is "the true model belongs to $\tilde{M}$". By means of his ambiguity rule, the receiver assigns $\tilde{M}$ a unique optimal action. Among all equilibria in a given setting, there is a maximum number N of induced actions. We provide an algorithm that constructs an equilibrium with $n-1$ distinct actions given one with n distinct actions, thus proving that for any $1 \leq n \leq N$ there is an equilibrium inducing n distinct actions. The size of the conflict of interest between the sender and the receiver controls how much information the sender is willing to reveal. We first show that the maximum number N of induced actions is decreasing in the conflict, and then characterize informativeness thresholds that mark when the most finely granulated information revelation is full disclosure or partial disclosure, respectively.

In order to derive further insights, we then consider a running example based on a modified beta-binomial model, in which each of the finitely many data points may either depend on the state of nature or be independent of it, and thus noise. We first show that the threshold regarding full disclosure need not be quasi-concave in a summary statistic of the observable data. Nevertheless, it attains its minimum at extreme histories, implying that these are the most difficult to 'explain'. Second, the threshold regarding partial disclosure also depends on how the receiver deals with the residual uncertainty, that is, on the ambiguity rule. Consequently, the threshold may not be symmetric in the summary statistic. Moreover, the higher the action induced by the ambiguity rule upon an upward deviation in an optimal bipartite partition, the higher the threshold for

3

partial disclosure.

Finally, we compare the outcome of equilibrium play under the maximum likelihood rule to the benchmark model of S&S, in which the receiver acts in a "naïve" fashion: Being offered a narrative of how to interpret the observed data, he adopts the proposed model if it fits the data better than a default model. The receiver thus takes the sender's suggestion at face value and ignores her strategic incentives. Unsurprisingly, a sender type who can do better than under truthful communication when facing a receiver who engages in equilibrium play will also be able to do so when facing a naïve receiver, as incentive constraints are weaker in the latter case. However, we show that the sender has more *persuasive power* facing a naïve receiver in an even stronger sense, in the spirit of state-wise dominance: She obtains larger expected utility facing a naïve receiver under *any* model. Thus, the sender generally prefers to deal with a naïve receiver. Interestingly, the receiver does not generally prefer to be rational himself. He sometimes would prefer to be naïve, as tight incentive constraints under equilibrium play imply missing out on some chances to mutually increase expected utility.

Revisiting the motivating example, our results show that a politician who runs for office and tries to convince voters who are aware of strategic incentives will provide coarse narratives. This leaves voters with residual uncertainty about the true data-generating process, which we propose to resolve by using an ambiguity rule. Voters' strategic sophistication constrains the politician's persuasive power compared to the scenario involving naïve voters.

Related literature. Shiller (2017) pointed out the necessity to study narratives from an economic point of view in order to better understand the effects of factual and non-factual information on decisions. The economic literature has thus far conceptualized narratives in different ways.

Our work ties up with the approach that views narratives as likelihood functions. These establish a probabilistic link between observable data and the parameter of interest. Most closely related to our work is S&S. The authors let a biased narrator provide narratives to a receiver, aiming at making observable data plausible in her favor. The receiver initially considers a default model, which can in fact be the true data-generating process, but adopts the narrative if it fits the observable data better than his default model. The receiver thus does not consider the strategic incentives of the sender. As a main result of theirs, the persuader finds it easier to manipulate the receiver's beliefs

if the default model fits the data poorly. Aina (2024) considers a similar setting but assumes the persuader commits to a set of narratives before the public data has been realized, while Jain (2023) introduces a third party who can strategically restrict the narrator's pool of feasible models. In a similar vein, Eliaz et al. (2021) allow the sender to strategically provide selective interpretations for her own messages. In a laboratory experiment, Barron and Fries (2024) find evidence that the likelihood of a narrative is a key determinant for persuasiveness and that individuals use this fact to tailor narratives to their own benefit.

Another way of formalizing narratives is by interpreting a directed acyclic graph as a causal model. Eliaz and Spiegler (2020) employ an equilibrium concept of long-run distribution over narrative-policy pairs and show that prevailing models can be misspecified causal models. Eliaz et al. (2025) and Bilotta and Manferdini (2025) take the approach to a political-economics framework. Bénabou et al. (2020) model narratives as messages about externalities in a psychological game; see Foerster and van der Weele (2021) for a closely related approach.

Our model combines the literature on narratives as likelihood functions with the one on cheap-talk games. Analogously to the seminal work of Crawford and Sobel (1982), the biased sender strategically partitions unknown information into intervals. While they do this for the infinite state space itself, assuming Bayesian agents, we do so for the discrete model space and introduce a general class of ambiguity rules to resolve the receiver's uncertainty, including the Bayesian approach. The combination of a cheap-talk game with ambiguity relates our paper to Colo (2024), who assumes the receiver's model uncertainty concerns the prior distribution of the state of the world directly instead of through the interpretation of public data. Colo thus considers cheap-talk communication over a classical interval of beliefs, while we study communication over a discrete set of narratives, making none of his major theorems carry over to our set-up.[1]

The importance of uncertainty when it comes to narratives has thus far been acknowledged by researchers on climate change as well as psychologists (Pedde et al., 2019; Constantino and Weber, 2021; Garcia-Lorenzo, 2010). We make ambiguity an integral part of our economic model. To resolve any remaining uncertainty in this setting, agents need a decision rule to determine their optimal action. While Colo (2024) separately studies the pessimistic max-min rule (Gilboa and Schmeidler, 1989) and a Bayesian ap-

---

[1]See the discussion below Proposition 2 on p. 20 for details.

proach, we introduce ambiguity rules as a unified framework to study a vast class of decision rules under uncertainty. Apart from the max-min rule and the Bayesian approach, it includes the maximum likelihood rule, dominant in the economic literature on narratives, as well as the smooth rule proposed by Klibanoff et al. (2005), which allows for different ambiguity attitudes.

To our knowledge, we are the first to model the communication of narratives as a cheap-talk game under model uncertainty, formalizing narratives as likelihood functions and introducing a general class of ambiguity rules, while the receiver takes the sender's strategic incentives into account.

The remainder of the article is organized as follows. Section 2 introduces the formal set-up of the game theoretic model, defines ambiguity rules and the equilibrium concept, and provides the running example. Section 3 characterizes equilibria and the informativeness thresholds. Within the running example, we provide further insights into the informative thresholds in Section 4. Section 5 classifies the persuasive power of the sender by comparing the equilibrium model to the one with a naïve receiver. Section 6 concludes and discusses some of our modelling assumptions.

2   Model and notation

Two agents, a *sender* (S or she) and a *receiver* (R or he), both observe a history of past outcomes (or a public signal) $h \in H$ about the state of nature $\theta \in \Theta = [0, 1]$. We assume that H is finite. The common prior over the state is a distribution F_0 on Θ with continuous and strictly positive density f_0. A *model* or *narrative* $m \in M$ parameterizes a likelihood function $\{\pi_m(\cdot|\theta)\}_{\theta \in \Theta}$, where $\pi_m(h|\theta)$ denotes the likelihood of history h given state θ under model m. The *expected fit* of a model m given data h is $\Pr(h \mid m, F_0) := \int_\Theta \pi_m(h \mid \theta) \, \mathrm{d}F_0(\theta)$. We assume that M is a finite set and under all $m \in M$ every history $h \in H$ has positive probability given the prior, i.e., $\Pr(h \mid m, F_0) > 0$.

The state θ is drawn according to F_0 while there is (model) uncertainty about the *true model* m^T. The history h is generated according to the true model m^T and the state θ, and observed by both agents. While the sender also learns the true model m^T and updates her prior to the posterior $F_{m^T,h}$ using Bayes' rule, the receiver faces full model uncertainty about the true model $m^T \in M$.[2]

---

[2]C.f. Muraviev et al. (2017) for a way to model extensive form games under ambiguity. The receiver's initial belief can formally be modeled by a *capacity* μ_0 (Choquet, 1954) such that $\mu_0(M') = 1$ if $M' = M$

After the sender has observed the history h and the true model m^T, she submits a cheap-talk report (or message) $r \in \mathcal{R}$, where $\mathcal{R}$ is a rich message space.[3] Upon observing the report r, the receiver can restrict the set of considered models to a *minimal feasible set* $\tilde{M}$. $\tilde{M}$ is the smallest collection of models the receiver can and does not want to exclude from his considerations, i.e., the indicative meaning of the submitted report is "$m^T \in \tilde{M}$". Finally, the receiver takes an action $a \in A = \mathbb{R}$.

2.1 Payoffs

The sender's payoff function $u_S(a, \theta, b)$ depends on the receiver's action $a \in A$, the state $\theta \in \Theta$, and the sender's bias $b > 0$, a commonly known constant that measures the conflict of interest between sender and receiver. We assume that $u_S(a, \theta, b)$ is twice continuously differentiable and strictly concave in a, with a unique maximum for fixed θ and b. Furthermore, we impose *supermodularity*, that is $\frac{\partial^2 u_S(a, \theta, b)}{\partial a \partial b} > 0$, a sorting condition common in cheap talk that ensures that the sender would like to induce a higher action with a higher bias. The receiver is considered unbiased and his payoff function is $u_R(a, \theta) := u_S(a, \theta, 0)$.

Note that the expected utility function $a \mapsto U_{m,h}(a, b) := \mathbb{E}[u_S(a, \theta, b) \mid m, h]$ after an update on the information revealed through the datum (m, h) is then also strictly concave in a and its unique maximizer, denoted by $a(m, h, b)$, is strictly increasing in b. For any fixed h, we assume *consistent ordering* across biases, meaning that both agents share the ordering of bliss points across models.

Assumption 1 (Consistent ordering). For any history h, models m, m' and bias b, we have $a(m', h, 0) \leq a(m, h, 0)$ if and only if $a(m', h, b) \leq a(m, h, b)$.

Since bliss points are equally ordered, we can reduce the structural complexity of the model set by projecting M on the real line. To this end, we define a total preorder of M by setting

$$m' \leq (<) \, m \text{ if and only if } a(m', h, 0) \leq (<) \, a(m, h, 0).$$

We call $M' \subseteq M$ an *interval* if for all $m, m', m'' \in M$ with $m < m' < m''$ and $m, m'' \in M'$, we also find $m' \in M'$. Our second assumption states that the set of models for which one action is better than another is an interval, cf. Kartik et al. (2024).

---

and $\mu_0(M') = 0$ if $M' \subsetneq M$, reflecting full uncertainty.

[3]A message space in our setting is rich if is has at least $\#M$ elements. Hence, communication is not artificially restricted and allows for a perfect discrimination between models.

Assumption 2 (Interval choice)**.** For any history h, bias b, ordered models $m < m' < m''$ and actions $a, a' \in A$, we have $U_{m',h}(a, b) \geq U_{m',h}(a', b)$ whenever $U_{m,h}(a, b) \geq U_{m,h}(a', b)$ and $U_{m'',h}(a, b) \geq U_{m'',h}(a', b)$.

Interval choice and similar assumptions, like convex choice, are common in many cheap-talk, observational learning, and collective choice settings, and closely related to single crossing properties, cf. Kartik et al. (2024); Kartik and Kleiner (2024).

Dating back to at least Crawford and Sobel (1982), the most pervasive utility function in communication games is the *quadratic loss functional*. It fulfills consistent ordering and interval choice, irrespective of any assumptions on M and its associated likelihood functions.

Example 1. Let $u_S(a, \theta, b) = -(\theta + b - a)^2$, which is a strictly concave function in a and supermodular. Under any history h and model m, the bliss point is $a(m, h, b) = \mathbb{E}[\theta \mid m, h] + b$. That is, bliss points are ordered by their expected state across biases and thus consistently ordered. Interval choice holds because $\mathbb{E}[u_S(a, \theta, b) \mid m, h]$ depends only on the Euclidean distance of a to $\mathbb{E}[\theta \mid m, h] + b$.

2.2 Ambiguity rules

After receiving the sender's message, the receiver still faces uncertainty in form of the minimal feasible set $\tilde{M}$. To resolve the remaining uncertainty, the receiver may adopt any utility function of the following class.

Definition 1 (Ambiguity rules)**.** An *ambiguity rule* (for the receiver) is a function $U \colon 2^M \times H \times A \to \mathbb{R}, (\tilde{M}, h, a) \mapsto U_{\tilde{M},h}(a)$, assigning a utility level to each action a when facing the minimal feasible set $\tilde{M}$ under history h with the following properties:

  (i) $U_{\{m\},h}(a) = \mathbb{E}[u_R(a, \theta) \mid m, h]$,

  (ii) $U_{\tilde{M},h}(\cdot)$ is strictly concave with unique maximizer $a(\tilde{M}, h)$ and the maximizers fulfill $a(M_1 \cup M_2, h) \in \mathrm{conv}(a(M_1, h), a(M_2, h))$ for all $M_1, M_2 \subseteq M$.

A decision maker takes an ambiguity rule as a measure for the anticipated utility. We require that such a rule is based on Bayesian expected utilities (part (i)) and retains a bliss point that hedges against increasing ambiguity by diversification (part (ii)). Note that

our notion of ambiguity rules satisfies a mild version of uncertainty aversion.[4] Our novel notion of ambiguity rules covers a large scope of preferences prominent in the literature on decision theory under uncertainty and allows us to study equilibrium properties in our framework in a unified way.

Example 2. Consider the minimal feasible set $\tilde{M}$ given a history h.

(i) *Maximum likelihood expected utility (MLEU).* Let $\succ_{M,h}$ be any *strict* ordering on M that respects the expected fit, i.e., $m \succ_{M,h} m'$ implies $\Pr(h \mid m, F_0) \geq \Pr(h \mid m', F_0)$. The maximum-likelihood expected utility (w.r.t. $\succ_{M,h}$) is given by the expected utility

$$U_{\tilde{M},h}(a) = \mathbb{E}[u_R(a, \theta) \mid \tilde{m}, h] = \int_0^1 u_R(a, \theta)dF_{\tilde{m},h}(\theta),$$

where $\tilde{m}$ is the largest element in $\tilde{M}$ w.r.t. $\succ_{M,h}$. Under the MLEU ambiguity rule, the receiver maximizes his expected utility w.r.t. a narrative from the minimal feasible set $\tilde{M}$ that is most likely to explain the observed data h. The strict ordering $\succ_{M,h}$ serves as a tiebreaker.[5]

(ii) *Max-min expected utility (MEU).* The max-min expected utility of Gilboa and Schmeidler (1989) is given by

$$U_{\tilde{M},h}(a) = \min_{m \in \tilde{M}} \mathbb{E}[u_R(a, \theta) \mid m, h] = \min_{m \in \tilde{M}} \int_0^1 u_R(a, \theta)dF_{m,h}(\theta).$$

Under the MEU ambiguity rule, the receiver seeks to maximize his worst-case expected utility given $\tilde{M}$.

(iii) *Smooth model.* The smooth model of decision making under ambiguity of Klibanoff et al. (2005) is given by

$$U_{\tilde{M},h}(a) = \sum_{m \in \tilde{M}} \upsilon(m \mid \tilde{M}, h) \cdot \phi\left(\mathbb{E}[u_R(a, \theta) \mid m, h]\right),$$

---

[4]For a fixed minimal feasible set $\tilde{M}$ and history h, every action $a \in A$ induces an act that maps models $m \in \tilde{M}$ to payoff consequences $U_{m,h}(a, 0)$. Uncertainty aversion à la Gilboa and Schmeidler (1989, Axiom A.5.) then follows from Definition 1 (ii).

[5]Using a tiebreaker is important for two reasons. First, without tiebreaking, the hedging property (ii) of an ambiguity rule might fail. Second, and in contrast to the literature, we cannot resolve ties by applying a "sender-preferred" action as the receiver faces ambiguity about the true model and thus about the sender's expected utility. Note also that a choice function $C_h \colon 2^M \to M$, $C_h(\tilde{M}) \in \arg\max_{m \in \tilde{M}} \Pr(h \mid m, F_0)$ with the property that if $M_1 \subseteq M_2$ and $C_h(M_2) \in M_1$, then $C_h(M_1) = C_h(M_2)$ would serve the same purpose as the strict ordering.

where $\phi\colon \mathbb{R} \to \mathbb{R}$ is a function capturing the ambiguity attitude and υ is a probability measure on M, with $\upsilon(m \mid \tilde{M}, h) = \frac{\mathbb{1}_{\tilde{M}}(m)\upsilon(m)\Pr(h|m,F_0)}{\sum_{\tilde{m}\in\tilde{M}}\upsilon(\tilde{m})\Pr(h|\tilde{m},F_0)}$ by Bayes' rule. Note that the smooth model becomes the standard Bayesian expected utility with prior υ if we use a linear ambiguity index ϕ.

2.3   Equilibrium concept

We now address equilibrium behavior. A strategy for the sender is a function $\sigma\colon H \times M \to \mathcal{R}$ that assigns a report to each history and model. A strategy for the receiver is a function $\rho\colon H \times \mathcal{R} \to A$ that assigns an action to each history and report. Note that the history h is known to both agents ex ante, i.e., before they make their move. Consequently, one can treat every h as indexing a different game.[6] In the following, we thus omit the dependence of the agents' strategies on h and consider $\sigma\colon M \to \mathcal{R}$ and $\rho\colon \mathcal{R} \to A$ for a given history h. Likewise, we frequently drop the dependence on h from the bliss points and simply write $a(m, b)$ and $a(\tilde{M})$ instead of $a(m, h, b)$ and $a(\tilde{M}, h)$, respectively, whenever h is clear from the context.

Upon observing a report $r \in \sigma(M)$, and given the sender strategy σ, the receiver can discard all models under which r is not sent. The minimal feasible set is thus $\tilde{M} = \sigma^{-1}(r)$ and the indicative meaning of message r is "$m^T \in \tilde{M}$".[7]

We restrict attention to pure strategies and assume that the receiver takes an on-equilibrium-path action if she observes an off-equilibrium message, which implies that we can ignore such deviations.[8]

An *equilibrium* (σ, ρ) (under ambiguity rule $U_{\tilde{M},h}$) then needs to entail mutual best replies:

(i) $\sigma(m^T) \in \arg\max_r U_{m^T,h}(\rho(r), b)$ for all $m^T \in M$,

(ii) $\rho(r) = \arg\max_a U_{\sigma^{-1}(r),h}(a)$ for all $r \in \sigma(M)$.

---

[6]While the restriction to an h-path is not a proper subgame as it cuts information sets, equilibria $(\overline{\sigma}, \overline{\rho})$ of the whole game and the considered "h-dependent equilibria" (σ_h, ρ_h) stand in a one-to-one correspondence: Fixing any (σ_h, ρ_h) for each h, we can define $\overline{\sigma}(m, h) := \sigma_h(m)$ and $\overline{\rho}(h, r) := \rho_h(r)$ to obtain an equilibrium. Vice versa, given $(\overline{\sigma}, \overline{\rho})$, we can define $\sigma_h(m) := \overline{\sigma}(m, h)$ and $\rho_h(r) := \overline{\rho}(m, h)$ to obtain an h-dependent equilibrium.

[7]Following footnote 2 on p. 7, we can use the *Dempster-Shafer updating rule* (Dempster, 1968; Shafer, 1976) or the full Bayesian updating rule (Fagin and Halpern, 1991) (if $0/0 := 1$) to obtain the updated capacity μ_1^r. The minimal feasible $\tilde{M}$ set can be obtained from μ_1^r since $\mu_1^r(M') = 1$ if $\tilde{M} \subseteq M'$ and $\mu_1^r(M') = 0$ otherwise.

[8]Note that the restriction to pure strategies is without loss of generality for the receiver, since the utility-maximizing action is unique.

A sender who observes the true model m^T sends a message r inducing the action most favorable for her out of $\rho(\mathcal{R})$. Being aware of the sender's communication strategy σ and observing a message r, the receiver first updates the considered set of models to the minimal feasible set $\tilde{M} = \sigma^{-1}(r)$. Then, he chooses the action maximizing his utility w.r.t. his ambiguity rule.

An immediate consequence of the cheap-talk setting is the existence of trivial, so-called *babbling equilibria* in which no information is transmitted: For instance, let the sender's strategy be constant, i.e., $\sigma(m) = r_0$ for all m. Then the minimal feasible set upon receiving r_0 is the set M of all models. The receiver thus responds optimally by playing the *pooling action* $\rho(r) = a(M, h)$ for any received message. Given that the receiver only plays a single action, the sender cannot change his behavior and thus improve her utility by transmitting a message different than r_0. Consequently, (σ, ρ) is an equilibrium inducing solely the pooling action $a(M, h)$.

2.4 Running example

We conclude this section by introducing our running example. Consider a politician (sender) who runs for office and wants to convince a representative voter (receiver) of her aptitude for office. Her ability is summarized by a parameter $\theta \in [0, 1]$ which is thought of as capturing her skill level in managing economic indicators, such as the unemployment rate or inflation, or exigencies of a pandemic. Without further information, the ability is believed to be drawn from a uniform distribution $F_0 = \mathcal{U}(0, 1)$. A record of past values of economic indicators, reproduction numbers, or death tolls is publicly available. The data is simplified and summarized by a vector $h = (h_1, h_2, \ldots, h_K) \in \{0, 1\}^K$, indicating public outcomes that are categorized as positive (1, success) or negative (0, failure).

In this example, a model is identified with a set of data points attributed to the politician's ability, i.e., $M = 2^{\{1,\ldots,K\}}$. The true model $m^T \in M$ indicates the instances in which the politician's ability contributed to the outcome. At all other times, the result is independent of her skill and equally likely to be a success or a failure, so that $h_1, \ldots, h_K$ are independent conditional on θ, with

$$\Pr(h_k = 1 \mid \theta) = \begin{cases} \theta, & \text{if } k \in m^T \\ \frac{1}{2}, & \text{else} \end{cases} \quad \text{for all } k = 1, 2, \ldots, K. \tag{1}$$

Note that (1) constitutes a modified beta-binomial model. The unbiased representative

voter is interested in learning the true ability of the politician in question, but faces uncertainty about which pieces of the data, e.g., a low unemployment rate, are relevant for this assessment. The biased politician, on the other hand, knows which of the observed data is truly connected to her own ability, i.e., the true model m^T. Following the leading example in Crawford and Sobel (1982), her payoff function is $u_S(a, \theta, b) = -(\theta + b - a)^2$, reflecting that she likes to raise the voter's opinion on her aptitude as a good politician. She does so by providing a *narrative* that rationalizes the observed data while also presenting herself in a good light. In our framework, the voter takes the politician's strategic incentives into account, following equilibrium behavior, and can thus counteract the bias.

Before proceeding with the equilibrium analysis, we provide some intuition for equilibria in this setting. Figure 1 depicts an equilibrium under the MLEU ambiguity rule (henceforth *MLEU equilibrium*) for $b = \frac{1}{25}$, $M = 2^{\{1,2,3\}}$, i.e., $K = 3$, and the history $h = (1, 0, 1)$. We associate each model $m \in M$ with the receiver's bliss point $a(m, h, 0) = \mathbb{E}[\theta \mid m, h]$ (cf. Example 1). As the communication is about models, the sender partitions the set of models into intervals, sending the same message for all models in a given partition element. Under the MLEU ambiguity rule, the receiver picks the model with the highest expected fit from each partition element. Given this action profile and the sender's bias, the sender has no incentives to deviate by sending a different message. For instance, if $m^T = \{1, 2\}$ such that the sender's bliss point is $a(m^T, h, b) = a(m^T, h, 0) + b = \frac{27}{50}$ (cf. Example 1), then the sender is better off sending $\sigma(m^T) = r_2$ than r_3, as

$$\left| a(m^T, h, b) - \rho(r_2) \right| = b < \tfrac{3}{50} = \left| a(m^T, h, b) - \rho(r_3) \right|.$$

3   Equilibrium analysis

In this section, we characterize equilibria in the general framework and derive informativeness thresholds on the conflict of interest. In a first step, we state necessary conditions for establishing equilibria. In equilibrium, messages span intervals of models, i.e., messages correspond to sets of narratives the respective bliss points of which are consecutive elements w.r.t. the set of all bliss points. This helps reducing the complexity of finding equilibria. Second, we prove that a property well-known from classic cheap-talk games
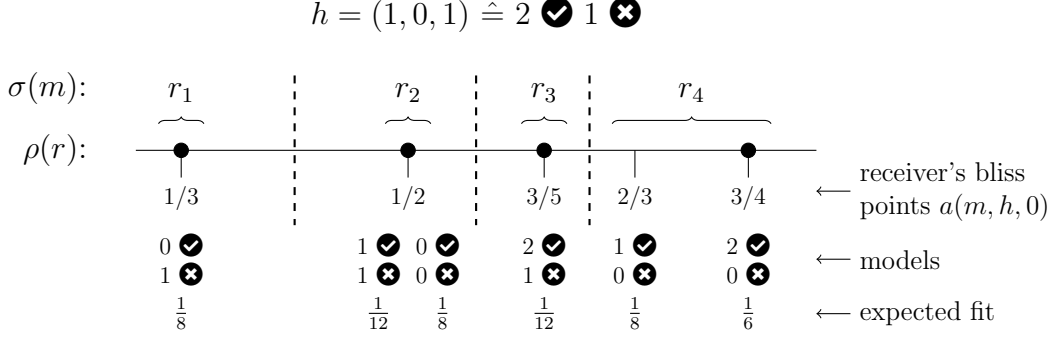
12

$$h = (1, 0, 1) \;\hat{=}\; 2 \;✅\; 1 \;❌$$

Figure 1: MLEU equilibrium in the setting of Section 2.4, where $b = \frac{1}{25}$, $M = 2^{\{1,2,3\}}$, and $h = (1, 0, 1)$ entails two successes (✅) and one failure (❌). Reports span consecutive bliss points, each of which corresponds to at least one model. Each model corresponds to a claimed number of relevant successes and failures and has an associated expected fit. The optimal receiver responses for each report are indicated by the solid dots.

surprisingly carries over to a finite setting: If an equilibrium induces $n+1$ different actions, there also is an equilibrium which induces n different actions. Third, we derive thresholds on the conflict of interest that mark when the most finely granulated information revelation is full disclosure or partial disclosure, respectively.

3.1 Necessary conditions

Our first straightforward observation establishes an upper bound for the number of actions induced in equilibrium. Observe that the sender can at most send $\#M$ different messages, which yields:

Lemma 1. *In any equilibrium, the number of distinct actions induced is finite and at most equal to $\#\{a(m, h, 0) \mid m \in M\}$.*

All proofs are relegated to the appendix. The next lemma describes important structural properties of the equilibria that help simplify the search for equilibria.

Lemma 2 (Reduction Lemma)**.** *Let (σ', ρ') be an equilibrium.*

(i) There is a corresponding reduced equilibrium *(σ, ρ) with $\rho \circ \sigma \equiv \rho' \circ \sigma'$ and whenever $\rho(\sigma(m)) = \rho(\sigma(m'))$ then $\sigma(m) = \sigma(m')$.*

(ii) We have $a(m', b) \leq a(m, b)$ if and only if $\rho'(\sigma'(m')) \leq \rho'(\sigma'(m))$.

Statement (i) lets us restrict attention to reduced equilibria in which messages and induced actions are in a one-to-one correspondence:[9] If an equilibrium has two different

---

[9]In that sense, messages can be identified with action recommendations, a property sometimes referred to as *straightforwardness*.

messages that induce the same action, one does not lose the equilibrium property by sending the same message for all corresponding narratives. Statement (ii) says that bliss points and induced actions of narratives must be ordered in the same way in equilibrium. This property is due to interval choice. Consequently, a message partitions narratives into convex sets w.r.t. the ordering of the bliss points.

3.2 Equilibria

In the following, we study the existence of equilibria with a certain number of induced actions and characterize the set of equilibria. We categorize equilibria according to the number of induced actions.

Definition 2 (*n*-step equilibrium)**.** An equilibrium (σ, ρ) is called *n-step equilibrium* if it induces n distinct actions, i.e., $\#(\rho \circ \sigma)(M) = n$.

Note first that by Lemma 1 any equilibrium induces at most $\#\{a(m, h, 0) \mid m \in M\}$ distinct actions. Second, a 1-step equilibrium is a trivial babbling equilibrium, which always exists. Third, by the Reduction Lemma 2 the reduced equilibrium corresponding to an n-step equilibrium has n distinct messages. We henceforth characterize equilibria by the corresponding reduced equilibria.

A classical result from the theory of cheap-talk games states that if there is an equilibrium which induces N distinct actions, then there is an equilibrium inducing n distinct actions for all $1 \leq n \leq N$, cf. Crawford and Sobel (1982). In their setting, messages correspond to intervals which partition the continuous state space. Surprisingly, this result also holds true in our more complex setting despite that messages correspond to intervals which partition the *finite* model space. Together with the Reduction Lemma 2, this characterizes the set of equilibria: An equilibrium is a partition of the set M into up to N intervals together with their actions induced by the considered ambiguity rule. We say that a partition of M is consecutive if it is consecutive w.r.t. the corresponding bliss points.

Theorem 1. *There is a natural number $N = N(h, b)$ such that there exists an n-step equilibrium if and only if $1 \leq n \leq N$. For any n-step equilibrium, the corresponding reduced equilibrium (σ, ρ) is characterized by a consecutive partition $\{M_1, M_2, \ldots, M_n\}$ of M and distinct $r_1, r_2, \ldots, r_n \in \mathcal{R}$ such that for all $i = 1, 2, \ldots, n$*

(i) $r_i = \sigma(m)$ if $m \in M_i$,

14

(ii) ρ *satisfies* $\rho(r_i) = \arg\max_a U_{M_i, h}(a)$, *and*

(iii) $r_i \in \arg\max_r U_{m^T, h}(\rho(r), b)$ *for all* $m^T \in M_i$.

By the Reduction Lemma 2, the set of narratives inducing an action is an interval in equilibrium. Furthermore, the actions induced in a reduced n-step equilibrium (Theorem 1 (ii)) are pairwise distinct. We can hence index these intervals M_i such that narratives from an interval with a higher index induce a higher action, i.e., $\rho(\sigma(m)) < \rho(\sigma(m'))$ whenever $m \in M_i$ and $m' \in M_j$ for $i < j$. Thus, reduced equilibria are characterized by a consecutive partition of M. Note that it is straightforward to reconstruct non-reduced equilibria from a reduced one:

Remark 1. Consider a reduced equilibrium (σ, ρ) characterized by the consecutive partition $\{M_1, M_2, \ldots, M_n\}$. Any n-step equilibrium (σ', ρ') of which (σ, ρ) is the corresponding reduced equilibrium is characterized by a refined consecutive partition $\{M'_1, M'_2, \ldots, M'_{n'}\}$ of $\{M_1, M_2, \ldots, M_n\}$ and distinct $r'_1, r'_2, \ldots, r'_{n'} \in \mathcal{R}$, $n' \geq n$, such that $r'_i = \sigma'(m)$ if $m \in M'_i$ for all $i = 1, 2, \ldots, n'$ and $\rho \circ \sigma \equiv \rho' \circ \sigma'$. If $n' = n$, then $M_i = M'_i$ for all $i = 1, \ldots, n$.

The proof of the theorem then gives at hand an explicit algorithm to arrive at an n-step equilibrium given an $(n+1)$-step equilibrium that we explain in the following.

Algorithm. Let (σ, ρ) be an $(n+1)$-step equilibrium characterized by the consecutive partition $\{M_1, M_2, \ldots, M_{n+1}\}$.

1. Define the sender strategy

$$
\sigma_a(m) := \begin{cases} r_n & , \text{ if } m \in M'_n \\ r_i & , \text{ if } m \in M'_i, i \neq n, n+1 \end{cases},
$$

where $M'_i = M_i$ for $i = 1, 2, \ldots, n-1$ and $M'_n = M_n \cup M_{n+1}$.

2. If σ_a together with the respective best reply ρ_a of the receiver is an n-step equilibrium, we are done. Otherwise go to Step 3.

3. Find i such that the sender has a profitable deviation for model $m = m^l_{M'_i}$ from σ_a,

where $m^l_{M'_i}$ denotes the narrative in M'_i with the lowest induced receiver action.[10]

Let

- σ'_a be such that $\sigma'_a(m) = r_{i-1}$ if $m = m^l_{M'_i}$ and $\sigma'_a(m) = \sigma_a(m)$ else,

- $M''_{i-1} = M'_{i-1} \cup \{m^l_{M'_i}\}$, $M''_i = M'_i \setminus \{m^l_{M'_i}\}$, and $M''_j = M'_j$ for $\neq i-1, i$.

Set $\sigma_a = \sigma'_a$ and $M'_i = M''_i$ and go back to Step 2. $\qquad\qquad\square$

In the first step of the algorithm, we merge the rightmost messages to obtain σ_a which uses n messages. If σ_a together with its best reply ρ_a does not form an equilibrium, we iteratively modify σ_a: There is a profitable deviation for the sender such that there is an interval M'_{i-1} that the sender can profitably enlarge by adding the leftmost model $m^l_{M'_i}$ in the subsequent interval M'_i (in the first step, this is because we have merged the rightmost messages, i.e., $i = n$). Choosing any such deviation, check whether the deviation together with the receiver's best reply constitutes an equilibrium. If not, continue as before. If this procedure continues to produce no equilibrium, σ'_a will eventually be equal to σ_b – the sender strategy resulting from merging the leftmost messages of σ. If the algorithm reaches that point, no further profitable deviation is possible (since by symmetry this would entail enlarging an interval by adding the rightmost model of the preceding interval). The procedure thus terminates at an σ'_a that forms an equilibrium together with ρ'_a. The key idea is summarized in Figure 2.

The algorithm also yields an intuitive comparative statics result on the conflict of interest. The smaller the conflict of interest, the more the sender is willing to discriminate between different models in terms of the step size in equilibrium.

Corollary 1. $N(h, b) \leq N(h, b')$ if $b' < b$.

3.3 Informativeness

As is common in cheap-talk games, there is a plethora of different equilibria. In this section, we introduce a partial order on sender strategies and thus equilibria, based on the (Blackwell-)informativeness of the sender's strategy, cf. Blackwell (1953). Informativeness measures how finely the sender discriminates between different models.

---

[10]For the sake of exposition, we ignore that $m^l_{M'_i}$ may not be uniquely defined when there are distinct $m, m' \in M$ with $a(m, h, 0) = a(m', h, 0)$. As we show in the proof of the theorem, we can essentially treat the set of such models as a single model.
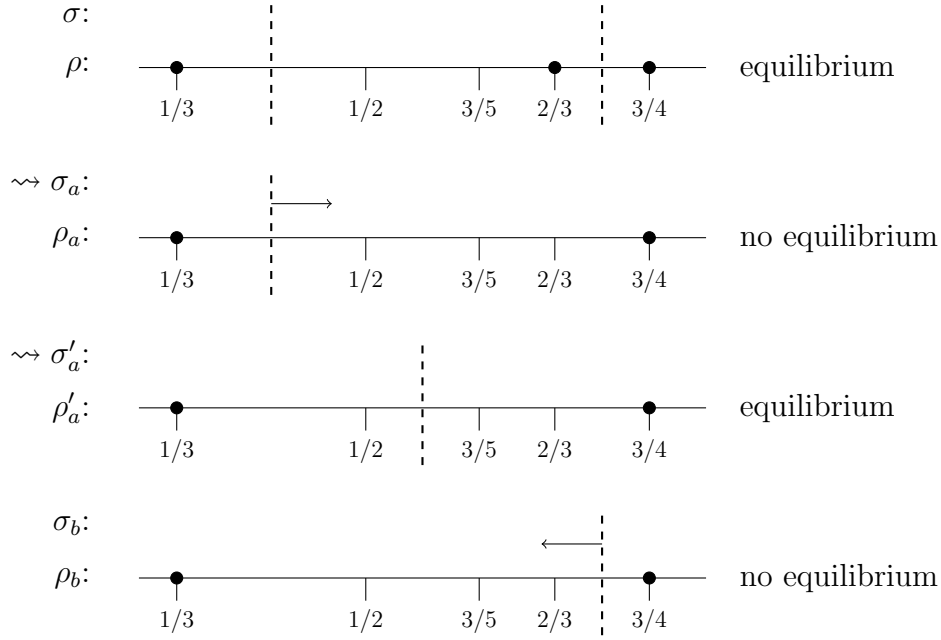
Figure 2: Sketch of the algorithm driving the proof of Theorem 1. Consider $b = \frac{1}{30}$, $M = 2^{\{1,2,3\}}$, $h = (1, 0, 1)$, and the MLEU ambiguity rule with any tiebreaker that favors the narrative with no success and one failure over the one with one success and no failure, and the latter over the one with no success and no failure. The dots mark the respective best responses of the receiver. Starting from the 3-step equilibrium σ, merge the two rightmost words to obtain σ_a. A profitable deviations must involve shifting the boundary to the right. By doing so, equilibrium (σ_a', ρ_a') is reached.

Definition 3 (Informativeness). (i) Strategy σ is *weakly more informative* than σ' (under h) if $\sigma(m) = \sigma(m')$ implies $\sigma'(m) = \sigma'(m')$ for all $m, m' \in M$.

(ii) Strategy σ is *fully informative (uninformative)* (under h) if σ (σ') is weakly more informative than σ' (σ) for any strategy σ'.

(iii) An equilibrium (σ, ρ) is *weakly more informative* than (σ', ρ') (under h) if σ is weakly more informative than σ'.

(iv) An equilibrium (σ, ρ) is *most informative* (under h) if for any equilibrium (σ', ρ') that is weakly more informative than (σ, ρ), it holds that (σ, ρ) is weakly more informative than (σ', ρ').

(v) An equilibrium (σ, ρ) is *fully informative* (under h) if σ is fully informative.

Some remarks seem in order. First, Definition 3 defines a partial order on strategy profiles and equilibria. In particular, note that a more informative equilibrium has a larger step size (cf. Theorem 1, Corollary 1). Second, the receiver can always infer the true model

$$h = (1,0,1) \mathrel{\hat{=}} 2 \; \checkmark \; 1 \; \times$$

$\sigma_1:$ r_1 r_2 r_3 r_4

$\rho_1:$

1/3   1/2   3/5   2/3   3/4

0 ✓   1 ✓   0 ✓   2 ✓   1 ✓   2 ✓

1 ✗   1 ✗   0 ✗   1 ✗   0 ✗   0 ✗

$\sigma_2:$ r'_1 r'_2 r'_3 r'_4

$\rho_2:$

1/3   1/2   3/5   2/3   3/4

0 ✓   1 ✓   0 ✓   2 ✓   1 ✓   2 ✓
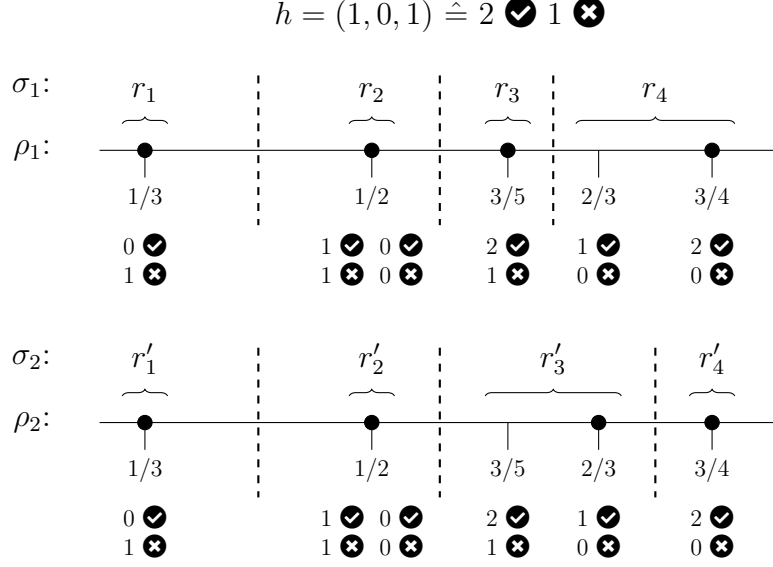
1 ✗   1 ✗   0 ✗   1 ✗   0 ✗   0 ✗

Figure 3: Two most informative MLEU equilibria that induce a different set of actions in the setting of Section 2.4 with $b = \frac{1}{25}$, $M = 2^{\{1,2,3\}}$, and $h = (1,0,1)$.

m^T if σ is fully informative. Formally, the corresponding reduced equilibrium is thus not fully informative if multiple models are associated with the same optimal action, i.e., $a(m,h,0) = a(m',h,0)$ for distinct $m, m' \in M$. Nevertheless, by Lemma 1:

Remark 2. A fully informative n-step equilibrium and its corresponding reduced equilibrium are such that $n = N(h,b) = \#\{a(m,h,0) \mid m \in M\}$.

Third, a fully informative equilibrium is most informative. Fourth, there might be several most informative equilibria which have different model-action profiles, i.e., equilibria (σ, ρ) and (σ', ρ') such that $\rho(\sigma(m)) \neq \rho'(\sigma'(m))$ for some $m \in M$. The last point is illustrated in Figure 3 by two 4-step equilibria.

In the following, we characterize thresholds on the conflict of interest that mark when fully or partially informative communication is possible in equilibrium. By the Reduction Lemma 2, a 2-step equilibrium consists of a partition of the set M into two intervals. Recall that for a subset $\tilde{M} \subseteq M$ the unique optimal receiver action given his ambiguity rule is denoted by $a(\tilde{M}, h)$. Defining

$$V(m^T, h, b) := U_{m^T, h}(a(\{m \mid m > m^T\}, h), b) - U_{m^T, h}(a(\{m \mid m \leq m^T\}, h), b),$$

we see that there is no 2-step equilibrium if $V(m^T, h, b) > 0$ for all m^T which are not maximal.[11] As a consequence of Theorem 1, the only equilibrium in this case is the

[11]Technically, we would need to allow different models m, m' with $a(m,h,0) = a(m',h,0)$ to end up

18

babbling equilibrium.

Definition 4 (Large conflicts of interest). We say that $U_{m,h}(a,b)$ *allows for large conflicts of interest* if for any h there exists a $\hat{b}(h) \geq 0$ such that $V(m^T, h, b) > 0$ for all $b > \hat{b}(h)$ and all non-maximal $m^T \in M$.

Under large conflicts of interest, the sender cannot induce two different actions without having an incentive to deviate: There will always be a true model under which she seeks to make the receiver pick a higher action than is designed. To illustrate this point, imagine that the bias, measuring the conflict of interest, is high enough to want her make the receiver always pick the highest possible action. Then there cannot be a 2-step equilibrium. Note that under a quadratic loss the bliss points of the sender are translations of $a(m^T, h, 0)$ by exactly b, thus allowing for large conflicts of interest.

Analogously to standard cheap-talk models in which the sender is imperfectly informed (e.g., Argenziano et al., 2016; Foerster, 2023), we obtain the following *informativeness thresholds* on the conflict of interest:

Proposition 1. *For any history h, there exists*

(i) $\underline{b}(h) > 0$, *such that a most informative equilibrium under history h involves a fully informative sender strategy if and only if $b \leq \underline{b}(h)$, and*

(ii) $\overline{b}(h) \geq \underline{b}(h)$, *such that any most informative equilibrium under history h is uninformative if and only if $b > \overline{b}(h)$, provided $U_{m,h}(a,b)$ allows for large conflicts of interest.*

To illustrate this result, we revisit the running example introduced in Section 2.4. In particular, it may indeed be the case that the two informativeness thresholds coincide under the MLEU ambiguity rule:

Example 3. Consider the setting of Section 2.4, where $F_0 = \mathcal{U}(0,1)$, $u_S(a,\theta,b) = -(\theta + b - a)^2$, and h is generated according to (1), and the MLEU ambiguity rule.

(i) If $h = (1,0,1)$, then $\underline{b}(h) = \frac{1}{24} < \frac{5}{24} = \overline{b}(h)$.

(ii) If $h = (0,0,0)$, then $\underline{b}(h) = \overline{b}(h) = \frac{1}{40}$ if the receiver breaks the tie in expected fit between the narrative indicating one failure and the one indicating no information

---

in either of the partition elements, which could be achieved by considering all strict refinements of $\leq$ on M. This formality is not restrictive in the following analysis.

in favor of the first one.[12] Since models that induce a smaller posterior belief are associated with a higher expected fit in this case, less than fully informative communication does not relax the incentive-compatibility constraints. As with a fully informative strategy, the sender has incentives to deviate from the partially informative strategy associated with the weakest incentive-compatibility constraints if $b > \frac{1}{40}$, see Figure 4 for an illustration.
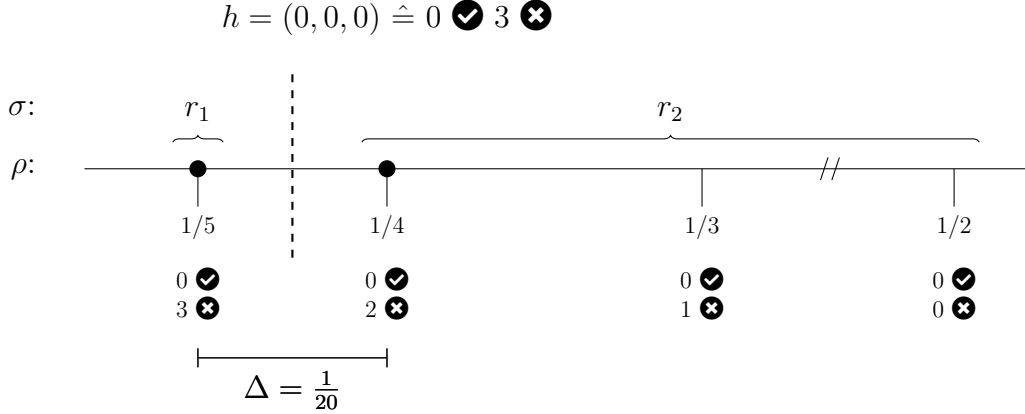

Figure 4: Partially informative strategy σ, where $M = 2^{\{1,2,3\}}$ and $h = (0, 0, 0)$ entails three failures (✖) in Example 3 (ii). The optimal receiver response is indicated by the solid dots.

The following proposition establishes that the sender always prefers to play a more informative equilibrium – and thus a most informative one – if facing a receiver who applies the MLEU ambiguity rule.

Proposition 2. *Let (σ, ρ) and (σ', ρ') be MLEU equilibria and σ weakly more informative than σ'. Then, the sender prefers the outcome of (σ, ρ) over (σ', ρ') for any m^T.*

A more informative equilibrium discriminates more finely between different models. Under the MLEU rule, splitting up an interval into subintervals means the sender can still induce the same actions as before (and an additional one), as the respective models will continue to maximize the expected fit across some interval. The sender thus could achieve the same outcome as before splitting up the interval, but has no incentives to do so by definition of equilibrium. A similar statement is not true for the MEU rule. As the example in Figure 5 shows, splitting up an interval into subintervals does not necessarily yield a superset of actions which can be induced under the MEU rule, and may

---

[12]If, instead of $M = 2^{\{1,2,3\}}$, we restrict the model set to $M = 2^{\{1,2,3\}} \setminus \{\emptyset\}$, the thresholds coincide independently of the employed tiebreaking rule.
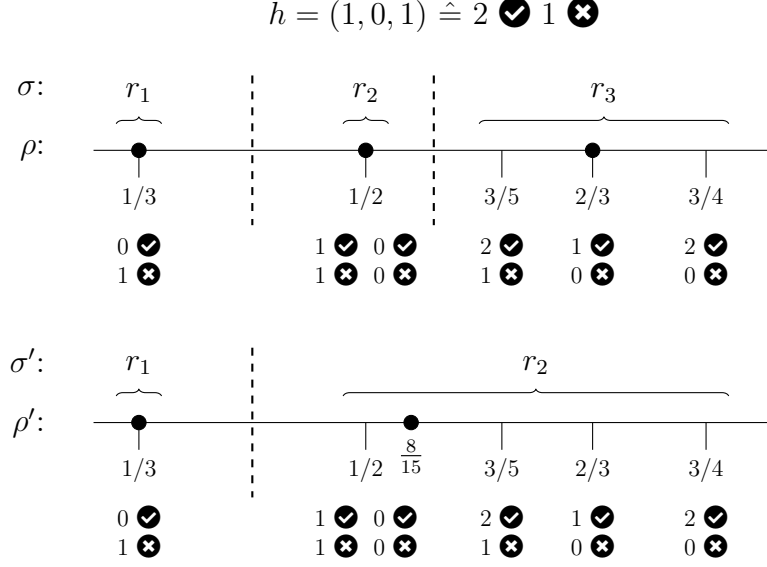
Figure 5: MEU equilibria in the setting of Section 2.4 for $b = 0.04$ and $h = (1, 0, 1)$. While σ is more informative than σ', the sender prefers the equilibrium action $\rho'(\sigma'(m^T)) = \frac{8}{15} \approx 0.53$ under σ' over $\rho(\sigma(m^T)) = \frac{1}{2}$ under σ for $m^T = \{1, 2\}$, since her bliss point is $a(m^T, h, b) = 0.54$.

thus be worse for the sender. The counterexample marks a stark contrast to Theorem 2 in Colo (2024), where the most informative equilibrium is interim dominant under MEU preferences. In further contrast to Colo's Theorem 1, not all 'cut-off' models defining the interval partition lie below the action under full uncertainty, as the rightmost interval of the MEU equilibrium (σ, ρ) in Figure 5 induces an action strictly greater than $a(M, h) = \frac{13}{25}$.

4 Analysis of the running example

In this section, we provide some more insights into the running example, characterizing the informativeness thresholds on the conflict of interest (cf. Proposition 1). Recall from Section 2.4 that the running example features a politician (sender) who wants to convince a representative voter (receiver) of her aptitude for office. The ability parameter θ is drawn from $F_0 = \mathcal{U}(0, 1)$ and the observable data is given by a sequence $h = (h_1, h_2, \ldots, h_K) \in \{0, 1\}^K$ of public outcomes categorized as positive (1, success) or negative (0, failure). The true model $m^T \in M = 2^{\{1,2,\ldots,K\}}$ indicates the instances in which the politician's ability contributed to the outcome, i.e., $h_1, \ldots, h_K$ are independent

conditional on θ, with

$$\Pr(h_k = 1 \mid \theta) = \begin{cases} \theta, & \text{if } k \in m^T \\ \frac{1}{2}, & \text{else} \end{cases} \quad \text{for all } k = 1, 2, \ldots, K.$$

Furthermore, the sender's payoff function is $u_S(a, \theta, b) = -(\theta + b - a)^2$.

Observe first that in our concrete setting $h^\Sigma := \sum_{k=1}^K h_k$ is a sufficient statistic for history h, counting the bare number of successes and failures. Second, the Reduction Lemma 2 reveals that in equilibrium the action induced by (respectively the message sent under) the narrative m must be induced by all models m' with $\#m' = \#m$ and $\sum_{i \in m} h_i = \sum_{i \in m'} h_i$. Consequently, they can be identified. Third, the expected value given m and h is the bliss point of the receiver,

$$a(m, h, 0) = \mathbb{E}[\theta \mid m, h] = \frac{\sum_{i \in m} h_i + 1}{\#m + 2}.$$

The biased sender chooses her strategy in order to induce a favorable posterior belief (regarding her own ability) of the receiver. We know from Proposition 1 that she is willing to reveal the true model to him if interests are sufficiently aligned, $b \leq \underline{b}(h^\Sigma) := \underline{b}(h)$, while she cannot credibly reveal information in a way that favors her if her bias is too large, $b > \overline{b}(h^\Sigma) := \overline{b}(h)$. We first establish that the former threshold concerning whether fully informative communication is feasible, $\underline{b}(h^\Sigma)$, is independent of the ambiguity rule, symmetric, and that extreme histories are the most difficult to 'explain', in the sense that $\underline{b}(h^\Sigma)$ attains its minimum at these histories.

Proposition 3. $\underline{b}(h^\Sigma)$ *is independent of the ambiguity rule, symmetric, i.e.,* $\underline{b}(h^\Sigma) = \underline{b}(K - h^\Sigma)$ *for any* h^Σ, *and* $\{0, K\} = \arg\min_{h^\Sigma = 0, 1, \ldots, K} \underline{b}(h^\Sigma)$.

Figure 6 illustrates our findings. Somewhat surprisingly:

Remark 3. $\underline{b}(h^\Sigma)$ *generally is not quasi-concave.*

Second, we consider the threshold $\overline{b}(h^\Sigma)$ concerning whether partially informative communication is feasible. We show that we no longer have symmetry in this case due to the receiver's residual uncertainty.

Proposition 4. *Under the MLEU and MEU ambiguity rule,* $\overline{b}(h^\Sigma)$ *is asymmetric if* $K \geq 2$. *In the case of*
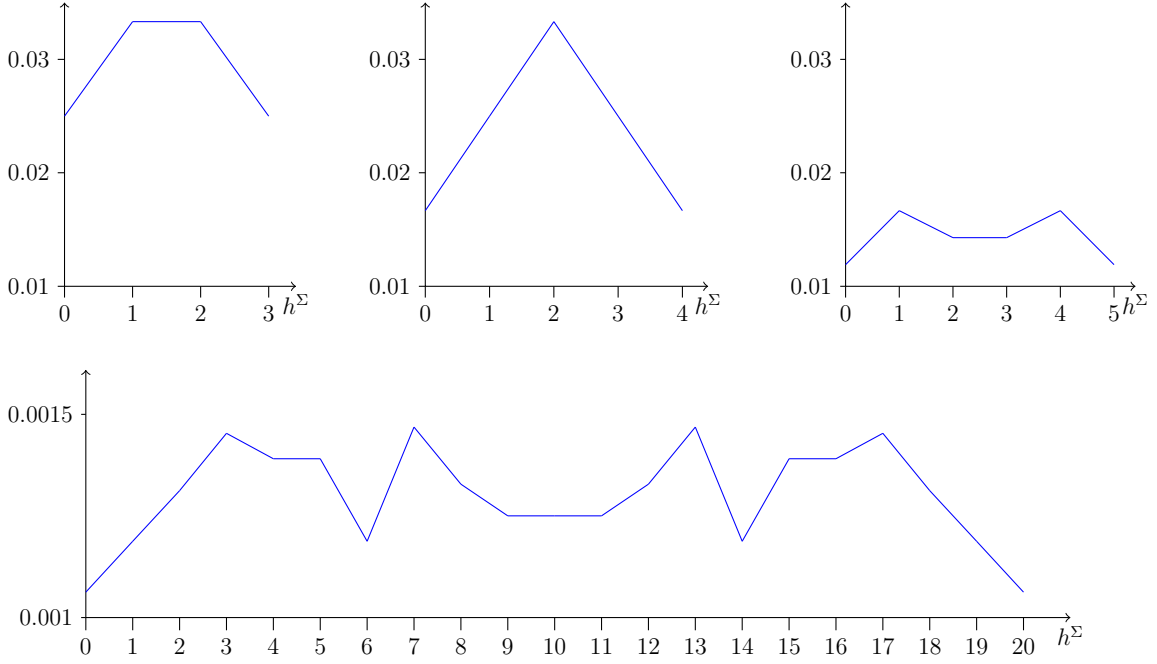
Figure 6: $\underline{b}(h^\Sigma)$ as a function of h^Σ for $K = 3, 4, 5$ (top left to right) and $K = 20$ (bottom).

(i) *MLEU, we have* $\bar{b}(K) = \frac{K}{4(K+2)} > \bar{b}(0) = \underline{b}(0) = \frac{1}{2(K+1)(K+2)}$,

(ii) *MEU, we have* $\bar{b}(K) = \max\{\frac{1}{12}, \frac{K}{8(K+3)}\} < \bar{b}(0) = \frac{K+1}{6(K+2)}$.

To understand the intuition, recall from Section 3.3 that we can characterize a 2-step equilibrium by a threshold model m^* such that $V(m^*, h, b) \leq 0$, i.e., the sender has no incentives to deviate if $m^T = m^*$. The proof establishes that for the MLEU and MEU ambiguity rule, the 2-step partition which determines the threshold $\bar{b}(h^\Sigma)$ is such that the first partition element is a singleton solely containing the minimal model with respect to the order of receiver bliss points, denoted $m^*(h^\Sigma)$. Thus, $\bar{b}(h^\Sigma)$ solves $V(m^*(h^\Sigma), h, b) = 0$ and depends on the induced actions, in particular the receiver's action $a(\{m \mid m > m^*(h^\Sigma)\}, h)$ upon a deviation at $m^T = m^*(h^\Sigma)$. Note that $a(\{m \mid m > m^*(h^\Sigma)\}, h)$ depends on the ambiguity rule and the concrete history h^Σ through the order of models it induces. In particular, $m^*(0) \neq m^*(K)$ and thus $a(\{m \mid m > m^*(0)\} \neq a(\{m \mid m > m^*(K)\}$. The higher the action chosen by the receiver, the more loose the sender's incentive-compatibility constraints and, thus, the higher $\bar{b}(h^\Sigma)$. Under the MLEU ambiguity rule, this action depends on the expected fit:

Example 4 (MLEU ambiguity rule). Suppose that $K = 3$, $b \in (\frac{1}{15}, \frac{3}{20}]$, and that the receiver behaves according to the MLEU ambiguity rule. As can be seen from Figure 7a, informative communication is feasible in this parameter range for $h^\Sigma = 3$ but not for $h^\Sigma = 0$:

(i) $h^\Sigma = 3$. The most informative equilibrium is partially informative, with $r = r_1$ if $m^T = \emptyset$, i.e., if no data point is relevant, and $r = r_2$ else. Note that a model's expected fit increases in the posterior it induces. After receiving report $r = r_2$, models that induce posteriors $\frac{2}{3}$, $\frac{3}{4}$, and $\frac{4}{5}$ are feasible, such that the receiver takes action $\frac{4}{5}$. Since $b \leq \frac{3}{20}$, the sender prefers action $a(m^T, h, 0) = \frac{1}{2}$ over $\frac{4}{5}$ if $m^T = \emptyset$.

(ii) $h^\Sigma = 0$. The most informative equilibrium is uninformative. In this case, a model's expected fit decreases in the posterior it induces, such that any partially informative communication strategy is associated with weakly tighter incentive-compatibility constraints than a fully informative strategy.

Under the MEU ambiguity rule, the action $a(\{m \mid m > m^*(h^\Sigma)\}, h)$ upon a deviation at $m^T = m^*(h^\Sigma)$ depends on the variance associated with models $m^T > m^*(h^\Sigma)$:

Example 5 (MEU ambiguity rule). Suppose that $K = 3$, $b \in (\frac{1}{8}, \frac{2}{15}]$, and that the receiver behaves according to the MEU ambiguity rule. As can be seen from Figure 7b, informative communication is feasible in this parameter range for $h^\Sigma = 0$ but not for $h^\Sigma = 3$:

(i) $h^\Sigma = 0$. The most informative equilibrium is partially informative, with $r = r_1$ if $m^T = \{1, 2, 3\}$, i.e., if all data points are relevant, and $r = r_2$ else. After receiving report $r = r_2$, models that induce posteriors $\frac{1}{4}$, $\frac{1}{3}$, and $\frac{1}{2}$ are feasible. Note that a higher posterior is associated with a higher variance, such that the receiver's hedging yields action $\frac{7}{15}$, close to $\frac{1}{2}$. Since $b \leq \frac{2}{15}$, the sender prefers action $a(m^T, h, 0) = \frac{1}{5}$ over $\frac{7}{15}$ if $m^T = \{1, 2, 3\}$.

(ii) $h^\Sigma = 3$. The most informative equilibrium is uninformative. Note that the partially informative strategy associated with the weakest incentive-compatibility constraints is such that $r = r_1$ if $m^T = \emptyset$, i.e., if no data point is relevant, and $r = r_2$ else. After receiving report $r = r_2$, models that induce posteriors $\frac{2}{3}$, $\frac{3}{4}$, and $\frac{4}{5}$ are feasible. Note that a higher posterior is associated with a lower variance, such that the receiver's hedging yields action $\frac{2}{3}$ equal to the lowest posterior induced by some feasible model. As a result, informative communication is precluded: Since $b > \frac{1}{8}$, the sender prefers action $\frac{2}{3}$ over $a(m^T, h, 0) = \frac{1}{2}$ if $m^T = \emptyset$.
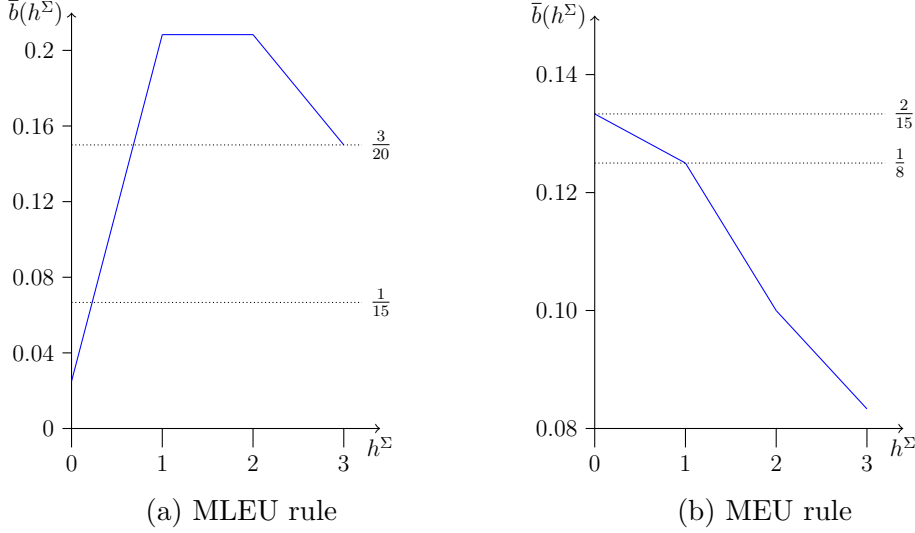
(a) MLEU rule   (b) MEU rule

Figure 7: $\bar{b}(h^{\Sigma})$ as a function of h^{Σ} for $K = 3$ for the MLEU and the MEU ambiguity rule. The dotted lines mark the thresholds on the bias in Example 4 and 5.

5   Comparison with a naïve receiver

In this section, we compare the MLEU equilibria in our framework with S&S, where the receiver is naïve in the sense that he does not take the sender's strategic incentives into account. Instead, the receiver takes the sender's message $r = m$ at face value and adopts the proposed model m if it has a higher expected fit than a default model m^d given the observed history h, i.e., $\Pr(h|m, F_0) \geq \Pr(h|m^d, F_0)$. Otherwise he sticks to the default model. Formally, let $\mathcal{R} = M$. The *sender-optimal profile* $(\sigma^{\dagger}, \rho^{\dagger})$ then satisfies

(i) $\sigma^{\dagger}(m^T) \in \arg\max_r U_{m^T, h}(\rho^{\dagger}(r), b)$ for all $m^T \in M$,

(ii) $\rho^{\dagger}(r) = \arg\max_a U_{\tilde{m}, h}(a)$, where $\tilde{m} \in \arg\max_{m = r, m^d} \Pr(h \mid m, F_0)$, for all $r \in \mathcal{R}$.[13]

First, note that indeed the receiver does not take into account the sender's strategy. Second, the restriction of the comparison to MLEU equilibria appears natural in light of (ii).

We henceforth refer to the model of S&S as the *naïve model* and adopt their assumption (from Section III) that the receiver's default model is the true model, $m^d = m^T$. This choice is reasonable for at least two reasons: First, as S&S point out, it allows to ask when the wrong story wins. Second, it can be seen as a 'best case' scenario, under

---

[13]Like S&S, we break ties in expected fit between the proposed model r and the default m^d in favor of the proposed model, i.e., $\tilde{m} = r$ if $\Pr(h \mid r, F_0) = \Pr(h \mid m^d, F_0)$.

which the truthful communication benchmark has the same model-action profile as the truthful communication benchmark in our framework.[14]

A weak measure of the persuasive power of the sender is the set of models m^T for which she prefers to induce an action $a(m^*, h, 0)$, associated to another narrative $m^* \neq m^T$, over $a(m^T, h, 0)$ and succeeds to do so under a profile (σ, ρ) given the history h:

$$M_{(\sigma,\rho)} := \left\{ m^T \in M \mid U_{m^T,h}(\rho(\sigma(m^T)), b) > U_{m^T,h}(a(m^T, h, 0), b) \right\}.$$

Take the naïve model, then $M_{(\sigma^\dagger,\rho^\dagger)}$ contains all models m^T under which the sender wants to and successfully can convince a naïve receiver to move away from the action he would have chosen by default. She does so although the default model corresponds to the true model m^T, i.e., the wrong story wins. Likewise, for any MLEU equilibrium (σ, ρ), $M_{(\sigma,\rho)}$ contains all models m^T under which the sender is able to make the receiver take an action that she prefers over the one the receiver would like to take if he knew m^T. In other words, these sets contain the models under which the sender does better than under truthful communication. The following lemma proves that this will more often be the case facing a naïve receiver.

Lemma 3. *For any MLEU equilibrium (σ, ρ), we have $M_{(\sigma,\rho)} \subseteq M_{(\sigma^\dagger,\rho^\dagger)}$.*

The key observation is that

$$\begin{aligned}
& U_{m^T,h}(\rho(\sigma(m^T)), b) \\
& \leq \max\{ U_{m^T,h}(a(m^*, h, 0), b) \mid m^* \in M : \Pr(h \mid m^*, F_0) \geq \Pr(h \mid m^T, F_0) \} \quad (2) \\
& = U_{m^T,h}(\rho^\dagger(\sigma^\dagger(m^T)), b)
\end{aligned}$$

for every true model $m^T \in M$, since we consider the MLEU ambiguity rule. Thus, if the sender manages to profitably delude a receiver in equilibrium (σ, ρ) into taking an action favorable to her, i.e., the left-hand side of Expression (2) exceeds $U_{m^T,h}(a(m^T, h, 0), b)$ (and thus $m^T \in M_{(\sigma,\rho)}$), then she can also do so when facing a naïve receiver. Moreover, Expression (2) shows that the sender's expected utility is weakly *larger* in the latter case, which gives rise to an arguably strong notion of persuasive power in the spirit of state-wise dominance:

---

[14]Note that the only other choice under which this is the case is a model with the lowest expected fit $m^d \in \arg\min_{m \in M} \Pr(h|m, F_0)$, which however depends on the observed data h.

Definition 5 (Persuasive power ordering)**.** The sender has *(strictly) more persuasive power* under profile (σ, ρ) than under (σ', ρ') if $U_{m^T, h}(\rho(\sigma(m^T)), b) \geq U_{m^T, h}(\rho'(\sigma'(m^T)), b)$ for all $m^T \in M$ (and with strict inequality for some m^T).

At first glance, the notion of persuasive power provided in Definition 5 may appear demanding: We cannot rank two profiles if the sender prefers one profile to the other under a model m and vice versa under another model $m' \neq m$. However, as described above, it nonetheless allows for a comparison of the MLEU equilibria in our framework and the naïve model, with persuasive power being greater in the latter. Moreover, it follows from Lemma 3 and Expression (2) that the ranking is strict if the inclusion relation is strict.

Proposition 5. *For any MLEU equilibrium (σ, ρ), the sender has (strictly) more persuasive power in $(\sigma^\dagger, \rho^\dagger)$ than in (σ, ρ) (if $M_{(\sigma, \rho)} \subsetneq M_{(\sigma^\dagger, \rho^\dagger)}$).*

The following proposition characterizes cases in which the sender has the same or strictly more persuasive power when facing a naïve receiver.

Proposition 6. *For any history h and MLEU equilibrium (σ, ρ), the following two statements hold.*

(i) *$M_{(\sigma, \rho)} = M_{(\sigma^\dagger, \rho^\dagger)} = \emptyset$ and the sender's expected payoff in (σ, ρ) and in $(\sigma^\dagger, \rho^\dagger)$ is identical if $b \leq \underline{b}(h)$.*

(ii) *$M_{(\sigma, \rho)} \subsetneq M_{(\sigma^\dagger, \rho^\dagger)}$ and the sender has strictly more persuasive power in $(\sigma^\dagger, \rho^\dagger)$ than in (σ, ρ) if $b > \bar{b}(h)$, provided $U_{m,h}(a, b)$ allows for large conflicts of interest and there is m^T, m^* with $\Pr(h \mid m^*, F_0) \geq \Pr(h \mid m^T, F_0)$ and $a(M, h) \leq a(m^T, h, 0) < a(m^*, h, 0)$.*

The intuition of the above result is as follows. On the one hand, if $b \leq \underline{b}(h)$, the best the sender can do is to reveal the true model and make the receiver take action $a(m^T, b, 0)$. On the other hand, if $b > \bar{b}(h)$, provided preferences allow for large conflicts of interest, the only equilibrium is babbling, inducing the pooling action $a(M, h)$, i.e., the action the receiver would opt for if no communication took place. If a true model's best reply $a(m^T, h, 0)$ exceeds the pooling action and the sender can credibly convince a naïve receiver to adopt a higher action, she strictly improves her expected utility in a case where equilibrium play cannot.

We conclude this section by providing an instance within our running example showing that the sender may also have strictly more persuasive power when facing a naïve receiver in case equilibrium play is informative.
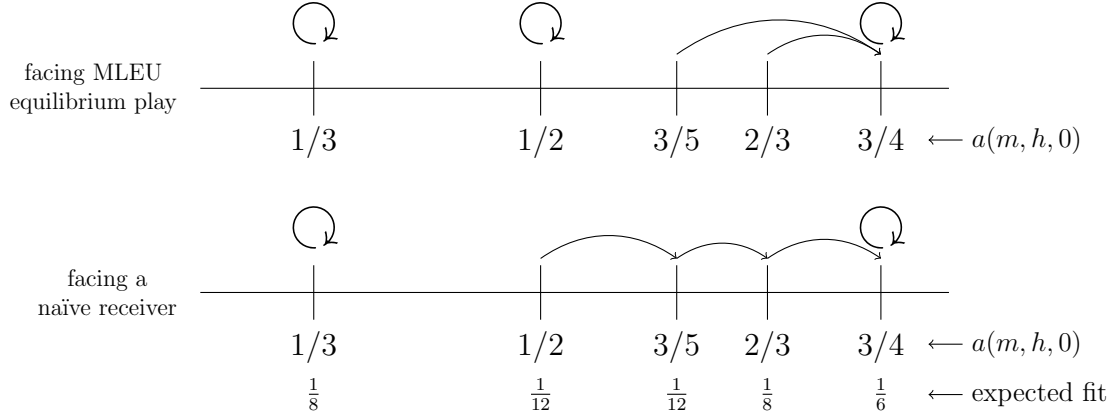
Figure 8: Illustration of equilibrium play and the naïve model in Example 6. The arrows indicate the action the sender manages to make the receiver take given the action optimal for the receiver under the true model.

Example 6. Consider the setting of Section 2.4, where $F_0 = \mathcal{U}(0,1)$, $u_S(a,\theta,b) = -(\theta + b - a)^2$, and h is generated according to (1), and the MLEU ambiguity rule. Suppose that $K = 3$, $h^\Sigma = 2$, $M = 2^{\{1,2,3\}} \setminus \{\emptyset\}$ and $b \in (\frac{1}{20}, \frac{3}{40}] \subseteq (\underline{b}(h), \overline{b}(h))$. A most informative MLEU equilibrium (σ, ρ) is given by $\sigma(m) = r_1$ if $a(m,h,0) = \frac{1}{3}$, $\sigma(m) = r_2$ if $a(m,h,0) = \frac{1}{2}$ and $\sigma(m) = r_3$ if $a(m,h,0) \in \{\frac{3}{5}, \frac{2}{3}, \frac{3}{4}\}$ (inducing posterior and action $\frac{3}{4}$). Facing a naïve receiver, the sender wants to and can use a narrative inducing a higher than the receiver's default action if $a(m^T,h,0) \in \{\frac{1}{2}, \frac{3}{5}, \frac{2}{3}\}$, see Figure 8 for an illustration.

We have $M_{(\sigma,\rho)} = \{m \mid a(m,h,0) = \frac{2}{3}\}$, a proper subset of $M_{(\sigma^\dagger,\rho^\dagger)} = \{m \mid a(m,h,0) \in \{\frac{1}{2}, \frac{3}{5}, \frac{2}{3}\}\}$. Furthermore, if $a(m^T,h,0) = \frac{1}{2}$ and $a(m^T,h,0) = \frac{3}{5}$, the sender prefers to induce action $\frac{3}{5}$ over action $\frac{1}{2}$ and action $\frac{2}{3}$ over action $\frac{3}{4}$, respectively. She is thus strictly better off facing a naïve receiver than facing equilibrium play. Finally, note that if $a(m^T,h,0) = \frac{3}{5}$, the sender would actually even prefer truthful communication over equilibrium play.

Our results and examples show that the sender generally prefers to deal with a naïve receiver. One might suspect that the receiver is always better off when following equilibrium play. However, this is not true as can be seen from Figure 8: On the one hand, if $a(m^T,h,0) = \frac{1}{2}$, the sender reveals the true model under equilibrium play, while she can induce action $\frac{3}{5}$ if the receiver was naïve. On the other hand, if $a(m^T,h,0) = \frac{3}{5}$, equilibrium play enforces the action $\frac{3}{4}$, whereas both would prefer the action $\frac{2}{3}$ which were induced if the receiver was naïve.

Remark 4. The receiver's expected payoff may be strictly higher in the naïve model

than following equilibrium play, or vice versa, depending on m^T. Hence, it is ex ante uncertain whether it is better for the receiver to be rational or naïve.

6   Conclusion

In this article, we model the communication of narratives as a cheap-talk game under model uncertainty. Our starting point was the literature initiated by S&S, which formalizes narratives as likelihood functions explaining observable data in the presence of model uncertainty. A key assumption in this literature is a form of receiver naïvité in the sense that the receiver does not take the sender's strategic incentives into account. In contrast, we provide a general game-theoretic framework in which the receiver is sophisticated and study equilibria under model uncertainty. To this end, we resolve any remaining model uncertainty by applying the novel concept of ambiguity rules, which allow for the study of MLEU, MEU, and smooth model preferences.

We characterize the set of equilibria by means of an algorithm: There is a positive integer N such that there is an n-step equilibrium which divides the set of narratives into n consecutive intervals if and only if $1 \leq n \leq N$. We then show that the maximum number N of consecutive intervals is decreasing in the conflict of interest and characterize informativeness thresholds that mark when full disclosure or partial disclosure is possible in equilibrium, respectively. Notably, these thresholds may depend in non-obvious ways on the history and the ambiguity rule. Finally, we compare our equilibrium framework under MLEU with S&S. We establish that the sender has more persuasive power in a strong sense – namely state-wise dominance – facing a naïve receiver than facing a sophisticated receiver of our framework. Interestingly, the receiver does not generally prefer to be rational himself.

Modeling assumptions.   Being based on S&S, our framework diverges in important ways from the classical literature on cheap-talk communication following the seminal work by Crawford and Sobel (1982): First, communication is not directly about the (payoff-relevant) state but about the model linking publicly observed outcomes and the state. Second, the model space is discrete. Third, the receiver faces ambiguity about the true model.

At least to us, it was thus not clear a priori whether well-known results from the literature carry over to our framework. Most importantly, there is our equilibrium char-

acterization, which closely resembles the one in Crawford and Sobel (1982). Moreover, we obtain precise informativeness thresholds like in other cheap-talk models with an imperfectly informed sender (e.g., Argenziano et al., 2016; Foerster, 2023). The relationship to the literature is less straightforward for the last set of results, as there is no obvious analogue to a naïve receiver à la S&S in the setting of Crawford and Sobel (1982). However, we could consider a receiver who takes the sender's message at face value. Such a receiver then follows the sender's recommendation, with the same outcome as under full delegation. While the sender clearly prefers delegation to cheap-talk communication, Dessein (2002) shows that the receiver also does so if the conflict of interest is small. On the contrary, for the receiver to prefer to be naïve in our framework not only requires the conflict of interest to be large enough to preclude truthful communication, but it also depends on the underlying true model.

Finally, one may wonder whether a typical equilibrium selection criterion, such as the NITS (no incentive to separate) criterion by Chen et al. (2008), has bite in our setting. The answer is yes and no. On the one hand, one can show that if an n-step equilibrium does not satisfy NITS, then there is an $n + 1$-step equilibrium, implying that an N-step equilibrium satisfies NITS. On the other hand, even a babbling equilibrium may satisfy NITS due to the model space being finite.

Outlook. In light of S&S, one might ask whether individuals take into account a sender's strategic incentives when being exposed to a narrative, as we assume they do, or simply adopt a narrative whenever it fits the data well, as they do. To this end, we conduct a laboratory experiment in a companion project. Different from Barron and Fries (2024), we focus on whether subject receivers take the size of the conflict of interest between them and the subject sender into account when evaluating the communicated narrative.

Our framework may also be extended in several directions. First, it does not embed the framework of Crawford and Sobel (1982) due to the finite model space, which precludes the sender from perfectly learning the state. One can show that this is possible in a tailor-made extension of our model to a specific continuous model space M. Second, one could introduce multiple senders who compete for convincing the receiver (or perhaps the receivers) along the lines of Krishna and Morgan (2001). S&S show that competition between persuaders pushes them to propose models that overfit the data. In turn, the

receiver underreacts to the data. In our framework with a sophisticated receiver, we suspect competition to benefit the receiver similarly to Krishna and Morgan (2001), with the extent depending on the exact protocol. Third, one could change the order of events, such that the data realizes only after the sender has communicated her narrative, similarly to Aina (2024). From the sender's point of view, the receiver's response to her narrative then is a random variable. Fourth, one could investigate alternatives to the NITS criterion in our setting. We leave these and related questions for future research.

References

Aina, C. (2024). Tailored stories.

Argenziano, R., S. Severinov, and F. Squintani (2016). Strategic information acquisition and transmission. *American Economic Journal: Microeconomics 8*(3), 119–155.

Barron, K. and T. Fries (2024). Narrative persuasion. Technical report, WZB Discussion Paper.

Bauch, G. and M. Foerster (2024). Strategic communication of narratives. Working paper available on https://arxiv.org/abs/2410.23259.

Bénabou, R., A. Falk, and J. Tirole (2020). Narratives, imperatives, and moral persuasion.

Bilotta, F. and G. Manferdini (2025). Coarse memory and plausible narratives.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics 24*(2), 265–272.

Chen, Y., N. Kartik, and J. Sobel (2008). Selecting cheap-talk equilibria. *Econometrica 76*(1), 117–136.

Choquet, G. (1954). Theory of capacities. In *Annales de l'institut Fourier*, Volume 5, pp. 131–295.

Colo, P. (2024). Communicating about confidence: Cheap talk with an ambiguity-averse receiver. *American Economic Journal: Microeconomics 16*(3), 43–75.

Constantino, S. M. and E. U. Weber (2021). Decision-making under the deep uncertainty of climate change: The psychological and political agency of narratives. *Current Opinion in Psychology 42*, 151–159.

Crawford, V. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1431–1451.

Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society: Series B (Methodological) 30*(2), 205–232.

Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies 69*(4), 811–838.

Eliaz, K., S. Galperti, and R. Spiegler (2025). False narratives and political mobilization. *Journal of the European Economic Association 23*(3), 983–1027.

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

Eliaz, K., R. Spiegler, and H. C. Thysen (2021). Strategic interpretations. *Journal of Economic Theory 192*, 105192.

Fagin, R. and J. Y. Halpern (1991). Uncertainty, belief, and probability 1. *Computational Intelligence 7*(3), 160–173.

Foerster, M. (2023). Strategic transmission of imperfect information: why revealing evidence (without proof) is difficult. *International Journal of Game Theory 52*(4), 1291–1316.

Foerster, M. and J. J. van der Weele (2021). Casting doubt: Image concerns and the communication of social impact. *The Economic Journal 131*(639), 2887–2919.

Garcia-Lorenzo, L. (2010). Framing uncertainty: Narratives, change and digital technologies. *Social Science Information 49*(3), 329–350.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Jain, A. (2023). Informing agents amidst biased narratives.

Kartik, N. and A. Kleiner (2024). Convex choice.

Kartik, N., S. Lee, and D. Rappoport (2024). Single-crossing differences in convex environments. *Review of Economic Studies 91*(5), 2981–3012.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Krishna, V. and J. Morgan (2001). A model of expertise. *Quarterly Journal of Economics 116*(2), 747–775.

Muraviev, I., F. Riedel, and L. Sass (2017). Kuhn's theorem for extensive form Ellsberg games. *Journal of Mathematical Economics 68*, 26–41.

Pedde, S., K. Kok, J. Onigkeit, C. Brown, I. Holman, and P. A. Harrison (2019). Bridging uncertainty concepts across narratives and simulations in environmental scenarios.

Regional Environmental Change 19, 655–666.

Schwartzstein, J. and A. Sunderam (2021). Using models to persuade. *American Economic Review 111*(1), 276–323.

Shafer, G. (1976). *A mathematical theory of evidence*, Volume 42. Princeton university press.

Shiller, R. J. (2017). Narrative economics. *American Economic Review 107*(4), 967–1004.

A Proofs

Proof of Lemma 1. In an equilibrium with sender strategy $\sigma \colon M \to \mathcal{R}$, the receiver can act upon the different messages in $\sigma(M)$, which yields $\#(\rho \circ \sigma)(M) = \#\{a(\sigma^{-1}(r), h, 0) \mid r \in \sigma(M)\} \leq \#\{a(m, h, 0) \mid m \in M\}$. $\qquad\square$

Proof of Lemma 2. (i) Fix any induced action $a^* \in \rho'(\sigma'(M))$. Define $M^* := \{m \in M \mid \rho'(\sigma'(m)) = a^*\}$, which is non-empty and contains some element m^*. Set $r^* := \sigma'(m^*)$ and define the sender strategy $\sigma(m) := r^*$ if $m \in M^*$ and $\sigma(m) := \sigma'(m)$ otherwise. Let ρ be the receiver's best reply to σ. Then

$$\rho(r^*) = a(M^*, h, 0) \in \mathrm{conv}(\{\underbrace{\rho'(\sigma'(m))}_{=a^*} \mid m \in M^*\}) = a^* = \rho'(r^*)$$

and $\rho(r) = a(\sigma'^{-1}(r), h, 0) = \rho'(r)$ if $r \in \sigma(M) \setminus \{r^*\}$. By construction, $\rho \circ \sigma = \rho' \circ \sigma'$ and thus σ is also a best reply to ρ. Note that the action a^* is induced solely by the report r^*. Repeating this construction for all other induced actions yields the result.

(ii) Let $a_1 < \cdots < a_N$ be the distinct action induced in equilibrium. By (i) we can pass to the corresponding reduced equilibrium (σ, ρ) and thus assume that the sender solely uses messages $r_1, \ldots, r_N$ with $\rho(r_i) = a_i$. Let $M_i := \sigma^{-1}(r_i) \subseteq M$. Since σ is a best reply to ρ, each M_i is an interval by interval choice. The M_i form a partition of M, and since $a_i = \rho(r_i) = a(M_i, h, 0) \in \mathrm{conv}(\{a(m, h, 0) \mid m \in M_i\})$ we find $m_i \leq m_j$, i.e., $a(m_i, h, 0) \leq a(m_j, h, 0)$, whenever $i < j$ and $m_i \in M_i, m_j \in M_j$. In other words, the M_i form a consecutive partition. Note in particular that induced actions $\rho(\sigma(m))$ and the receiver's bliss points $a(m, h, 0)$ are equally ordered. Consistent ordering concludes the proof. $\qquad\square$

Proof of Theorem 1. Fix any history h and bias b in the following. By Lemma 1, an equilibrium induces at most $\#\{a(m, h, 0) \mid m \in M\}$ different actions. Consequently, there is a maximum number $N \leq \#\{a(m, h, 0) \mid m \in M\}$ such that an N-step equilibrium exists. For any n-step equilibrium, the corresponding reduced equilibrium (σ, ρ) uses exactly n different messages $r_1, \ldots, r_n$ without changing the model-action profile $(m, \rho(\sigma(m)))$ by the Reduction Lemma 2. The proof of the Reduction Lemma furthermore establishes that $M_i = \sigma^{-1}(r_i)$ is a consecutive partition of M with $\sigma(m) = r_i$ for all $m \in M_i$ and the equilibrium conditions.

It remains to show that for every $1 \leq n \leq N$ there exists an n-step equilibrium. We do so by means of an algorithm that constructs an n-step equilibrium out of an $(n+1)$-step equilibrium. Starting from any N-step equilibrium, the claim follows. In the following, we explain the algorithm and show its properties.

Let (σ, ρ) be an $(n+1)$-step equilibrium. If $n = 1$ there is nothing to prove, as a babbling equilibrium always exists. Assume thus $n \geq 2$. By passing to the corresponding reduced equilibrium (σ, ρ), we find a consecutive partition $\{M_1, M_2, \ldots, M_{n+1}\}$ and messages $r_1, \ldots, r_{n+1}$ such that $\sigma(m) = r_i$ for all $m \in M_i$ and $\rho(r_i) < \rho(r_{i+1})$ for $i = 1, \ldots, n$. Note that, even though the M_i form a consecutive partition, we may have $m_i \in M_i$ and $m_{i+1} \in M_{i+1}$ with $a(m_i, h, 0) = a(m_{i+1}, h, 0)$, where, necessarily, $a(m_i, h, 0) = \max\{a(m, h, 0) \mid m \in M_i\}$ and $a(m_{i+1}, h, 0) = \min\{a(m, h, 0) \mid m \in M_{i+1}\}$. In order to account for these indifferences, it is useful to introduce the following notation. Define inductively

$$\mathfrak{m}_1 := \{m \in M_1 \mid a(m, h, 0) \leq a(m', h, 0) \, \forall m' \in M_1\},$$

$$\mathfrak{m}_{\ell+1} := \left\{m \in M \setminus \bigcup_{l=1}^{\ell} \mathfrak{m}_l \,\middle|\, \left(a(m, h, 0) \leq a(m', h, 0) \, \forall m' \in M \setminus \bigcup_{l=1}^{\ell} \mathfrak{m}_l\right) \right.$$
$$\left. \wedge \left(m \in M_{i+1} \Rightarrow M_i \subseteq \bigcup_{l=1}^{\ell} \mathfrak{m}_l \, \forall i\right)\right\},$$

and set $L := \max\{\ell \in \mathbb{N} \mid \mathfrak{m}_\ell \neq \emptyset\}$. I.e., we order models w.r.t. the receiver's bliss points $a(m, h, 0)$ and cluster those together that induce the same receiver action and lie in the same partition element M_i. Hence, $M_i = \bigcup_{l=\ell_i^0}^{\ell_{i+1}^0 - 1} \mathfrak{m}_l$ for according indices $1 =: \ell_1^0 < \ldots < \ell_i^0 < \ldots < \ell_{n+2}^0 := L + 1$. In the following, we will consider different partitions of M by identifying them with partitions of $\{1, \ldots, L\}$ that mark "cutoff points" for the $\mathfrak{m}_\ell$. This aids writing down the iterative procedure described later on.

Define now the two strategies σ_a and σ_b which result from merging the messages r_n, r_{n+1} respectively r_1, r_2 from σ. Formally, define the new cutoffs $\ell_i^a := \ell_i^1 := \ell_i^0$ for $i = 1, \ldots, n$ and $\ell_{n+1}^a := \ell_{n+1}^1 := \ell_{n+2}^0 = L + 1$, and $\ell_1^b := \ell_1^0 = 1$ and $\ell_i^b := \ell_{i+1}^0$ for $i = 2, \ldots, n+1$. The respective partition elements are thus $M_i^a := M_i^1 = \bigcup_{l=\ell_i^1}^{\ell_{i+1}^1 - 1} \mathfrak{m}_l$ and $M_i^b = \bigcup_{l=\ell_i^b}^{\ell_{i+1}^b - 1} \mathfrak{m}_l$. Note that $M_i = M_i^1$ for $i = 1, \ldots, n-1$ and $M_n \cup M_{n+1} = M_n^1$ as well

as $M_1 \cup M_2 = M_1^b$ and $M_{i+1} = M_i^b$ for $i = 2, \ldots, n$. In terms of the original partitions,

$$\sigma_a(m) := \sigma_a^1(m) := \begin{cases} r_i & , \text{ if } m \in M_i, i \neq n, n+1, \\ r_n & , \text{ if } m \in M_n \cup M_{n+1} \end{cases},$$

$$\sigma_b(m) := \begin{cases} r_1 & , \text{ if } m \in M_1 \cup M_2, \\ r_{i-1} & , \text{ if } m \in M_i, i \neq 1, 2 \end{cases}.$$

If $\sigma_a = \sigma_a^1$ together with the unique best reply $\rho_a = \rho_a^1$ of the receiver is an n-step equilibrium, we are done. Suppose now that (σ_a, ρ_a) is not an n-step equilibrium. Then, there is a model m^* for which the sender has a profitable deviation from σ_a by inducing an action $\rho_a(r) \neq \rho_a(\sigma_a(m^*))$. Since (σ, ρ) is an equilibrium and ρ agrees with ρ_a on the partition elements $M_1, \ldots, M_{n-2}$, we must have $m^* \in M_{n-1} \cup M_n \cup M_{n+1}$. As $M_n \subseteq M_n \cup M_{n+1}$, we have $\rho(r_n) \leq \rho_a(r_n)$ by the hedging property of the ambiguity rule. By strict concavity of the utility function $U_{m^*,h}(\cdot, b)$, we conclude $m^* \in M_n \cup M_{n+1} = M_n^1 = \bigcup_{l=\ell_n^1}^{\ell_{n+1}^1 - 1} \mathfrak{m}_l$. By interval choice and consistent ordering, we find that a profitable deviation must be possible for all $m \in \mathfrak{m}_{\ell_n^1}$ by inducing $\rho_a^1(r_{n-1})$ instead. Similarly, one can see that a profitable deviation from (σ_b, ρ_b) can only be to send r_2 instead of r_1 for some $m^* \in M_1 \cup M_2$, which will become important later.

In the next step (and the following), we modify strategy σ_a^1 according to the profitable deviation by defining $\ell_n^2 := \ell_n^1 + 1$ and $\ell_i^2 := \ell_i^1$ for all other i as well as $M_i^2 := \bigcup_{l=\ell_i^2}^{\ell_{i+1}^2 - 1} \mathfrak{m}_l$. Then, $\rho_a^2(r_i) = a(M_i^2, h, 0)$ is the unique best reply to $\sigma_a^2(m) = r_i$ if $m \in M_i^2$. Note that $\rho_a^2(r_i) \geq \rho_a^1(r_i)$ for all $i = 1, \ldots, n$. Hence, if (σ_a^2, ρ_a^2) is still not an equilibrium, a profitable deviation must again concern some, and indeed all, $m \in \mathfrak{m}_{\ell_{i_2}^2}$ for some $i_2 \in \{1, \ldots, n\}$ by means of sending r_{i_2-1} instead of r_{i_2}. Picking any such deviation, we set $\ell_{i_2}^3 := \ell_{i_2}^2 + 1$ and $\ell_i^3 := \ell_i^2$ for all other i, which defines new partition elements M_i^3 and thus a new σ_a^3 and ρ_a^3. That way, an iterative process is defined, specifying new cutoffs ℓ_i^k and strategies σ_a^k and ρ_a^k in each step k.

We claim that this process delivers an n-step equilibrium at some iteration step and thus terminates. First observe that ρ_a^k will always induce n different actions: Assume by means of contradiction that $\rho_a^k(r_i) < \rho_a^k(r_{i+1})$, but $\rho_a^{k+1}(r_i) = \rho_a^{k+1}(r_{i+1}) =: a^*$ at some iteration step for some i. By the hedging property of an ambiguity rule, we must have $a(M_i^{k+1}, h, 0) = a^* = a(M_{i+1}^{k+1}, h, 0) = a(m^*, h, 0)$ for any $m^* \in \mathfrak{m}_{\ell_{i+1}^k - 1} \cup \mathfrak{m}_{\ell_{i+1}^k}$. Again by the hedging property, we find $\rho_a^k(r_{i+1}) = a(M_{i+1}^k, h, 0) = a^*$ as $M_{i+1}^k = M_{i+1}^{k+1} \cup \mathfrak{m}_{\ell_{i+1}^k}$.

36

Consider $m^* \in \mathfrak{m}_{\ell_{i+1}^k} \subseteq M_{i+1}^k$ and note that $a(m^*, h, b) \geq a(m^*, h, 0) = a^* = \rho_a^k(r_{i+1}) > \rho_a^k(r_i)$, contradicting that the iteration step $i \to i+1$ was a profitable deviation.

Second, if the process keeps not terminating at an equilibrium, we eventually find $\sigma_a^k = \sigma_b$ for some k and, in that case, show that (σ_b, ρ_b) is an equilibrium: Note that $\ell_i^k \leq \ell_i^{k+1}$, being strict for exactly one $i \neq 1$ at each iteration step k. Consequently, there is going to be a minimal k^* such that $\ell_i^{k^*} = \ell_i^b$ for some $i \neq 1$. We now prove that $\ell_i^k = \ell_i^b$ for all $k > k^*$, i.e., there will not be a profitable deviation from shifting the cutoff farther to the right. Note that $1 = \ell_1^{k^*} = \ell_1^b$ and $\ell_i^{k^*} \leq \ell_i^b$ for all $i \neq 1$, thus $\rho_a^{k^*}(r_i) \leq \rho_b(r_i)$. If there was a profitable deviation from σ_a^k for $m^* \in \mathfrak{m}_{\ell_{i+1}^{k^*}} \subseteq M_{i+1}^{k^*}$ to send r_i instead, i.e., $U_{m^*, h}(\rho_a^k(r_i), b) > U_{m^*, h}(\rho_a^k(r_{i+1}), b)$, then $U_{m^*, h}(\rho_b(r_i), b) > U_{m^*, h}(\rho_b(r_{i+1}), b)$ since $a(m^*, h, b) > \rho_a^k(r_i)$, and thus this would already be the case for σ^b. However, recall that (σ_b, ρ_b) can only have profitable deviations in which the cutoffs are shifted to the left, i.e., sending a higher message. Thus, $\ell_i^{k^*} = \ell_i^{k^*+1}$ in the next iteration. In the further iteration process, we thus keep the property $\rho_a^k(r_i) \leq \rho_b(r_i)$ for all i, showing $\ell_i^k = \ell_i^b$ for $k > k^*$. As soon as $\ell_j^k = \ell_j^b$ for another $j \neq i$, there again is no profitable deviation increasing this cutoff. As one cutoff always increases if there is a profitable deviation, we eventually find $\ell_i^{\tilde{k}} = \ell_i^b$ for all i at some iteration step $\tilde{k}$, i.e., $(\sigma_a^{\tilde{k}}, \rho_a^{\tilde{k}}) = (\sigma_b, \rho_b)$. Note that now no further deviation is profitable. Hence, (σ_b, ρ_b) must be an equilibrium and the iteration process terminates no later than $\tilde{k}$. $\qquad\square$

Proof of Corollary 1. Applying arguments from the algorithm, we show how an n-step equilibrium (σ, ρ) under b can be transformed into an n-step equilibrium under b'. Start with σ and the receiver's best reply ρ' under b'. If (σ, ρ') is also an equilibrium under b', we ware done. If not, the sender has an incentive to deviate. As $b' < b$, a profitable deviation entails sending a message r_l that induces a smaller action $\rho'(r_l)$ for a left-most model-package $\mathfrak{m}_{l+1} \subseteq M_{l+1}$. If the resulting sender strategy together with the receiver's best reply to it does not constitute an equilibrium, we argue similarly to the proof of Theorem 1 that any profitable deviation is of the type described above and that the step-size does not go down. The algorithm has to terminate, and thus produce an n-step equilibrium under b', as otherwise we eventually reach the sender strategy whose partition consists of $n-1$ singletons to the right and one interval which bunches all other models to the left, precluding further profitable deviations. $\qquad\square$

Proof of Proposition 1. (i) Consider any fully informative strategy σ of the sender. Then, σ is injective and receiving r, the minimal feasible set is $\tilde{M} = \sigma^{-1}(r) = \{m^T\}$,

the singleton containing the true model. The best response ρ is thus given by

$$\rho(r) = a(\sigma^{-1}(r), h, 0) = \arg\max_a \mathbb{E}[u_R(a, \theta) \mid \sigma^{-1}(r), h].$$

For now, consider any $m^T \in M$ with $a^T := a(m^T, h, 0)$ and let $a^* := \min\{a = a(m, h, 0) \mid m \in M, a^T < a\}$ and define $H_{m^T, h} \colon \mathbb{R}_{\geq 0} \to \mathbb{R}, b \mapsto U_{m^T, h}(a, b) - U_{m^T, h}(a', b)$, which is a continuous function. Note that the sender has an incentive to deviate from revealing m^T if and only if $H_{m^T, h}(b) < 0$ as $\rho(\sigma(m^T)) = a(m^T, h, 0) < a(m^T, h, b)$ by supermodularity. Define $\underline{b}(m^T, h) := \min\{b \geq 0 \mid H_{m^T, h}(b) = 0\} \cup \{\infty\}$. Since $H_{m^T, h}(0) > 0$ we have $\underline{b}(m^T, h) > 0$. Now set $\underline{b}(h) := \min\{\underline{b}(m^T, h) \mid m^T \in M \colon a(m^T, h, 0) < \max_{m \in M} a(m, h, 0)\}$. We have $\underline{b}(h) > 0$ as M is finite and for all $b \leq \underline{b}(h)$ the sender has no profitable deviation from the fully revealing strategy.

(ii) $V(m^T, h, b)$ is increasing in b by supermodularity. Under a large conflict of interest, we let $\bar{b}(h)$ be the infimum of all $\hat{b}(h)$ with $V(m^T, h, b) > 0$ for all $b > \bar{b}(h)$ and all non-maximal m^T. For $b > \bar{b}(h)$, any equilibrium is babbling while for $b \leq \bar{b}(h)$ there exists a 2-step equilibrium. $\qquad\square$

Proof of Proposition 2. Fix any h and let m^T be arbitrary. In order to show that the sender is better off in (σ, ρ), it suffices to show that the sender can induce the action $\rho'(\sigma'(m^T))$ in equilibrium (σ, ρ) as well. As σ is weakly more informative than σ', the set of models $M' := \sigma'^{-1}(\sigma'(m^T))$ decomposes into a (possibly singleton) partition $\dot\bigcup_{r \in \tilde{\mathcal{R}}} \sigma^{-1}(r)$ for a subset $\tilde{\mathcal{R}} \subseteq \mathcal{R}$. The MLEU rule of the receiver selects some $m' \in M'$ under σ' with highest expected fit and chooses action $a(m', h, 0)$. Let $r' \in \tilde{\mathcal{R}}$ be such that $m' \in \sigma^{-1}(r')$. By the consistent tiebreaking MLEU rule, m' is also chosen under the minimal feasible set $\sigma^{-1}(r')$ and thus $\rho(\sigma(r')) = a(m', h, 0)$. $\qquad\square$

Proof of Proposition 3. Consider any h^Σ. By the Reduction Lemma 2, we can identify the set of models with the set of induced actions (and beliefs) under h^Σ,

$$\mathcal{A}(h^\Sigma) = \left\{ a(m, h', 0) \mid m \in M, \sum_{k=1}^{K} h'_k = h^\Sigma \right\} = \left\{ \frac{\sum_{i \in m} h'_i + 1}{\#m + 2} \mid m \in M, \sum_{k=1}^{K} h'_k = h^\Sigma \right\}.$$

Note that $\underline{b}(h^\Sigma) = \frac{1}{2} \min_{a, a' \in \mathcal{A}(h^\Sigma) \colon a \neq a'} |a - a'|$, and thus $b(h^\Sigma)$ is independent of the ambiguity rule. Symmetry then follows since, for any given $k \in \{1, 2, \ldots, K\}$,

$$\frac{s+1}{k+2} \in \mathcal{A}(h^\Sigma) \iff \max\{0, h^\Sigma - K + k\} \leq s \leq \min\{k, h^\Sigma\}$$

$$\Longleftrightarrow \max\{0, K - h^\Sigma - K + k\} \le k - s \le \min\{k, K - h^\Sigma\}$$

$$\Longleftrightarrow \frac{k - s + 1}{k + 2} \in \mathcal{A}(K - h^\Sigma) \Longleftrightarrow 1 - \frac{s + 1}{k + 2} \in \mathcal{A}(K - h^\Sigma).$$

For the second claim, note first that the smallest common divisor of distinct $a, a' \in \mathcal{A}(h^\Sigma)$ is at most $(K+1)(K+2)$, which implies $|a - a'| \ge \frac{1}{(K+1)(K+2)}$. This inequality is an equality only if $a = \frac{s+1}{K+2}$ and $a' = \frac{s'+1}{K+1}$ (or vice versa) for $0 \le s \le K$ and $0 \le s' \le K - 1$. Next, observe that

$$|a - a'| = \frac{1}{(K+1)(K+2)} \Longleftrightarrow |(s+1)(K+1) - (s'+1)(K+2)| = 1$$

$$\Longleftrightarrow |s(K+1) - s'(K+2) - 1| = 1$$

$$\Longleftrightarrow s'(K+2) = s(K+1) \lor s = s' + \frac{2 + s'}{K+1}$$

$$\Longleftrightarrow s = s' = 0 \lor (s = K \land s' = K - 1),$$

where the last step follows from $s(K+1) > (s-1)(K+2)$ for all $s \le K$ and $\left(\frac{2+s'}{K+1} \in \mathbb{N} \land 0 \le s' \le K - 1\right)$ $s' = K - 1$. The claim then follows since $\frac{1}{K+2} \in \mathcal{A}(h^\Sigma) \Longleftrightarrow h^\Sigma = 0$ and $\frac{K+1}{K+2} \in \mathcal{A}(h^\Sigma) \Longleftrightarrow h^\Sigma = K$. $\qquad\square$

Proof of Proposition 4. Fix any $K \ge 2$ and consider first the MLEU ambiguity rule. Consider $h^\Sigma = K$, i.e., , i.e., $h = \vec{1}$, and note that in this case $a(m, h, 0) = \frac{\#m+1}{\#m+2}$ for any m. Thus, the expected fit of model m is

$$\Pr(h \mid m, F_0) = \int_0^1 \theta^{\#m} \left(\frac{1}{2}\right)^{K - \#m} d\theta = \left(\frac{1}{2}\right)^{K - \#m} \frac{1}{\#m + 1},$$

which is (strictly) increasing in $\#m$ (if $\#m \ge 1$). Moreover, any 2-step partition can be written as $M_1^{\vec{1}}(k) := \{m \mid \#m \le k\}$ and $M_2^{\vec{1}}(k) = \{m \mid \#m > k\}$ for some $k \in \{0, 1, \ldots, K - 1\}$; the induced actions then are $a(M_1^{\vec{1}}(k), \vec{1}) = \frac{k+1}{k+2}$ and $a(M_2^{\vec{1}}(k), \vec{1}) = \frac{K+1}{K+2}$.[15] The sender has no incentives to deviate if and only if she has no incentives to deviate when $m^T \in \{m | \#m = k\}$ (then $a(m, h, b) = \frac{k+1}{k+2} + b$), which is equivalent to

$$a(m, h, b) - a(M_1^{\vec{1}}(k), \vec{1}) \le a(M_2^{\vec{1}}(k), \vec{1}) - a(m, h, b) \Longleftrightarrow b \le \frac{1}{2}\left(\frac{K+1}{K+2} - \frac{k+1}{k+2}\right).$$

Thus, we obtain $\bar{b}(h) = \max_{k=0,1,\ldots,K-1} \frac{1}{2}\left(\frac{K+1}{K+2} - \frac{k+1}{k+2}\right) = \frac{K}{4(K+2)}$.

---

[15]Note the we break the tie in expected fit between a narrative m such that $\#m = 1$ and one such that $\#m = 0$ (both 2^{-K}) by adopting the first one.

Next, consider $h^\Sigma = 0$, i.e., $h = \vec{0}$. Analogously to the first case, $a(m, h, 0) = \frac{1}{\#m+2}$ for any m and the expected fit of model m is (strictly) increasing in $\#m$ (if $\#m \geq 1$). Therefore, any 2-step partition can be written as $M_1^{\vec{0}}(k) := \{m \mid \#m \geq k\}$ and $M_2^{\vec{0}}(k) = \{m \mid \#m < k\}$ for some $k \in \{1, 2, \ldots, K\}$; the induced actions then are $a(M_1^{\vec{0}}(k), \vec{0}) = \frac{1}{K+2}$ and $a(M_2^{\vec{0}}(k), \vec{0}) = \frac{1}{k+1}$. The sender has no incentives to deviate if and only if she has no incentives to deviate when $m^T \in \{m \mid \#m = k\}$, which is equivalent to

$$a(m, h, b) - a(M_1^{\vec{0}}(k), \vec{0}) \leq a(M_2^{\vec{0}}(k), \vec{0}) - a(m, h, b)$$
$$\iff b \leq \frac{1}{2(k+1)} + \frac{1}{2(K+2)} - \frac{1}{k+2}.$$

Thus, we obtain $\bar{b}(h) = \max_{k=1,2,\ldots,K} \frac{1}{2(k+1)} + \frac{1}{2(K+2)} - \frac{1}{k+2} = \frac{1}{2(K+1)(K+2)}$. Finally, a fully informative strategy is part of an equilibrium if and only if

$$b \leq \frac{1}{2} \min_{m,m' \in M : a(m,h,0) \neq a(m',h,0)} |a(m, h, 0) - a(m', h, 0)| = \frac{1}{2(K+1)(K+2)}.$$

Second, consider the MEU ambiguity rule. Consider $h^\Sigma = 0$, i.e., $h = \vec{0}$, and recall that in this case $a(m, h, 0) = \frac{1}{\#m+2}$ for any m. By considering $k \in \{1, 2 \ldots, K\}$, we go through all bipartitions in the order of their bliss points by setting $M_1^{\vec{0}}(k) := \{m \mid \#m \geq k\}$ and $M_2^{\vec{0}}(k) = \{m \mid \#m < k\}$. We now state the optimal actions under the MEU ambiguity rule for the relevant partition elements, providing a derivation at the end of this proof.

$$a(M_1^{\vec{0}}(k), \vec{0}) = \min\left\{ \frac{1}{k+2}, \frac{K+k+5}{(K+3)(k+3)} \right\}, \tag{3}$$
$$a(M_2^{\vec{0}}(k), \vec{0}) = \min\left\{ \frac{1}{2}, \frac{k+4}{3(k+2)} \right\}. \tag{4}$$

A bipartition, indexed by k, is an equilibrium if and only if b satisfies the inequality $\left| \frac{1}{k+2} + b - a(M_1^{\vec{0}}(k), \vec{0}) \right| \leq \left| \frac{1}{k+2} + b - a(M_2^{\vec{0}}(k), \vec{0}) \right|$, i.e., the sender has no incentive to deviate to the rightmost bliss point in $M_1^{\vec{0}}(k)$. Thus, with (3) and (4) we obtain

$$\bar{b}(0) = \max_{k=1,2,\ldots,K} \frac{a(M_1^{\vec{0}}(k), \vec{0}) + a(M_2^{\vec{0}}(k), \vec{0})}{2} - \frac{1}{k+2}$$
$$= \max_{k=1,2,\ldots,K} \frac{1}{2} \left(\min\left\{ \frac{1}{k+2}, \frac{K+k+5}{(K+3)(k+3)} \right\} + \min\left\{ \frac{1}{2}, \frac{k+4}{3(k+2)} \right\} \right) - \frac{1}{k+2}$$
$$= \frac{K+1}{6(K+2)}.$$

We now turn towards $h^\Sigma = K$, i.e., $h = \vec{1}$. Recall that in this case $a(m, h, 0) = \frac{\#m+1}{\#m+2}$. The subsequent analysis can be based on the previous calculations by observing the following symmetry to the case $h = \vec{0}$: Define the bipartition via $M_1^{\vec{1}}(k) := \{m \mid \#m < k\}$ and $M_2^{\vec{1}}(k) = \{m \mid \#m \geq k\}$ for $k \in \{1, 2 \ldots, K\}$. The expected utility functions for the singleton models are

$$U_{\{m\},\vec{0}}(a) = -(a - \tfrac{1}{\#m+2})^2 - \frac{\#m+1}{(\#m+2)^2(\#m+3)}$$

$$U_{\{m\},\vec{1}}(a) = -(a - \tfrac{\#m+1}{\#m+2})^2 - \frac{\#m+1}{(\#m+2)^2(\#m+3)},$$

i.e., parabolas the maxima of which have the relation $a(\{m\}, \vec{0}) = 1 - a(\{m\}, \vec{1})$ and who share the same height in their maxima. Consequently, also $a(\tilde{M}, \vec{0}) = 1 - a(\tilde{M}, \vec{1})$ for each $\tilde{M} \subseteq \{0, \ldots, K\}$. In analogy to before, we thus see that

$$\begin{aligned}
\bar{b}(K) &= \max_{k=1,2,\ldots,K} \frac{a(M_1^{\vec{1}}(k), \vec{1}) + a(M_2^{\vec{1}}(k), \vec{1})}{2} - \frac{k}{k+1} \\
&= \max_{k=1,2,\ldots,K} \frac{1}{k+1} - \frac{a(M_1^{\vec{0}}(k), \vec{0}) + a(M_2^{\vec{0}}(k), \vec{0})}{2} \\
&= \max_{k=1,2,\ldots,K} \frac{1}{k+1} - \frac{1}{2}\left(\min\left\{\frac{1}{k+2}, \frac{K+k+5}{(K+3)(k+3)}\right\} + \min\left\{\frac{1}{2}, \frac{k+4}{3(k+2)}\right\}\right) \\
&= \max\left\{\frac{1}{12}, \frac{K}{8(K+3)}\right\}.
\end{aligned}$$

Finally, it is left to derive the optimal actions under the MEU ambiguity rule (3) and (4). Consider the following fact: Let $f_i(x) := f_{\alpha_i, \beta_i}(x) = -(x - \alpha_i)^2 - \beta_i$ for real numbers $\alpha_1 > \ldots > \alpha_n$ and $\beta_1 \geq \ldots \geq \beta_n$. Note that these functions are all strictly quasi-concave: f_i is strictly increasing for $x < \alpha_i$ and strictly decreasing for $\alpha_i < x$. The minimum function $f := \min\{f_1, \ldots, f_n\}$ is also strictly quasi-concave with unique maximizer $\alpha^* := \arg\max_x f(x)$. For $x > \alpha_1$, f is strictly decreasing as all f_i are, thus $\alpha^* \leq \alpha_1$. Likewise, $\alpha_n \leq \alpha^*$. Note that α^* is either one of the α_i or marks the intersection of at least two functions. By assumption on the β_i's and the previous insight, α^* is thus either α_1 or an intersection on the interval $[\alpha_n, \alpha_1]$. If α^* is an intersection of a set $I \subseteq \{1, \ldots, n\}$ of functions, there must be $i, j \in I$ with $\alpha_i \leq \alpha^* \leq \alpha_j$, since otherwise f was either strictly decreasing or increasing in an open neighborhood of α^* (the set of intersections is finite and thus contains no accumulation points). Now, observe that for each $i \in \{2, \ldots, n\}$ and $x \leq \alpha_i$ we have $f_1(x) < f_i(x)$, for the functions in question,

since $\alpha_i < \alpha_1$, $f_1(\alpha_1) = \beta_1 < \beta_i = f_i(\alpha_i)$ and $\frac{\partial}{\partial x} f_1 = -2x + 2\alpha_1 < -2x + 2\alpha_i = \frac{\partial}{\partial x} f_i$. Together, the last two insights reveal that if α^* is an intersection point, it must intersect f_1. As a result, α^* is the minimum among α_1 and the intersection points of f_1 with the other f_i. In our concrete case, $f_k(x) := -(x - \frac{1}{k+2})^2 - \frac{k+1}{(k+2)^2(k+3)}$. If we consider $m' \in M_1^{\vec{0}}(k) = \{m \mid \#m \geq k\}$ and let $k' = \#m'$, the intersection of $f_{k'}$ with f_k is $\frac{k'+k+5}{(k+3)(k'+3)}$, which is strictly decreasing in k' and thus minimized for $k' = K$. With $\alpha_1 = \frac{1}{k+2}$ we hence obtain (3). Similarly for $m' \in M_2^{\vec{0}}(k) = \{m \mid \#m < k\}$ with $k' = \#m'$ we find the intersections at $\frac{k'+5}{3(k'+3)}$ which is minimized by $k' = k - 1$. With $\alpha_1 = \frac{1}{2}$ we thus obtain (4). $\qquad\square$

Proof of Lemma 3. Let $m^T \in M_{(\sigma,\rho)}$. Then $\rho(\sigma(m^T)) \neq a(m^T, h, 0)$. Since we consider an MLEU equilibrium, we have $\rho(\sigma(m^T)) = a(m^*, h, 0)$ for some $m^* \in \sigma^{-1}(\sigma(m^T))$ with $\Pr(h \mid m^*, F_0) \geq \Pr(h \mid m^T, F_0)$. Thus, $m^T \in M_{(\sigma^\dagger, \rho^\dagger)}$. $\qquad\square$

Proof of Proposition 5. By (2), $(\sigma^\dagger, \rho^\dagger)$ has more persuasive power than (σ, ρ). If now $m^T \in M_{(\sigma^\dagger, \rho^\dagger)} \setminus M_{(\sigma,\rho)}$, we have strictly more persuasive power as

$$U_{m^T, h}(\rho(\sigma(m^T)), b) = U_{m^T, h}(a(m^T, h, 0), b) < U_{m^T, h}(\rho^\dagger(\sigma^\dagger(m^T)), b).$$

$\qquad\square$

Proof of Proposition 6. First, let $b \leq \underline{b}(h)$. At any true model m^T, the sender cannot do better than inducing $a(m^T, h, 0)$. Consequently, $M_{(\sigma,\rho)} = M_{(\sigma^\dagger, \rho^\dagger)} = \emptyset$. Second, if $b > \bar{b}(h)$, provided preferences allow for large conflicts of interest, the only equilibrium is babbling, inducing the constant $a(M, h)$. For every true model m^T with $a(M, h) \leq a(m^T, h, 0)$, the sender's utility thus stays at $U_{m^T, h}(a(M, h), b) \leq U_{m^T, h}(a(m^T, h, 0), b)$ in equilibrium. If there exists m^* with $m^T < m^*$ and $\Pr(h \mid m^*, F_0) \geq \Pr(h \mid m^T, F_0)$, a naïve receiver can be convinced to take action $a(m^*, h, 0)$ under the true model m^T, which strictly improves the sender's expected utility. $\qquad\square$