

The Power of Referential Advice*

Steven Callander[†] Nicolas S. Lambert[‡] Niko Matouschek[§]

April 2019

Abstract

Expert advice is often rich and broad, going beyond a simple recommendation. A doctor, for example, often provides information about treatments beyond the one that she recommends. In this paper we show that this additional, referential information plays an important strategic role in expert advice and, in fact, that it is vital to an expert’s power. We develop this result in the context of the canonical model of strategic communication with hard information, enriching the model with a notion of expertise that allows for a meaningful distinction between a recommendation and referential information. We identify an equilibrium in which, with probability one, the expert is strictly better off by providing referential advice than she is in any equilibrium in which she provides a recommendation alone. The benefit of referential advice to the expert is non-monotonic in the complexity of her expertise, reaching its peak when expertise is moderately complex.

*We thank Dirk Bergemann, Ben Brooks, Wouter Dessein, Marina Halac, Gilat Levy, Meg Meyer, David Myatt, Motty Perry, Andrea Prat, Luis Rayo, Bruno Strulovici, and participants at various conferences and seminars for their comments and suggestions. Lambert thanks Microsoft Research and the Cowles Foundation at Yale University for their hospitality and financial support.

[†]Stanford University, Stanford Graduate School of Business; sjc@stanford.edu.

[‡]Stanford University, Stanford Graduate School of Business; nlambert@stanford.edu.

[§]Northwestern University, Kellogg School of Management;
n-matouschek@kellogg.northwestern.edu

1 Introduction

Advice takes many forms. A common form is for an expert to simply offer a recommendation: a librarian recommends a book, a travel agent recommends a tour, or a sales assistant recommends a pair of shoes. In many situations, however, an expert does not limit herself to only a recommendation. Instead, the expert provides advice that is more expansive and that conveys information about decisions beyond the one recommended. This richer, contextual advice—what linguists refer to as *referential* information (Jakobson 1960; Sobel 2013)—is particularly relevant when the expert possesses complex knowledge. For instance, in addition to recommending a treatment, a doctor will often discuss alternative treatments and why she does not recommend them. Similarly, a mechanic might detail likely outcomes should a car owner undertake only superficial repairs instead of the more extensive ones she does recommend.

The role that referential information plays in the supply of expert advice has not previously been examined. As such, it is unclear whether referential information plays a meaningful role in communication or whether it is superfluous or even babbling. The objective of this paper is to offer the first analysis of referential information and address these questions.

Our central insight is that referential information plays an important strategic role in communication and, in fact, is vital to an expert’s power. By supplying referential information, and by supplying it in just the right way, the expert is able to leverage her expertise and systematically sway decisions in her favor. We identify an equilibrium in which the expert is strictly better off when she provides referential advice, regardless of the true state of the world, than she is in *any* equilibrium in which she provides only a recommendation. Thus, a mechanic, regardless of the true damage to a car, is able to induce the owner to spend more money on repairs if she provides referential information in addition to a recommendation.

The model we analyze is one of hard, or verifiable, information (Milgrom 1981; Grossman 1981). We follow that literature in assuming that the receiver’s preferences are defined over outcomes, whereas the expert—the sender—cares only about the decision taken. To this framework we introduce a novel conception of expertise, one that allows for a meaningful distinction between a recommendation and referential information. Specifically, we suppose that each decision is associated with a unique state that determines the outcome of that decision, and that the states are imperfectly correlated. The novelty of this approach is that it allows the expert to communicate precisely yet imperfectly—by revealing the outcome of a state variable, the expert reveals fully the outcome of that decision, but does not reveal all of her information.

This informational structure captures an important dimension of expertise in practice. It allows a doctor, for instance, to reveal precisely the outcome a patient can expect from a treatment without revealing all her information and rendering the patient an expert. This implies that the patient can trust the doctor’s recommendation, but at the same time be unable to perfectly evaluate that recommendation against alternatives. The patient is not without information, however, as he can glean some information from the doctor’s recommendation. For example, the recommendation to bandage an ankle injury suggests that simply keeping weight off of the ankle might work fine, even though it reveals little about the likely outcome of, say, invasive surgery. The degree of informational spillover across decisions reflects the *complexity* of an expert’s knowledge and is captured (and parameterized) in our model by the correlation across the state variables.

The construction of expertise in our model contrasts with the literature in which the expert’s knowledge is relevant across multiple decisions, even when that knowledge is rich and multidimensional. This distinction is easiest to see when expertise is modeled in its simplest form as a single piece of information, as in the classic formulations of Milgrom (1981) and Crawford and Sobel (1982). To communicate precisely with simple expertise, the expert must also communicate perfectly. Because the same, single piece of information affects all decisions (often in an identical way), advice about one decision is necessarily advice about all decisions, leaving no room for referential advice. This structure generates the seminal result of the hard information literature—the famous “unravelling” result—that in the unique equilibrium the expert’s information advantage unravels and she fully reveals her information to the receiver (Milgrom 1981; Grossman 1981; Milgrom and Roberts 1986; Matthews and Postlewaite 1985). As a result, the expert retains no leverage and the decision that is made aligns fully with the preferences of the decision maker.¹

Our richer notion of expertise provides the expert with greater ability to keep her information private and we show that, in equilibrium, the expert is able to do this to the maximal extent. We identify a continuum of equilibria in which the expert reveals the minimum amount of information to influence decision making—i.e., she reveals only the outcome of a single state variable. The decision that is revealed constitutes a recommendation—what linguists refer to as the *conative* function of language (Jakobson 1960)—and the receiver follows the recommendation.

¹Subsequent literature has generalized the informational structure of the classic models and explored the limits of unraveling, although in directions different from ours. We discuss this work below.

The existence of these equilibria—which we refer to as *conative equilibria*—imply that rich expertise does not necessarily compel the expert to communicate in a rich way. A doctor can, despite her extensive knowledge, simply recommend a treatment and know that the patient will follow her advice. As these equilibria maximize the expert’s ability to shield her information from the receiver, it may be reasoned that they are the expert’s preferred equilibria. Our main result is that this is not true. We identify a *referential equilibrium* in which the expert reveals strictly more of her knowledge to the receiver than a single point. Doing so leaves the expert strictly better off but the receiver strictly worse off. That referential information has this impact is not immediate. By construction, the referential information that is revealed cannot change the decision maker’s beliefs about the recommendation itself. How then does referential information have such an impact?

To understand the power of referential information it is important to understand that effective communication is a process of both persuasion and dissuasion. To *persuade* the receiver to follow a recommendation, the expert must simultaneously *dissuade* him from taking any other decision. Although referential information does not change the decision maker’s beliefs about the recommended decision, it creates leverage by changing beliefs about alternative ones. Deployed strategically, referential information can render a recommendation relatively more appealing to the receiver which, in turn, allows recommendations more favorable to the expert to be supported in equilibrium.

The influence of referential information is not that the expert reveals bad outcomes when the realized states of the world are unfavorable to the receiver and stays quiet otherwise (which would fall to standard adverse selection arguments). Referential information works through a different channel. In our model, by construction, an ideal decision for the receiver almost surely exists. Therefore, no amount of bad information will convince him otherwise. The decision maker’s core problem, however, is that he does not know which decision is his ideal. By strategically providing referential information, and by exploiting the correlation across states, the expert is able to manipulate, or spread out, the decision maker’s uncertainty such that no single decision is particularly attractive. By combining this ability with a recommendation that is also strategically chosen, the expert is able to present a picture of the world that guides the receiver to her recommendation, a recommendation that is more favorable to her, and that the receiver would not follow were it presented alone and without referential information. The power of this result is that the logic holds with probability one. Regardless of the realized state of the world, the expert is able to induce a more favorable decision with referential advice than with a recommendation alone.

The view of expertise we present here resonates with classic accounts of expertise in practice. In his famous treatise on the political power of bureaucratic experts, Weber (1958) emphasized the knowledge gap between an expert and the policymaker, observing that “the ‘political master’ finds himself in the position of the ‘dilettante’ who stands opposite the ‘expert.’” Weber argued that this gap is the source of the expert’s power—as it is in our model—and that the strategic provision of referential information is what enables this power to manifest: “A bureau’s influence rests...as Weber noted, [in] its control of information, its ability to manipulate...information about the consequences of different alternatives.” (Bendor, Taylor, and Van Gaalen 1985, p. 1042).

Related Literature

Experts play a role in almost every aspect of economic, social, and political life. Indeed, Weber (1958) concluded that in politics the power of experts was preeminent: “Under normal conditions, the power position of a fully developed bureaucracy is always overtowering.” Documenting this advantage empirically, however, can be challenging. Nevertheless, over time, broad and compelling evidence has accumulated that experts not only influence decisions but that they shape them to their personal advantage: Division managers manipulate headquarters into funding too many projects (Milgrom and Roberts 1988); realtors manipulate homeowners into selling too quickly and cheaply (Levitt and Syverson 2008); and OBGYNs manipulate patients into having too many C-sections (Gruber and Owings 1996); among other evidence. The contribution of our model is to provide a novel theoretical foundation for this expert advantage even when the decision maker knows the expert does not have his interests at heart.

The core departure of our model from the literature is, as noted, the notion of expertise we introduce and the form of advice it gives rise to. Formally, we analyze a decision space that is finite but we allow that space to become arbitrarily large such that in the limit it approximates the real line. The realization of the state for each decision corresponds, then, to a mapping from decisions to outcomes that is the realized path of a discrete stochastic process. We construct the correlation across states such that in the limit this path approximates a Brownian motion. This construction offers advantages in richness, tractability, and realism, and we leverage these to provide insights into strategic communication.² To see the generalization that this represents, it

²The Brownian motion representation of uncertainty has also recently found application

is again helpful to contrast it to expertise in its simplest form as a single piece of information (Milgrom 1981; Crawford and Sobel 1982). This corresponds to the special case of our model in which the correlation across states is perfect. Graphically, this is a linear function of known slope and expertise is knowing the value of the intercept. In our setting, in contrast, both players know the intercept (what we will refer to as the default option) but, with a much richer family of possible functions in the Brownian paths, this knowledge does not translate into complete expertise.

We are not the first paper to enrich the informational structure of the canonical model, although the focus and intention of those papers is very distant from ours, and we are the first to identify a role for referential information. The expert's knowledge is generalized to multiple dimensions in Glazer and Rubinstein (2004), Shin (2003), and Dziuda (2011). More recently, Hart, Kremer, and Perry (2017) generalize further and assume only that knowledge satisfies a partial order (see also Ben-Porath, Dekel, and Lipman (2017) and Rappoport (2017)). In these settings, the information that is strategically withheld from the receiver is decision relevant and, in equilibrium, the receiver is uncertain of the outcome he will receive from his decision. In equilibrium in our model, in contrast, the receiver is certain of the outcome he will receive from his decision as all unrevealed information is irrelevant to that decision. This provides a sharp distinction in interpretation. In our setting, the extra information provided is purely referential and aimed at dissuasion (and persuasion indirectly, as explained above), whereas in these other models, the conative and referential functions of language are intermingled and all information that is provided constitutes part of the recommendation and is aimed at persuasion directly.

A further distinction is that the literature typically focuses on the receiver-optimal equilibrium, whereas we are interested in equilibria that serve the expert's interests.³ The receiver-optimal equilibrium in our setting is trivial and is purely conative (the expert reveals a decision that produces the receiver's ideal outcome). The expert-optimal equilibrium, however, proves elusive due to the richness of our information structure. Instead, our approach is to establish a dominance result for referential advice. Our main contribution is to identify a referential equilibrium in which the expert performs strictly better almost surely than she does in *all* conative equilibria. Thus, while we do not know if this equilibrium is expert-optimal, we do know that, whatever is, it must

in models of search (Callander 2011; Garfagnini and Strulovici 2016).

³One strand of this literature (Glazer and Rubinstein 2004; Hart, Kremer, and Perry 2017; see also Sher 2011) adopts a mechanism design approach and explores the role of commitment.

involve referential advice.

A separate, prominent strand of the hard information literature, due to Dye (1985), incorporates the possibility that the sender is uninformed and, thus, not an expert (see also Dziuda 2011). This implies that should a receiver not receive some information, he is unsure whether the expert is deliberately withholding the information or whether she doesn't have it at all. This concern is not present in our model. Throughout our analysis, the expert is informed and the receiver knows this with certainty.

The alternative approach to strategic advice is, of course, cheap talk communication (Crawford and Sobel 1982). The analysis with cheap-talk in our environment is trivial and immediate: no informative equilibria exist. If different messages induce different decisions by the receiver then, because the expert cares only about the action taken, it follows that at least one message must be strictly suboptimal. Callander (2008) shows that informative communication is possible if the expert cares instead about the outcome. Analyzing the limit case of our model in which the mapping is a Brownian motion, he shows that this creates a common interest—both expert and receiver wish to avoid extreme outcomes—and that this common interest supports informative advice. Nevertheless, in the equilibrium he identifies, referential advice plays no role and communication is purely conative: in equilibrium, the expert recommends an action that maps into her own ideal point and, as long as the expert's preferences are not too different from his, the receiver implements that decision. The expert has no incentive to deviate as she obtains her ideal outcome, and the receiver implements the recommendation as he prefers the expert's ideal outcome than face the risk of choosing a decision on his own. This balance is not relevant for the preferences we analyze here and the equilibria we identify are logically distinct.

Our model is also distinct from the flourishing literature on information design (Kamenica and Gentzkow 2011; Rayo and Segal 2010; Brocas and Carrillo 2007). Our core difference with that literature is an absence of commitment. In our model, neither the receiver nor the expert can commit to any particular course of action. Similarly, communication in our model is without institutional constraint. In political economy, the influential model of Gilligan and Krehbiel (1987) demonstrates how legislatures can organize themselves by committing to formal institutional structure and rules that incentivize and leverage expertise in policymaking. Our model, instead, contributes to our understanding of how and why experts can wield power even in absence of commitment or institutional structure.

2 Model

A sender and a receiver play a game of strategic communication that is described by the following assumptions.

Technology: There is a set of decisions $\mathcal{D} = \{d_0, d_1, \dots, d_n\}$, where $d_0 = 0$ and

$$d_i = d_{i-1} + \frac{1}{\sqrt{n}} \quad \text{for } i = 1, 2, \dots, n.$$

The $n+1$ decisions, therefore, span the interval $[0, \sqrt{n}]$, with each being equally far from its neighbors.

Decision $d \in \mathcal{D}$ generates outcome $x \in \mathbf{R}$ according to outcome function $x(d)$. The outcome function is the realization of a random walk with drift. Specifically, $x(d_0) = 0$ and

$$x(d_i) = x(d_{i-1}) + \frac{\mu}{\sqrt{n}} + \sqrt{\frac{\sigma^2}{\sqrt{n}}} \theta_i \quad \text{for } i = 1, \dots, n,$$

where each θ_i is independently drawn from the standard normal distribution, $\mu > 0$ measures the expected rate of change from one decision to the next, and $\sigma > 0$ scales the variance of each decision relative to its neighbors. Our focus below is on the limit as n goes to infinity, in which case the set of feasible decisions becomes the nonnegative half line $[0, \infty)$ and the outcome function becomes the realization of a Brownian motion with drift μ and scale σ .

We denote the state of the world by $\theta = (\theta_1, \dots, \theta_n)$ and the set of possible states by $\Theta \equiv \mathbf{R}^n$. Since d_0 is the only decision whose outcome is fixed, we refer to it as the *default decision* and its outcome $x(d_0) = 0$ as the *default outcome*.

Preferences: The receiver's utility is given by $u_R(x) = -(x - b)^2$, where b is a strictly positive parameter. The sender's utility is given by $u_S(d) = -d$. The incentive conflict between the sender and the receiver is, therefore, extreme: while the receiver only cares about the outcome, the sender only cares about the decision. Note also that the assumption that the sender prefers smaller rather than larger decisions is not just a normalization. We discuss this issue and explore the opposite case in which she prefers larger decisions in Section 5.

Information: The sender is an expert. She is privately informed about states θ , and, thus, outcomes $x(d_i)$, for $i = 1, \dots, n$. All other information is public,

including the outcome of the default decision d_0 and the parameters of the process that generates the outcome function.

Communication: For each decision, the sender has a piece of hard information that verifies its outcome. The sender can hide or reveal any number of these pieces of information but she cannot fake them. Her message is, therefore, given by $m = (m_0, m_1, \dots, m_n)$, where $m_i \in \{x(d_i), \emptyset\}$.

Timing: First, nature draws the outcome function and the sender learns the realization. Second, the sender sends her message. Third, the receiver updates his beliefs and makes a decision. Finally, the sender and the receiver realize their payoffs and the game ends. Note that neither party has commitment power: the sender cannot commit to a message rule and the receiver cannot commit to a decision rule.

Solution Concept: The solution concept is perfect Bayesian equilibrium. A strategy for the sender is a mapping M from the set of all possible states Θ to the set of all possible messages \mathcal{M} . A strategy for the receiver is a mapping D from the set of all possible messages \mathcal{M} to the set of all possible decisions \mathcal{D} . The receiver's beliefs are described by a belief mapping B that assigns belief $B(m)$ —a probability distribution over states, $B(m) \in \Delta(\Theta)$ —to every possible message $m \in \mathcal{M}$. Strategies M and D and belief mapping B form a perfect Bayesian equilibrium if (i.) the sender's strategy M maximizes her utility given D , (ii.) the receiver's strategy D maximizes his expected utility given B , and (iii.) given M , on path, the receiver's beliefs satisfy Bayes' rule whenever possible and, off path, they are consistent with any hard information that has been revealed.

Off-Path Beliefs: We follow the literature on hard information and characterize equilibria in which the receiver's off-path beliefs are skeptical. To define *skeptical beliefs*, suppose the sender deviates in a way the receiver can detect and let d^* denote the best decision for the receiver among all the decisions he is informed about. We say that the receiver's off-path beliefs are *skeptical* if the following holds: (i.) if the sender did not reveal all decisions to the right of d^* , the receiver believes that the largest unrevealed decision is best for him and (ii.) if the sender did reveal all decisions to the right of d^* , the receiver believes that d^* is the best decision for him. These beliefs are akin to the “skeptical posture” in Milgrom and Roberts (1986).

Parameter Restriction: We assume that the scale parameter σ is sufficiently small, in the following sense:

$$b - \frac{\sigma^2}{2\mu} \geq \frac{\mu}{2\sqrt{n}}. \quad (1)$$

If this assumption did not hold, the sender's problem would be trivial: she could induce the receiver to make her preferred decision d_0 by simply not revealing any information. As $n \rightarrow \infty$, and the number of decisions becomes large, this condition approaches $b - \sigma^2/(2\mu) > 0$.

Remarks: The modeling choices we just described reflect our goal to incorporate our notion of expertise as knowledge about the realization of a stochastic process into a model of strategic communication that is otherwise as standard and familiar as possible. The most standard model of strategic communication is the linear-quadratic model that goes back to Crawford and Sobel (1982). Apart from the outcome function, our model is essentially the same as the version of the linear-quadratic model that is commonly used to illustrate communication with hard information (Meyer 2017; Gibbons, Matouschek, and Roberts 2013), in which the sender's preferences are assumed to be a linear function of decision rather than a quadratic function of the outcome.

The only other difference is that we work with a discrete rather than continuous decision space. We do so for technical reasons. Skeptical beliefs discipline expert behavior in equilibrium by inducing the receiver to believe the expert is the “worst” type. Yet with the Brownian motion, the set of equilibrium decisions can be unbounded if the space is unbounded, and thus no matter how the receiver responds following a deviation, there is always an expert who prefers this punishment to the decision she would otherwise receive in equilibrium. The discrete decision space alleviates this problem. Nevertheless, to stay as close as possible to the standard example, and avoid the standard integer problems, we focus throughout on the limit in which the space of decisions approaches the real line.

The message technology captures the notion that the sender can hide information but cannot fake it. This technology is simple, natural, and avoids distractions. It does, however, limit the type of messages the sender can communicate. For instance, it does not allow the sender to communicate that the outcome of a decision is in some range. Ruling out such vague messages does not drive our results. In Appendix A, we show that we can expand the message technology so that, for each decision, the sender can either say nothing or send any message that includes the true outcome, without changing our

results.

Finally, a note on the style of our presentation. To capture the simple intuition of referential advice most clearly, we are deliberately informal in the descriptions of strategies, beliefs, and results. Everything, of course, can be stated precisely and we do so in the formal statements of our results and the appendices.

3 Nonstrategic Advice

We start by exploring a simplified version of the game in which the sender acts nonstrategically and reveals the outcome of just one, arbitrarily chosen decision $d' \in \mathcal{D}$. After observing the outcome of d' , the receiver updates his beliefs about the outcomes of all other decisions. We refer to these beliefs as *neutral*, to distinguish them from other beliefs the receiver may form when the sender acts strategically. We characterize the receiver's neutral beliefs in our first lemma.

Lemma 1 *Suppose nature reveals the outcome $x(d')$ of one, arbitrarily chosen decision $d' \in \mathcal{D}$. The receiver then forms the following “neutral” beliefs:*

- (i.) *For any decision $d \geq d'$, he believes that outcome $x(d)$ is normally distributed with mean and variance*

$$E[x(d) | x(d')] = x(d') + (d - d')\mu \quad \text{and} \quad \text{Var}[x(d) | x(d')] = (d - d')\sigma^2.$$

- (ii.) *For any decision $d \leq d'$, he believes that outcome $x(d)$ is normally distributed with mean and variance*

$$E[x(d) | x(d')] = \frac{d}{d'}x(d') \quad \text{and} \quad \text{Var}[x(d) | x(d')] = \frac{d(d' - d)}{d'}\sigma^2.$$

The lemma is illustrated in Figure 1. In all our figures, red indicates the outcomes of known decisions, in this case the default decision d_0 and the revealed decision d' . Figure 1 also indicates the expected outcomes of all other decisions, by the dashed blue lines, and their variances, by the vertical distance between the dashed green curves. Since drift μ is positive, the receiver expects the outcome $x(d)$ of any decision $d > d'$ to be larger, the larger the decision is. For any $d < d'$, instead, the receiver expects outcome $x(d)$ to be a convex combination of the two known outcomes $x(0)$ and $x(d')$. In either case, the receiver is more uncertain about the outcome of a decision, the further it is from the closest known decision d_0 or d' .

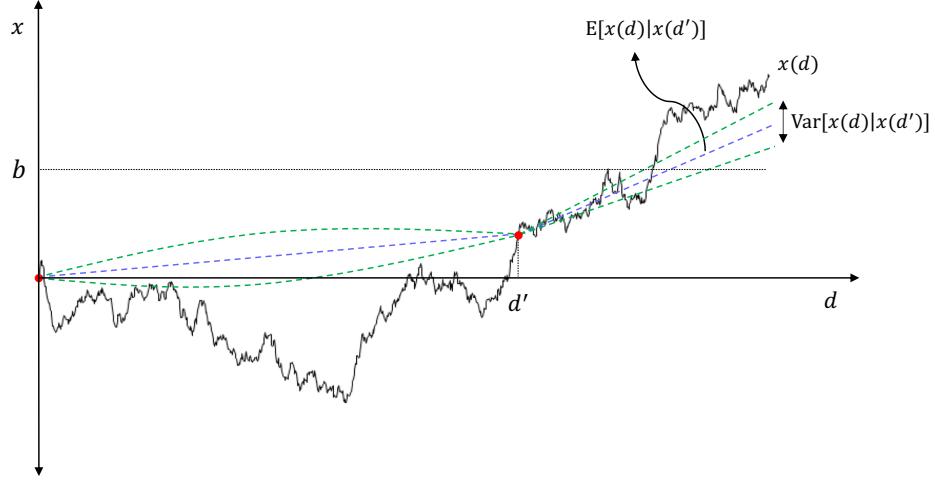


Figure 1

After updating his beliefs, the receiver makes the decision that maximizes his expected utility

$$E[u_R(x(d)) \mid x(d')] = -E[(x(d) - b)^2 \mid x(d')].$$

Rewriting expected utility in its mean-variance form

$$E[u_R(x(d)) \mid x(d')] = - (E[x(d) \mid x(d')] - b)^2 - \text{Var}[x(d) \mid x(d')] \quad (2)$$

highlights the receiver's two, potentially conflicting, goals: he wants an expected outcome that is as close as possible to b but he also wants to avoid the risk that comes with making a decision he is not perfectly informed about.

To streamline the discussion of how the receiver best pursues these goals, we focus on the case in which $x(d') \leq b$. The alternative case in which $x(d') > b$ is similar and not relevant for our discussion of strategic advice below.

Suppose first the receiver has to choose between d' and one of the decisions to its right. Making a decision to the right of d' exposes him to risk but may generate a better expected outcome. The receiver's willingness to take on more risk in return for a better expected outcome depends on how far the outcome of d' is from his ideal one: if $x(d')$ is far from b , the benefit of a better expected outcome is large and outweighs the cost of more risk, but if $x(d')$ is already close to b , the opposite holds. A few steps of standard algebra show that the critical outcome level at which the receiver is just indifferent between

a marginally better expected outcome and the risk that comes with making a marginally larger decision is given by

$$\bar{x} \equiv b - \frac{\sigma^2}{2\mu} > 0, \quad (3)$$

where the inequality follows from assumption (1). For reasons that will become apparent, we refer to \bar{x} as the *upper threshold*. If $x(d')$ is closer to the upper threshold than the expected outcome of any decision to its right, the receiver settles for $x(d')$ and makes decision d' . If, instead, $x(d')$ is further from the upper threshold than the expected outcome of some of the decisions to its right, the benefit of a better expected outcome is worth the additional risk. In this case, the receiver makes the decision $d > d'$ whose expected outcome comes closest to \bar{x} . In summary, the receiver's preferred decision to the right of d' is

$$d_r(d') \equiv \arg \min_{d \in \{d', \dots, d_n\}} |E[x(d) | x(d')] - \bar{x}|, \quad (4)$$

where the expected outcome $E[x(d) | x(d')]$ and the upper threshold \bar{x} are given by (2) and (3) and where the subscript r stands for "to the right of d' ." If there are multiple minimizers, $d_r(d')$ is the smallest one.

The receiver's preferred decision to the right of d' is best illustrated in the limit in which the number of decisions goes to infinity. In this case, the receiver makes decision d' if $x(d')$ is at or above the upper threshold. If, instead, $x(d')$ is strictly below the upper threshold, he makes the decision to the right of d' that generates the expected outcome \bar{x} exactly. Figures 2a and 2b provide illustrations of each case. Formally, in the limit, the receiver's preferred decision to the right of d' is given by $d_r(d') = d'$ if $x(d') \in [\bar{x}, b]$ and $d_r(d') = d' + (\bar{x} - x(d'))/\mu > d'$ if $x(d') < \bar{x}$.

The receiver, of course, does not have to choose a decision to the right of d' and can, instead, choose one to its left. The only such decision the receiver may be tempted by, however, is the default decision d_0 . He will never make a decision strictly between d_0 and d' since doing so would expose him to risk without generating an expected outcome that is any better than what he can get by making either decision d_0 or d' . For the sender to be tempted by a decision to the left of d' , the outcome of the default decision, therefore, has to be better than the outcome of the revealed decision, as in the example in Figure 3. If this is the case, the receiver prefers d_0 to d' and to all the decisions between them.

The receiver may prefer the default decision, not just to d' and to the decisions between them, but to *any* decision. And he may do so no matter

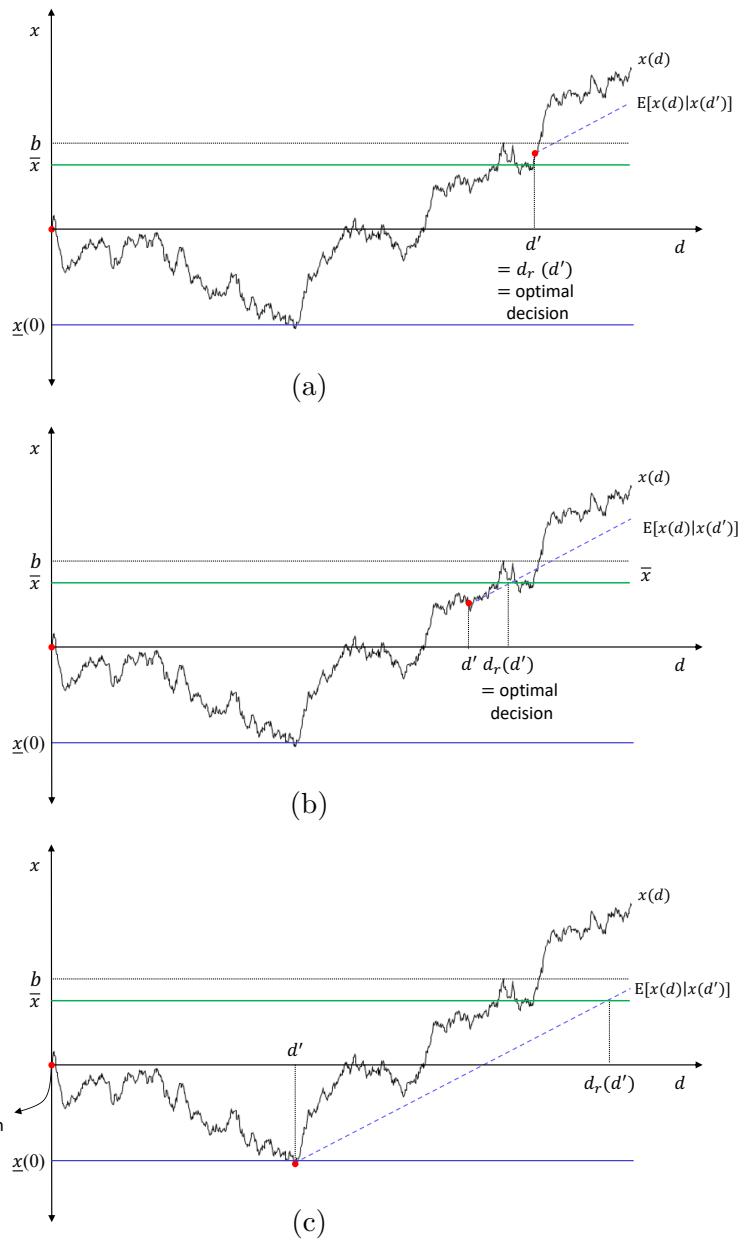


Figure 2

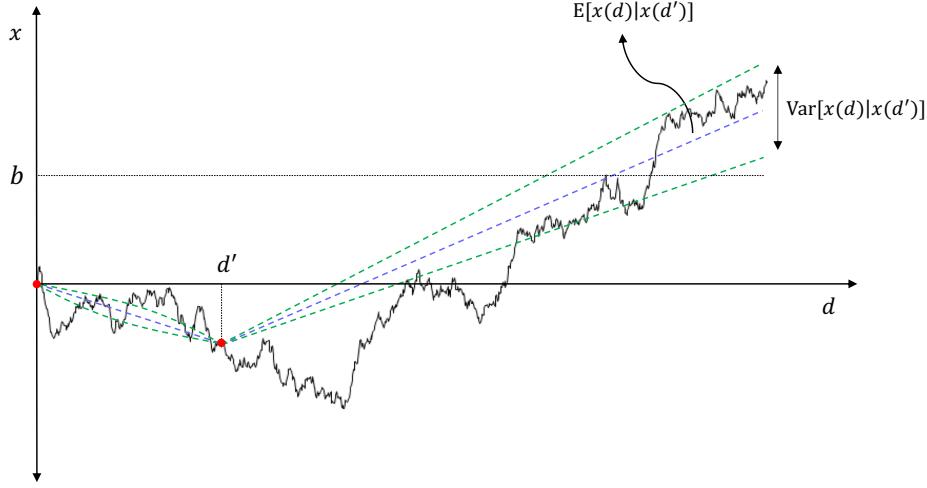


Figure 3

how far its outcome is from his ideal one. To see why, note that the receiver's expected utility from making decision $d_r(d')$ depends on the outcome of the revealed decision d' , even when $d_r(d')$ is different from d' . No matter how far $x(d')$ is below the threshold \bar{x} , decision $d_r(d')$ always aims for the same expected outcome. The further $x(d')$ falls below \bar{x} , though, the more risk the receiver has to take on to obtain the desired expected outcome. For sufficiently low values of $x(d')$, decision $d_r(d')$ becomes too risky and the receiver is better off forgoing an expected outcome close to his ideal one for the safety of the default decision. We refer to the critical value of $x(d')$ at which the receiver is just indifferent between $d_r(d')$ and d_0 as the *lower threshold* $\underline{x}(0)$, where the '0' indicates the default outcome. Formally, the lower threshold is given by the value of $\underline{x}(0) \leq 0$ such that

$$E [u_R(x(d_r(d')) \mid x(d') = \underline{x}(0))] = -b^2, \quad (5)$$

where the right-hand side is the utility from the default decision. The next proposition summarizes the discussion.

Proposition 1 *Suppose the sender acts nonstrategically and reveals the outcome $x(d')$ of one, arbitrarily chosen decision $d' \in \mathcal{D}$. The receiver's optimal decision is given by*

$$d_0 \quad \text{if} \quad x(d') \in (-\infty, \underline{x}(0)]$$

and

$$d_r(d') \quad \text{if } x(d') \in [\underline{x}(0), b],$$

where $d_r(d')$ and $\underline{x}(0)$ are defined in (4) and (5).

In the limit in which the number of decisions goes to infinity, the lower threshold is given by

$$\underline{x}(0) = -\frac{\mu}{\sigma^2} \bar{x}^2. \quad (6)$$

Applied to this case, the proposition says that the receiver makes the revealed decision d' if $x(d') \in [\bar{x}, b]$, as in the case illustrated in Figure 2a. If $x(d') \in [\underline{x}(0), \bar{x}]$, the receiver makes decision $d_r(d') = d' + (\bar{x} - x(d'))/\mu$, as illustrated in Figure 2b. Finally, if $x(d') < \underline{x}(0)$, the receiver makes the default decision, as illustrated in Figure 2c.

To conclude this section, we use the proposition to establish a corollary that deals with the natural benchmark in which the receiver has to make a decision without being able to consult the sender. This benchmark corresponds to the case in which $d' = d_0$.

Corollary 1 *Suppose the receiver has to make a decision without being able to consult the sender. In this no-advice benchmark, the receiver makes decision $d_r(d_0)$. As the number of decisions goes to infinity, this decision is given by*

$$d_r(d_0) = \frac{\bar{x}}{\mu} > 0,$$

where the sign follows from assumption (1).

In the absence of any advice, the receiver, therefore, would not settle for the default decision and, instead, make an uncertain decision to its right. The size of this decision will provide the benchmark for how strategic advice can—or cannot—move outcomes in favor of the sender.

4 Strategic Advice

We can now build on our understanding of nonstrategic advice to explore the parties' incentives to provide, and respond to, strategic advice. We start by showing that a strategic sender will always reveal at least some information. Never providing any advice cannot be an equilibrium.

Proposition 2 *In any equilibrium, the sender reveals at least some information.*

To see why this claim holds, suppose to the contrary that the sender's strategy is to never reveal any information. We know from the previous section that the receiver's best response is to make decision $d_r(d_0) > d_0$. Given this decision rule, however, the sender will sometimes find it optimal to deviate and induce a decision that is better for both parties. For some realizations of the outcome function, for instance, there is a single decision $d < d_r(d_0)$ that generates the receiver's ideal outcome b exactly, in which case the sender can induce the receiver to make this smaller decision by revealing the outcome of all decisions. The question, therefore, is not *if* the sender will give advice but *how* she will do so.

4.1 Conative Advice

A simple and common form of advice is a recommendation or instruction of what to do. Following Jakobson (1960), we refer to such advice as “conative” advice. Formally, we say that an equilibrium is *conative* if, as the number of decisions goes to infinity, the sender almost surely reveals the outcome of only one decision and the receiver makes the revealed decision.

We start by constructing a particular conative equilibrium. In this equilibrium, the sender's strategy is to reveal the smallest decision whose outcome is sufficiently good for the receiver for him to be willing to make that decision. Specifically, the sender reveals the smallest decision d_C whose outcome is in

$$\left[b - \frac{\sigma^2}{2\mu}, b + \frac{\sigma^2}{2\mu} \right], \quad (7)$$

where the subscript “C” stands for “conative.” Note that this band is symmetric around the receiver's ideal outcome b and that the smallest outcome in the band is equal to \bar{x} , the upper threshold we defined in (3). If there is no decision whose outcome is in the band, the sender reveals the outcomes of all decisions. The receiver's strategy, in turn, is to make the best revealed decision, unless he detects a deviation by the sender, in which case he acts in accordance with his skeptical beliefs.

Figure 4 illustrates the equilibrium in the limit in which the number of decisions goes to infinity. In this case, the outcome function approaches a continuous function that almost always generates any positive outcome, including \bar{x} . The sender reveals the decision d_C at which the outcome function first reaches \bar{x} and the receiver makes the revealed decision.

The fact that d_C is the *smallest* decision that generates \bar{x} is crucial for why the above strategies form an equilibrium. One implication is that the



Figure 4

outcomes of decisions to the left of d_C are further from the receiver's ideal outcome b than $x(d_C)$ is. The receiver, therefore, prefers d_C to any decision to its left. Another implication is that the sender's message reveals no additional information about decisions to the right of d_C than if the message had been sent nonstrategically. For decisions $d \geq d_C$ the receiver's beliefs are, therefore, neutral, as in the nonstrategic benchmark and as illustrated in Figure 4. It then follows from our discussion of nonstrategic advice that if $x(d_C)$ is in the bottom half of the band (7), d_C is "good enough" and getting an expected outcome that is even closer to b is not worth the additional risk. The case for d_C is even clearer if $x(d_C)$ is in the top half of the band (7), as might be the case when the number of decisions is limited. Decisions to the right of d_C then not only expose the receiver to more risk but also generate a worse outcome on average. If the sender recommends d_C , the sender's best response, therefore, is to rubber-stamp it.

Turn now to the sender's strategy. The sender can always deviate in a way the receiver can detect, for instance, by revealing the outcomes of more than one decision. Such a deviation, however, would induce the receiver to form skeptical beliefs, causing him to act in ways that are worse for the sender. Alternatively, the sender may be able to deviate in a way the receiver cannot detect. The only way to do so, however, is to reveal the outcome of a decision that is larger than d_C but also generates \bar{x} . Since this deviation would induce the receiver to make a larger decision than he otherwise would, the sender has

no incentive to engage in it.

The strategies we described above, therefore, support an equilibrium. Moreover, the equilibrium is conative: as the number of decisions goes to infinity, the sender almost surely recommends a single decision and the receiver finds it optimal to rubber-stamp the recommendation. The next proposition summarizes our discussion so far.

Proposition 3 *There exists a conative equilibrium that, as the number of decisions goes to infinity, almost always implements outcome \bar{x} .*

There are other conative equilibria than the one we have focused on so far. To see this, suppose we augment the above strategies by making the band (7) more narrow, while keeping it symmetric around b . In the limit in which the number of decisions goes to infinity, the sender will once again recommend the smallest decision that generates the outcome at the bottom of the band, which is now even closer to b than \bar{x} is. Since the receiver is willing to rubber-stamp the sender's recommendations when they generate outcome \bar{x} , he is certainly willing to do so when they generate outcomes that are even better for him. It is, therefore, not hard to construct conative equilibria that almost surely implement any outcome $x_C \in [\bar{x}, b]$. The next proposition shows that while this is true, the reverse also holds: it is not possible to construct conative equilibria that, with positive probability, implement outcomes that are outside of $[\bar{x}, b]$.

Proposition 4 *In every conative equilibrium, as the number of decisions goes to infinity, the outcome that is implemented is almost surely in $[\bar{x}, b]$.*

To get an intuition for this proposition, suppose now that, instead of narrowing the band (7), we widen it, while still keeping it centered around b . Suppose further that there is a decision that generates the outcome at the bottom of the band, as will almost always be the case in the limit. One problem that now arises is that, even if the sender sticks to her strategy and reveals the decision at which the outcome function first intersects with the lower bound, the receiver is no longer willing to rubber-stamp it. Since the outcome of the recommended decision is now so far from the receiver's ideal one, he is better off making the decision to its right that generates \bar{x} on average.

Moreover, given this response, the sender will sometimes deviate from her strategy. Suppose, for instance, that there is a single decision that generates the receiver's ideal outcome and that this decision is between the decision the sender is supposed to recommend and the one the receiver will make if she does so. The sender can then deviate in a way that makes both parties better off, for instance by revealing the outcomes of all decisions. Doing so would

induce the receiver to make the decision that is ideal for him which, since it is smaller than the decision he would otherwise have made, is also better for the sender. If the band is wider than in (7), the above strategies, therefore, do not form an equilibrium, let alone a conative one. The proposition shows that there are also no other strategies that form a conative equilibrium and implement outcomes outside of $[\bar{x}, b]$.

The fact that there are no conative equilibria that implement outcomes outside of $[\bar{x}, b]$ implies that, from the receiver's perspective, the worst conative equilibrium is the one that almost surely implements \bar{x} . Even in this equilibrium, however, the receiver is better off than without advice: he gets \bar{x} for sure while, in the no-advice benchmark, he gets \bar{x} only on average. At a minimum, therefore, conative advice protects the receiver from risk. And since there are other conative equilibria that implement outcomes closer to b , conative advice can also allow the receiver to realize better outcomes on average.

While the conative equilibrium that implements \bar{x} is the worst one for the receiver, it is the best conative equilibrium for the sender. In the sender's case, however, the comparison with the no-advice benchmark is less clear. Sometimes the outcome function will cross \bar{x} early and the receiver will make a smaller decision than he would have without advice. Other times, however, the outcome function will cross \bar{x} later, leaving the sender worse off than if she had not given any advice. This raises the question whether, overall, conative advice makes the sender better off. In the next proposition we answer this question in the negative.

Proposition 5 *As the number of decisions goes to infinity, the sender performs no better in her preferred conative equilibrium than in the no-advice benchmark.*

We saw above that there are conative equilibria that almost always implement any $x_C \in [\bar{x}, b]$. In the proof of this proposition we show that in any such equilibrium, the receiver's average decision is given by x_C/μ . In the sender's preferred conative equilibrium $x_C = \bar{x}$, in which case the average decision is \bar{x}/μ . This is the same as the decision $d_r(d_0)$ the receiver makes on his own. In any other conative equilibrium, the outcome is closer to the receiver's ideal outcome and the average decision is larger. In the receiver's preferred conative equilibrium, for instance, $x_C = b$ and the average decision is b/μ , which, as illustrated in Figure 5, is strictly larger than $d_r(d_0)$.

Fundamentally, the challenge for the sender is that advice does not only reveal information about the recommended decision. Since outcomes are correlated, the recommendation also reveals information about all other decisions. As a result, by revealing a decision that is more appealing to the receiver, the



Figure 5

sender is helping the receiver see where other good decisions might lie, and this makes it less risky for the receiver to take his chances with one of them.

This logical process iterates—the sender reveals a decent outcome, which emboldens the receiver, so the sender must reveal an even better outcome, and so on—until the sender has revealed an outcome sufficiently close to the receiver’s ideal that he is satisfied. In this way, it is the receiver who extracts all of the benefit from conative advice, leaving the sender no better off than if she hadn’t communicated at all.

To do better, the sender has to induce the receiver to follow her recommendation even if its outcome is quite far from his preferred one. The only way she can do so is to convince him that his alternatives are even worse. In the next section we show that the sender can always induce such a pessimistic outlook by complementing her recommendation with information about an appropriately chosen set of alternatives. This referential information provides the context for her recommendation and convinces the receiver that even though the recommendation may not be very good, it still beats his alternatives.

4.2 Referential Advice

We now go beyond simple, conative advice and allow the sender to complement her recommendation with referential information. Specifically, we say that an equilibrium is *referential* if, as the number of decisions goes to infinity,

the sender almost surely reveals the outcomes of more than one decision and the receiver makes one of the revealed decisions.

As suggested at the end of the last section, the purpose of providing context for the recommendation is to cast doubt on the alternatives and convince the receiver that even though the recommendation may not be very good, his alternatives are even worse. The difficulty in doing so is that the receiver knows there almost surely is a decision that generates his ideal outcome, and many that come close. Below we explore how the sender can overcome this challenge by complementing her recommendation with information about similar decisions. To set the stage for this *interval equilibrium*, we first revisit and adapt our model of nonstrategic advice.

Nonstrategic Advice, Revisited: In the model of nonstrategic advice we explored above, the receiver observes the outcome of one, arbitrarily chosen decision d' . Suppose now that, in addition to observing the outcome of d' , the receiver also learns the outcomes of all decisions $d < d'$. In line with our previous discussion of nonstrategic advice, we focus on the case in which $x(d') \leq b$.

Learning the outcomes of decisions $d < d'$ does not reveal any information about decisions $d > d'$ beyond what the receiver already knows from observing $x(d')$. The receiver's preferred decision to the right of d' , therefore, is still given by $d_r(d')$, the decision $d \geq d'$ whose expected outcome comes closest to the upper threshold \bar{x} .

What learning the outcomes of decisions $d < d'$ does allow the receiver to do, is to identify a better alternative to $d_r(d')$. He will no longer dismiss decisions between d_0 and d' since he now observes their outcomes and may well find that some of them are better than those of both d_0 or d' . The receiver's best alternative to $d_r(d')$, therefore, is now given by the decision to the left of d' that generates the best outcome. We denote this decision by

$$d_l(d') \equiv \arg \min_{d \leq d'} |x(d) - b| \quad (8)$$

and the associated outcome by

$$x_l(d') \equiv x(d_l(d')), \quad (9)$$

where the subscript l stands for “to the left of d' .” If there are multiple minimizers, $d_l(d')$ is the smallest one.

The choice between the two relevant alternatives once again depends on $x(d')$. In (5) we defined the lower threshold $\underline{x}(0)$ as the value of $x(d')$ at

which the receiver is indifferent between the decision $d_r(d')$ and the default outcome. Since the receiver now has more information and may know a better alternative, he is more willing to abandon $d_r(d')$. For the lower threshold to still be applicable, it, therefore, now has to be higher. To adapt the lower threshold, we now define it as the value of $x(d')$ at which the receiver is indifferent between the decision $d_r(d')$ and any certain outcome $x \in \mathbf{R}$. Formally, the lower threshold is now given by the value $\underline{x}(x) \leq b$ for which

$$E[u_R(x(d_r(d'))) | x(d') = \underline{x}(x)] = -(x - b)^2, \quad (10)$$

where the right-hand side is the utility of getting outcome x for sure. The next proposition uses this generalized version of the lower threshold to characterize the receiver's optimal decision.

Proposition 6 *Suppose the sender acts nonstrategically and reveals the outcomes of all decisions $d \leq d'$ for an arbitrarily chosen $d' \in \mathcal{D}$. The receiver's optimal decision is then given by*

$$d_l(d') \quad \text{if } x(d') \in (-\infty, \underline{x}(x_l(d'))]$$

and

$$d_r(d') \quad \text{if } x(d') \in [\underline{x}(x_l(d')), b],$$

where $d_r(d')$ is defined in (4) and $d_l(d')$, $x_l(d')$, and $\underline{x}(\cdot)$ are defined in (8), (9), and (10).

In the limit in which the number of decisions goes to infinity, the lower threshold is given by

$$\underline{x}(x_l(d')) = \begin{cases} x_l(d') - \frac{\mu}{\sigma^2} (\bar{x} - x_l(d'))^2 & \text{if } x_l(d') < \bar{x}, \\ x_l(d') & \text{if } x_l(d') \in [\bar{x}, b]. \end{cases} \quad (11)$$

In our previous discussion of nonstrategic advice, we used Figures 2a–2c to illustrate the receiver's optimal response when he only learns the outcome of decision d' . Figures 6a–6c revisit the same examples to illustrate his optimal response when he also learns the outcomes of all decisions $d < d'$. The key difference between Figures 2 and 6 is that the lower threshold $\underline{x}(\cdot)$ is now increasing in d' . This is also reflected in the above expression of the lower threshold, which shows that $\underline{x}(x_l(d'))$ is increasing in $x_l(d')$, which, in turn, is increasing in d' . Intuitively, the larger d' is, the more information the receiver has about alternatives to its left, and the better, thus, the best such alternative

$x_l(d')$ has to be. For the receiver to be willing to make decision $d_r(d')$ rather than settle for the safe alternative $x_l(d')$, decision $d_r(d')$ then has to be better, which requires $x(d')$ to be closer to b .

The figures illustrate another property of the lower threshold $\underline{x}(x_l(d'))$ that will play an important role below: it meets the upper threshold \bar{x} at $d' = d_C$, the smallest decision at which the outcome function reaches \bar{x} . Since d_C is the *smallest* decision at which the outcome function reaches \bar{x} , decisions to its left have to be worse, and so $x_l(d_C) = \bar{x}$. The fact that $\underline{x}(x_l(d_C)) = \bar{x}$ then follows immediately from the expression of the lower threshold in (11).

To see the intuition for why the two thresholds are the same if $d' = d_C$, note that in this case the receiver prefers d_C to all decisions to its left, since it generates the best known outcome, and to all decisions to its right, since its outcome is close enough to the ideal one. The best decisions to the left and the right of d_C are, therefore, the same, they are both given by d_C . But if $d_l(d_C)$ and $d_r(d_C)$ are the same, the receiver is obviously indifferent between them, which means that $x(d_C)$ has to be equal to the lower threshold. Since, by definition, $x(d_C)$ is also equal to the upper threshold, the two thresholds have to be the same.

Interval Equilibrium: The interval equilibrium builds on this logic. The sender's strategy is to reveal the outcomes of all decisions $\{d_R, \dots, \bar{d}\}$, where the subscript “ R ” stands for “referential.” The right endpoint \bar{d} is the smallest decision d whose outcome $x(d)$ falls below the lower threshold $\underline{x}(x_l(d))$. If no such decision exists, \bar{d} is the largest decision d_n . To get an intuition for \bar{d} , note that in the nonstrategic benchmark, \bar{d} is the smallest value of d' for which the receiver prefers the safety of the best known decision to any risky decision. The left endpoint d_R , instead, is the decision to the left of \bar{d} whose outcome is closest to the receiver's ideal outcome b , that is, $d_R = d_l(\bar{d})$. The receiver's strategy, in turn, is to make decision d_R , unless he detects a deviation by the sender, in which case he acts in accordance with his skeptical beliefs.

A feature of the interval equilibrium is that the length of the interval depends on the realization of the outcome function. The interval could be short or it could be long, the sender may reveal a lot or a little of her expertise. Regardless of the length of the interval, its left endpoint d_R is the sender's preferred revealed decision and constitutes her recommendation. All other knowledge she reveals—the outcomes of the decision to the right of d_R all the way through to \bar{d} —is referential information, its only purpose is to stir the receiver towards following the recommendation, which he ends up doing. Figure 7 illustrates the equilibrium, with the sender's revealed information marked by

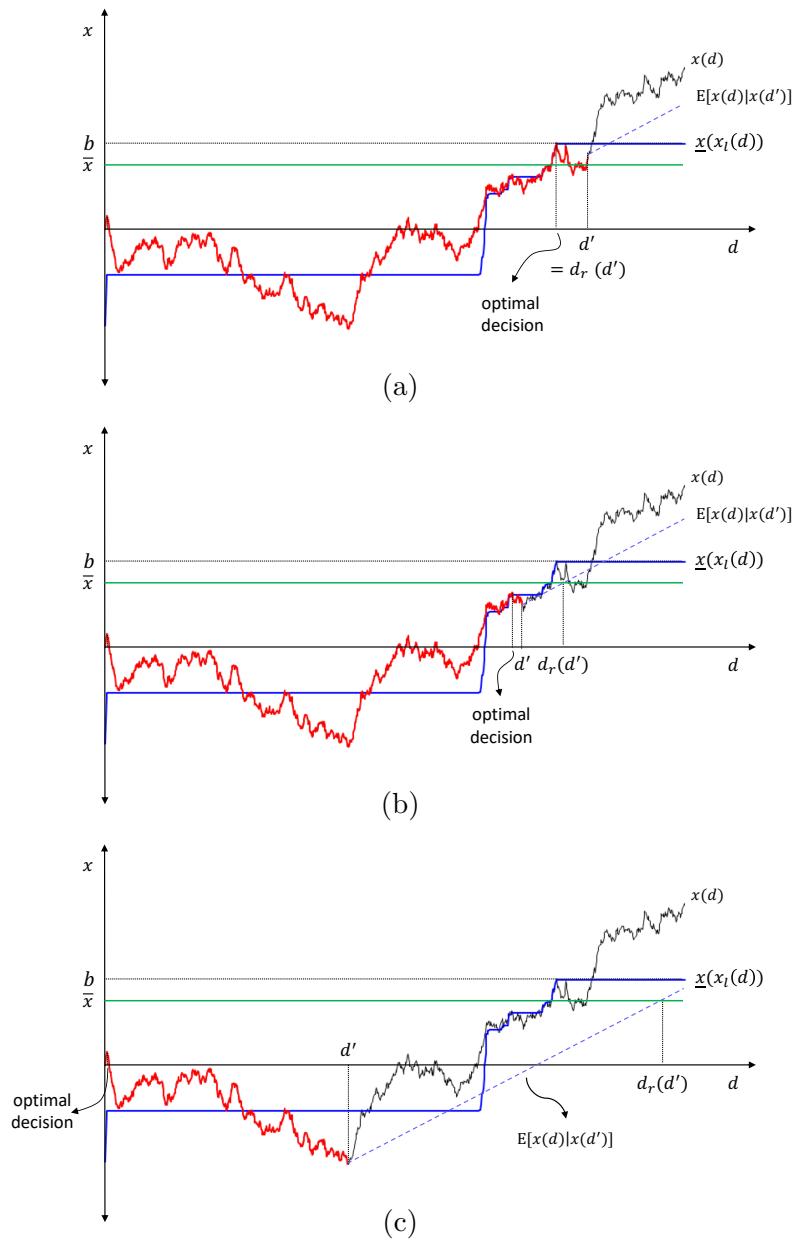


Figure 6

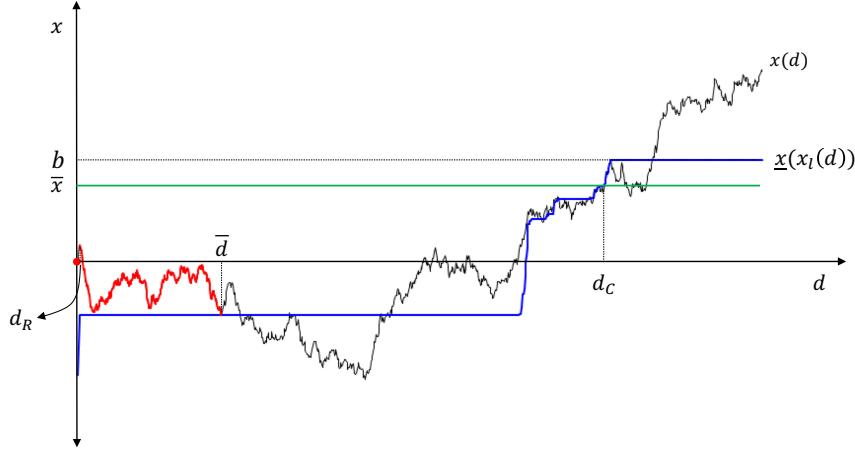


Figure 7

the red interval.

To see that these strategies support an equilibrium, consider the receiver's incentive to follow the recommendation. By construction, the receiver prefers d_R to all the other revealed decisions. As for decisions to the right of \bar{d} , the sender's strategy implies that the receiver learns no information about these decisions beyond what he would learn if the same information were communicated nonstrategically. Given the sender's neutral beliefs, it then follows from the definition of the lower threshold that the sender prefers d_R to all decisions to the right of \bar{d} .

The final possibility is that the receiver prefers a decision to the left of d_R . Since the sender does not share any information about such decisions, the receiver cannot be sure about their outcomes. Given the sender's strategy, however, he rightly infers that all of them generate outcomes that are worse for him than the recommended one.

Turn then to the sender's incentive to follow her strategy. As in the conative equilibrium, the sender can always deviate in ways the receiver can detect and may be able to deviate in ways he cannot. Once again, however, she has no incentive to do either. Any deviation the receiver cannot detect will necessarily induce him to make a larger decision. And any deviation he can detect will lead him to form skeptical beliefs and act in ways that make the sender worse off.

The strategies we described above, therefore, support an equilibrium. Moreover, the equilibrium is referential: as the number of decisions goes to infinity, the sender almost surely reveals the outcomes of more than one decision and

the receiver makes the best revealed one. The next proposition summarizes our discussion.

Proposition 7 *The interval equilibrium exists and is referential.*

The key property of the interval equilibrium is the receiver's willingness to follow recommendations he would not follow if advice were conative. In Figure 7, for instance, the outcome of the recommended decision d_R is well below the upper threshold \bar{x} . As we saw above, the receiver would not follow such a recommendation if it didn't come with additional information. The receiver's willingness to follow worse recommendations, in turn, allows the sender to induce decisions that are better for her. In Figure 7, not only does d_R generate an outcome well below \bar{x} , it is also much smaller than d_C , the decision the receiver would make in the sender's preferred conative equilibrium.

To get the receiver to follow poor recommendations, the sender convinces him that his alternatives are even worse. As we noted above, the challenge in doing so is that the receiver knows there almost always is a decision that generates his ideal outcome, and many that come close. There is, therefore, no point in trying to convince the receiver that these decisions don't exist. Instead, the sender's strategy is to provide the receiver with just enough information to prove that even though these decisions may very well exist, they are far from the ones he is informed about and thus very risky. The sender's strategy, in other words, is to cast doubt on the receiver's alternatives by creating uncertainty about his best one.

For the sender to be able to pursue this strategy successfully, the outcome function has to have the right shape: it has to, early on, exhibit a local peak that is followed by a sufficiently pronounced fall. Revealing the outcomes of all decisions from the peak down to the trough then allows the sender to move the decision in her favor. It is by now evident that this is possible. The power of the interval equilibrium is that doing so is not just a possibility, or even achievable on average, but that it is feasible for almost all realizations of the outcome function. Almost surely, the sender, by revealing a carefully chosen interval of decisions, is able to induce a decision that is better for her, and worse for the receiver, than in any conative equilibrium.

Proposition 8 *As the number of decisions goes to infinity, the decision and the outcome that are implemented in the interval equilibrium are almost surely strictly smaller, and never strictly larger, than in any conative equilibrium.*

To get an intuition for this proposition, take the sender's preferred conative equilibrium. We saw earlier that, in this equilibrium, the receiver makes the

decision at which the outcome function first meets the upper threshold \bar{x} . In the interval equilibrium, instead, he makes a decision to the left of the one at which the outcome function first meets the lower threshold $\underline{x}(x_l(d))$. Now recall from our discussion of nonstrategic advice above, that when $x(d)$ first intersects with \bar{x} , it also intersects with $\underline{x}(x_l(d))$. The receiver's decision in the interval equilibrium, therefore, can never be strictly larger than his decision in the sender's preferred conative equilibrium. To see that it will almost surely be strictly smaller, consider Figure 7 and note that as $x(d)$ approaches \bar{x} , $\underline{x}(x_l(d))$ also approaches \bar{x} . As $x(d)$ gets closer and closer to the upper threshold \bar{x} , it, therefore, only takes a small downturn in $x(d)$ for it to hit the lower threshold $\underline{x}(x_l(d))$. The main result in Proposition 8 is that such a downturn almost surely occurs.

The final issue we need to address is how much referential information the sender needs to communicate. In the interval equilibrium, the sender may have to reveal a lot of information and it may seem that she can do as well by revealing the outcomes of only two decisions, d_R and \bar{d} . If revealing the outcomes of only d_R and \bar{d} were credible, the receiver would indeed act exactly as in the interval equilibrium. The problem is that if the receiver acts as in the interval equilibrium, the sender will sometimes deviate in a way the receiver cannot detect. For some realizations of the outcome function, for instance, the sender can do so by revealing the outcomes of $d_R - \epsilon$ and of any decision to its right that generates outcome $\underline{x}(x(d_R - \epsilon))$. The purpose of complementing the recommendation d_R with information about decision \bar{d} , therefore, is to create uncertainty about the receiver's best alternative. And the purpose of complementing it with information about decisions between d_R and \bar{d} is to make this strategy credible.

Implications: If referential advice is so effective in swaying the receiver's decision, as we just saw, then may the receiver be worse off receiving referential advice than not receiving any advice at all? Since the receiver is a rational Bayesian, information has to make him better off on average, no matter how selectively it is presented to him. Even though the receiver benefits from referential advice on average, though, it may well leave him worse off most of the time. Figure 8 illustrates this point. It depicts the distribution of the benefits of advice for both the interval equilibrium and the sender's preferred conative equilibrium, for 10,000 randomly generated outcome functions. The figure shows that in more than 60% of the cases, the benefits of referential advice are negative: the receiver is worse off in the interval equilibrium than in the no-advice benchmark. The reason the receiver still benefits from advice

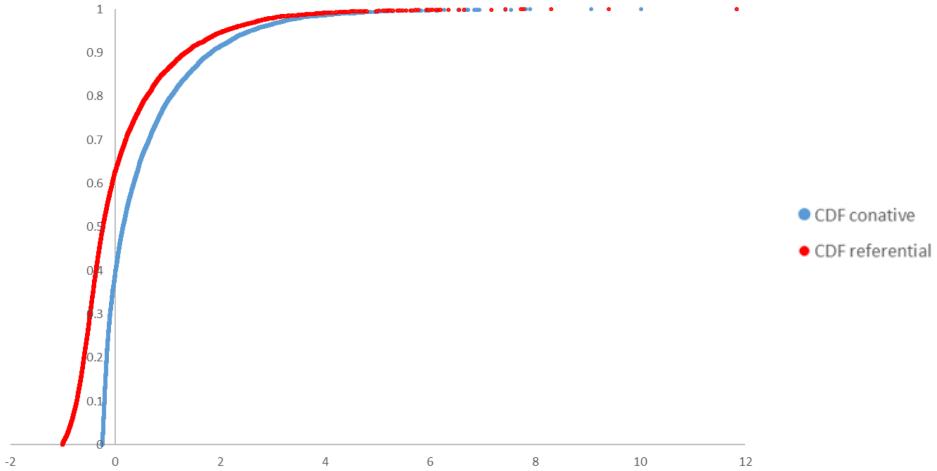


Figure 8: Distributions of numerical estimates of the difference between the receiver's utility with advice (for the sender's preferred conative equilibrium, and the interval equilibrium, respectively) and without advice, for parameter values $\mu = b = \sigma = 1$, $n = 1000$, and sample size 10,000.

on average is that in a small number of cases, it keeps him from making very bad decisions. The distribution of the benefits of advice is different if advice is conative. The receiver then benefits from advice in more than 60% of the cases. And when he doesn't benefit, his losses are much smaller than in the interval equilibrium.

The sender's ability to sway the receiver's decision depends on how she structures her advice, which is what we have focused on so far. It also, though, depends on the economic environment she operates in and, in particular, on the complexity of her informational advantage, as captured by σ^2 . To illustrate the issues, consider Figure 9, which summarizes another simulation. In this simulation, we randomly generated 10,000 outcome functions each for different values of σ^2 . On average, the decision d_R the receiver makes in the interval equilibrium is smaller, the larger σ^2 is. As one might expect, complexity helps the sender induce decisions that are more favorable to her. More complexity, however, also induces the receiver to make smaller decisions in the absence of any advice. Relative to not advising the receiver at all, the sender's ability to sway his decision in her favor, therefore, peaks when complexity is neither too high nor too low. This is illustrated in Figure 9 and shown more generally in the next proposition.

Proposition 9 *As the number of decisions goes to infinity, the difference in*

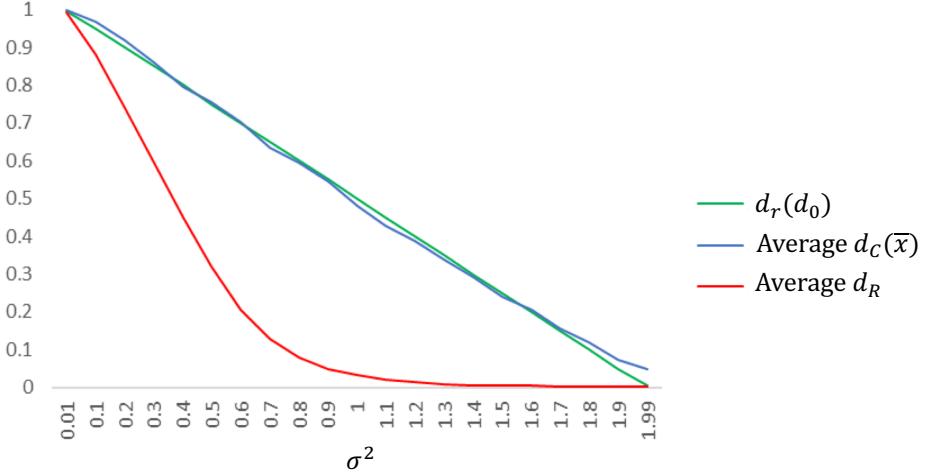


Figure 9: Numerical estimates of the receiver’s decision as a function of σ^2 , without advice, in the sender’s preferred conative equilibrium, and in the interval equilibrium, for parameter values $\mu = b = 1$, $n = 1000$, and sample size 10,000 for every value of σ^2 .

the sender’s expected utility in the interval equilibrium and in the no-advice benchmark is maximized for an intermediate degree of complexity $\sigma^2 \in (0, 2b\mu)$.

Having discussed the interval equilibrium and its implications, we conclude this section by discussing how it relates to other referential equilibria.

Other Referential Equilibria: The interval equilibrium demonstrates that referential information matters and illuminates the process through which it does. It shows that referential information allows the sender to improve on a simple recommendation and that it does so by creating uncertainty about the receiver’s best alternative. In the interval equilibrium, the sender complements the recommendation with information about similar, larger decisions. There may be other referential equilibria in which referential information takes a different form. The next proposition shows that among all referential equilibria in which the sender reveals the outcomes of a single, convex set of decisions, she cannot do better than in the interval equilibrium.

Proposition 10 *Take any equilibrium in which the sender always reveals the outcomes of consecutive decisions and, on the equilibrium path, the receiver makes one of the revealed decisions. The sender’s expected utility in any such equilibrium cannot be strictly larger than in the interval equilibrium.*

For the sender to do better than in the interval equilibrium, she would have to go beyond complementing her recommendation with information about similar decisions. Such equilibria may exist. And if they do, they strengthen our point about the power of referential advice. The simulation that we summarized in Figure 9, however, suggests that at least for sufficiently high degrees of complexity, the sender would be hard pressed to do much better than in the interval equilibrium.

5 Extensions

Having explored the model, we now turn to extensions. Our goal is to address potential concerns about the robustness of our results and to show that the model can easily be adapted and extended in a variety of ways.

5.1 Sender Prefers Larger Decisions

As we noted above, the assumption that the sender prefers smaller decisions is not a normalization. It implies that the decision that is best for the sender is the one that, in the absence of any advice, is safest for the receiver. It applies, for instance, to situations in which there is an expensive repair that solves a problem for sure—a new engine say—and the challenge for the mechanic is to talk the owner out of going with a more limited repair. We now explore the opposite case in which the sender wants to convince the receiver to make the decision about which, in the absence of any advice, he is most uncertain about. In this section, we describe these results informally and exclusively for the case in which the number of decisions goes to infinity. The formal statements and proofs are in Appendix C.1.

Suppose that the sender’s preferences are given by $u_S(d) = d$ rather than $u_S(d) = -d$. Since the sender plays no role in the no-advice benchmark, this change does not affect the decision the receiver would make in the absence of any advice, which is still given by $d_r(d_0) = \bar{x}/\mu$.

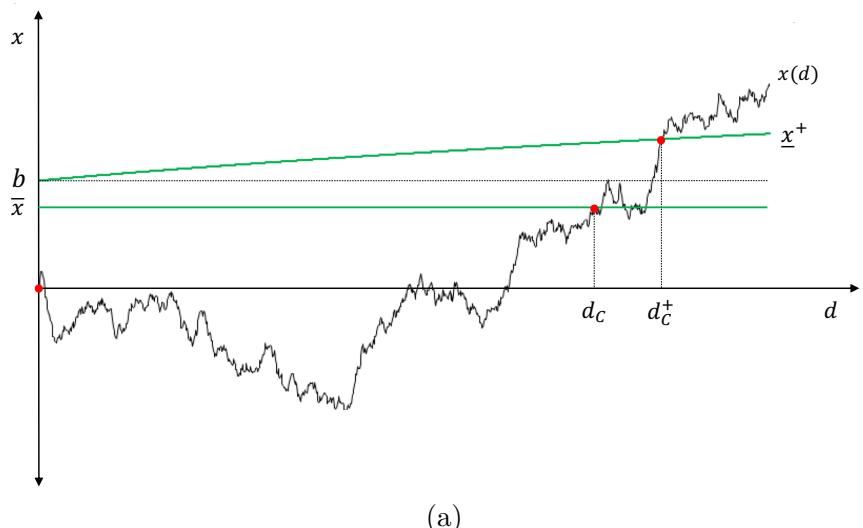
The change in the sender’s preferences does change conative advice. We saw above that when the sender prefers smaller decisions, she cannot do better than to recommend the smallest decision whose outcome is sufficiently close to the receiver’s ideal one. Similarly, when the sender prefers larger decisions, she cannot do better than to recommend the *largest* decision whose outcome is sufficiently close to the ideal one. Figure 10a illustrates both cases. If the sender prefers smaller decisions, she recommends d_C , the smallest decision whose outcome is equal to threshold $\bar{x} < b$. If, instead, she prefers larger

decisions, the sender recommends d_C^+ , the *largest* decision whose outcome is equal to the threshold $\underline{x}^+ > b$. The figure also illustrates that the average decision $E[d_C^+]$ is strictly larger than the decision $d_r(d_0)$ the receiver makes in the absence of advice. In contrast to what we saw above, therefore, conative advice does allow the sender to sway decisions in her favor when she has a preference for larger decisions.

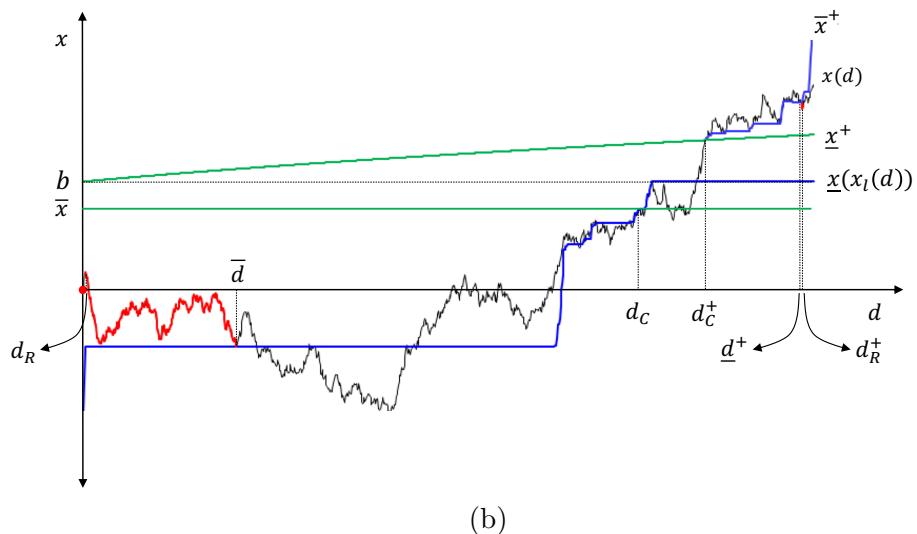
It is still the case, though, that referential advice allows the sender to sway decisions in her favor more effectively. We saw above that when the sender prefers smaller decisions, she can improve on conative advice by complementing her recommendation with information about neighboring decisions to its right. Similarly, when the sender prefers larger decisions, she can improve on conative advice by complementing her recommendation with information about neighboring decisions to its *left*. Figure 10b illustrates both cases. If the sender prefers smaller decisions, she reveals the outcomes of decisions in the interval $\{d_R, \dots, \bar{d}\}$, where \bar{d} is the smallest decision at which the outcome function is equal to the threshold $\underline{x}(\cdot)$ and d_R is the decision to the left of \bar{d} whose outcome is best for the receiver. The threshold $\underline{x}(\cdot)$ is chosen such that the receiver is indifferent between following the recommendation d_R and making any decision to the right of \bar{d} .

If, instead, the sender prefers larger decisions, she reveals the outcomes in the interval $\{\underline{d}^+, \dots, d_R^+\}$, where \underline{d}^+ is the *largest* decision at which the outcome is equal to the threshold \bar{x}^+ and d_R^+ is the decision to the *right* of \underline{d}^+ whose outcome is best for the receiver. The threshold \bar{x}^+ is chosen such that the receiver is indifferent between following the recommendation d_R^+ and making any decision to the *left* of \underline{d}^+ .

Above we saw that when the sender prefers smaller decisions, she does better in the interval equilibrium than in any conative equilibrium *with probability one*. This is no longer the case when the sender prefers larger decisions. In particular, it is no longer the case that the interval equilibrium almost always implements a decision that is strictly better for the sender than the decision that is implemented in the sender's preferred conative equilibrium. It is, however, still the case that the interval equilibrium implements a decision that is strictly better for the sender with positive probability and never implements a decision that is strictly worse. Referential advice, therefore, still allows the sender to do strictly better than conative advice *on average*. And it allows her to do so through the same mechanism, by creating uncertainty and doubt about the wisdom of not following the recommendation.



(a)



(b)

Figure 10

5.2 Receiver Cares about Outcomes & Decisions

Our model follows the literature on hard information in assuming a very clean-cut incentive conflict: the sender only cares about the decision while the receiver only cares about the outcome. This approach has the benefit of clarity and familiarity but limits the applicability of the model. A patient with limited health insurance, for instance, cares not only about how treatments impact his health but also about how they affect his pocket book. The model can easily be adapted to such applications and doing so allows us to generate additional empirical predictions. In this section, we describe this extension informally and exclusively for the case in which the number of decisions goes to infinity. The formal statements and proofs are in Appendix C.2.

Suppose the receiver's preferences are given by $u_R(x) = -(x - b)^2 + cd$, where the new parameter $c \in [0, \sigma^2]$ captures the marginal cost to the receiver of a larger decision, holding constant the outcome. The introduction of c affects the model by increasing the upper threshold and reducing the lower one. Both changes reflect the same underlying factor: the more the receiver has a preference for larger decisions, the more attractive it is for him to make a decision to the right of the largest revealed one.

The increase in the upper threshold implies that in the sender's preferred conative equilibrium, the sender now has to recommend a decision whose outcome is closer to the receiver's preferred one. As a result, the receiver is able to obtain a better outcome and, on average, makes a decision that is worse for the sender. Intuitively, since the sender prefers smaller decisions, a higher c makes it more difficult for her to sway decisions in her favor. And since it is more difficult for her to sway decisions in her favor, the receiver is able to obtain a better outcome.

A similar logic applies in the interval equilibrium, where a higher c also leads to an increase in the average decision. Since a higher c reduces the lower threshold, it induces the sender to reveal the outcomes of decisions to the right of the largest one he would otherwise reveal. This increase in the largest revealed decision, in turn, implies an increase in the decision the sender recommends and the receiver makes.

A higher c , therefore, increases the average decision, even in the interval equilibrium. Nevertheless, it is still the case that, on average, the sender is strictly better off in this equilibrium than in any conative one, and that in all states, she is never strictly worse off. The introduction of c , therefore, generates additional, potentially testable predictions. But it does not affect the nature of the different types of advice and the main results we derived above.

5.3 Sender is Imperfectly Informed

Another direction in which the model can be extended is to allow for the sender to be imperfectly informed about the outcomes. In our context, it is natural to suppose that, just like the receiver, the sender is better informed about decisions that are closer to the default decision. To capture this notion, in this extension we assume that the sender observes the realization of a signal function that is correlated with the outcome function. For each decision, she then decides whether to reveal her signal. As in the previous extensions, we describe our results informally and focus on the case in which the number of decisions goes to infinity. The formal statements of our results and the proofs are in Appendix C.3.

Suppose that, instead of observing the realization of the outcome function $x(d)$, the sender observes the realization of signal function $y(d) = x(d) + \varepsilon\xi(d)$, where ξ is an independent standard Brownian motion and where $\varepsilon \geq 0$ captures how informed the sender is. If $\varepsilon = 0$, the sender is perfectly informed about the outcomes, as in our main model. As ε grows larger, the sender becomes gradually less informed until, in the limit in which $\varepsilon = \infty$, she knows as little as the receiver does. Finally, for a fixed ε , the sender is more uncertain about the outcomes of decisions that are further from the default decision, just like the receiver.

This extension can be analyzed in much the same way as our main model by noticing that, given the sender's decision, the expected outcome of decision d is given by $z(d) = (1 - \gamma)y(d) + \gamma\mu d$, where $\gamma \equiv \varepsilon^2/(\sigma^2 + \varepsilon^2)$. The outcome estimates $z(\cdot)$, therefore, form a Brownian motion with drift μ and scale $(1 - \gamma)\sigma$. Rather than revealing outcomes depending on when $x(\cdot)$ hits different thresholds, the sender then reveals signals depending on when $z(\cdot)$ hits suitably adjusted thresholds.

The key difference between this extension and our main model is that there are no longer any conative equilibria. To see why, suppose the sender's strategy were to reveal the smallest decision at which the estimate function $z(\cdot)$ hits the upper threshold, which is still defined by $\bar{x} = b - \sigma^2/(2\mu)$. If the estimate function hits the upper threshold early enough, it is once again a best response for the receiver to make the revealed decision. But if the estimate function hits the upper threshold very late, the receiver is now better off making a decision to the left of the revealed one, perhaps even the default decision. In such cases, the receiver understands that the sender is very unsure about the outcome of the decision she is recommending. Rather than follow such a risky recommendation, he turns it down for something closer to the safe default decision.

There are, however, equilibria that are similar to the conative equilibria we explored above. In these *one-or-all equilibria* the sender reveals the outcomes of either one or all decisions. In the appendix, we show that in one such equilibrium, as ε vanishes, the sender almost surely reveals the outcome of only one decision, which converges to d_C —the decision the receiver makes in the sender’s preferred equilibrium in the main model—and which the receiver finds optimal to make. For sufficiently small values of ε , one-or-all equilibria are, therefore, *almost conative*.

The fact that the receiver is uncertain about the outcomes of large revealed decisions does not cause any issues for referential advice. We show, in particular, that there is an interval equilibrium in which the sender reveals outcomes over a range of decisions up to the point where the outcome estimate $z(\cdot)$ is equal to the upper threshold or to a lower cutoff whose expression is analogous to the interval equilibrium in the main model. On average, the sender is strictly better off in this interval equilibrium than in any one-or-all equilibrium.

6 Concluding Discussion

Expertise is everywhere and its importance verges on being self-evident. Yet grasping why and how it matters, and teasing its effects out empirically, has proven more elusive than one may have expected. The objective of this paper has been to shed more light on the role of expertise, how it manifests in advice, and when it matters. At an abstract level, we hope that the use of referential information resonates with intuition and experience. At a practical level, we aim for our results to open up new channels of understanding and new interpretations of old data. Before concluding, we offer briefly here several areas where this opportunity is most promising.

A striking implication of our results is the benefit to the receiver of simple advice. In comparing the conative and referential equilibria, it stands out that the receiver is better off receiving fewer rather than more pieces of information from the expert. This is surprising as Bayes’ rule implies the receiver can be no worse off receiving more information than less. This is why the conative equilibrium must, by necessity, leave the receiver no worse off than if the expert offered no advice at all. Yet by offering two or more pieces of information, the expert can leave the receiver worse off than if only one piece were provided. Immediately, this suggests that the decision maker would be better off if he can restrict the ability of the expert to give advice. Specifically, if the receiver can restrict the expert to *only* provide a recommendation, the receiver would be better off. This incentive is perhaps evident in the “executive summaries”

favored by CEOs that ease their cognitive load but possibly also restrict the ability of subordinates to shape their decision making. It may also provide the logic to other institutional features, such as the U.S. Congress' requirement that committees submit to the floor only bills for debate and not supporting reports, or the common law requirement that only judicial decisions serve as binding precedent and not the supporting argument contained in a judge's written opinion.

Our results suggest also an explanation for why we rely on experts yet so often feel ripped-off in doing so. In the referential equilibrium the receiver knows he is getting persuaded—ripped-off—by the mechanic, yet he nevertheless persists in taking her advice. Indeed, the decision maker takes the mechanic's advice even though he knows that, more often than not, he will be worse off than he would be if he ignored the expert and took a decision independently. The reason why he persists with the expert is that, by doing so, he avoids the rare catastrophe. The complexity of the decision making environment leaves open the possibility that, without an expert, he might really screw up, as is evident in the simulations of Section 3. It is the possibility of screwing up that the expert exploits, and why the receiver follows advice willingly but unenthusiastically. Ironically, in our interpretation, the receiver knows he is being ripped off yet, at the same time, he knows that the expert is being truthful to him. It suggests that our trust in some experts may be high, such as for doctors, not out of mistaken belief that they have our interests at heart, but because they are telling us the truth even when their preferences diverge.

The importance of rare disasters also suggests new patterns with which to identify the impact of experts in data, such as in studies that compare outcomes with and without experts (Hendel, Nevo, and Ortalo-Magne 2009). Our results suggest that the impact of experts should be evident not only in the average outcomes obtained by decision makers but also in the distribution of these outcomes. Expert advice, even when it persuades the decision maker, removes the possibility of a disaster, and decreases the variance in outcomes obtained relative to when no advice is received. Examining the distribution of outcomes, in addition to the average outcome, should provide more evidence of the impact of experts.

Another empirical application to which our results speak is the example we have carried throughout the paper, that of doctors and medical expertise. A remarkable feature of health care in the United States is how much variation exists in the incidence and quality of health care across the country, and considerable effort has been put into understanding the reasons why (Chandra, Cutler, and Song 2012). Our results suggest a novel interpretation for the complicated interdependency between insurance coverage and medical care. In

addition to the demand and supply side factors documented in the literature, our results suggest that differences in insurance produce different degrees of persuadability of patients, and that this may explain some of the variation in the data. This relationship can be seen in our extension of the model in which the receiver cares about both outcomes and decisions (which captures well the importance of quality medical care and its high cost in the U.S.). In that extension we show that a receiver who cares less about the decision—a patient with good insurance and lower copays—is more susceptible to persuasion and, as a result, will receive worse outcomes. Examining this connection more closely in the data, and separating it from other demand side effects, offers an interesting and promising direction to investigate.

Finally, our results imply an obvious but little explored property: that to identify the role of experts, we should look directly at what, and how much, they say. Our results imply that an expert's power comes not only from what she knows, but how she expresses that knowledge, and that the most powerful experts provide more than the minimal amount of information necessary to influence a decision maker. This implies that by looking at how much information experts provide, we can identify the nature of expertise and the leverage the expert wields. Turning this around, evidence of expertise should also be evident in the beliefs held by decision makers. Referential advice is persuasive as it changes beliefs about many actions, and not just beliefs about the recommended decision. Decision makers presented with referential information form different views of the environment in which they live from those who aren't. Looking more closely at what experts say, therefore, and how it shapes the beliefs of decision makers, as well as their actions, offers the opportunity to identify more clearly the role and impact of experts.

Appendices

To avoid ambiguity, throughout the appendices we denote by uppercase X the full decision-outcome path, and by lowercase x a single outcome. The numbering for equations and formal results are preceded by the letter ‘A’ as in Appendix to distinguish them from the main text.

A Partial Equilibrium

A perfect Bayesian equilibrium requires to specify both on- and off-equilibrium path beliefs for the receiver. When constructing an equilibrium, it is convenient to focus on the on-path messages—the messages that the receiver expects to observe given the sender’s strategy—and complete the equilibrium specification with adequate beliefs and decisions for off-path messages—the messages that the receiver does not expect to observe—in a way that captures the idea of skeptical posture for the receiver. The purpose of this appendix is to give sufficient conditions under which one can dispense with a full equilibrium specification, leaving the part that concerns off-path messages unspecified.

The concept of partial equilibrium applies to *partial strategy profiles* of the form $(\mathcal{M}^*, D^*, M^*)$, in which:

- $\mathcal{M}^* \subseteq \mathcal{M}$ is the set of messages that the sender communicates for at least one state of the world, i.e., it is the set of “on-path” messages;
- D^* is a mapping from \mathcal{M}^* to the set of decisions \mathcal{D} , it represents the strategy of the receiver, but only accounts for on-path messages; the receiver’s decision for off-path messages is left unspecified;
- M^* is a surjective mapping from Θ to \mathcal{M}^* which represents the strategy of the sender; its surjective nature means that every message of \mathcal{M}^* can be expected under some state of the world, which makes \mathcal{M}^* the set of on-path messages for D^* .

We refer to M^* and D^* as *partial strategies with associated message space* \mathcal{M}^* . The triple $(\mathcal{M}^*, D^*, M^*)$ is called a *partial equilibrium* when the following conditions are satisfied.

- (i.) the sender’s strategy M^* maximizes her utility in every state given D^* and \mathcal{M}^* ,

- (ii.) in every state, the sender is never strictly better off revealing the outcomes of every decision, if she believes that, doing so, the receiver will make a decision that maximizes his utility,
- (iii.) letting $B^*(m)$ be the probability distribution over states θ given $M^*(\theta) = m$, the receiver's strategy M^* maximizes his utility for every message of \mathcal{M}^* given his belief captured by B^* .

Unlike a perfect Bayesian equilibrium, a partial equilibrium does not account for the possibility that the sender may deviate by communicating a message outside of \mathcal{M}^* , except for the case of full disclosure. The following result asserts that, as long as the receiver chooses a decision among the ones whose outcome has been revealed, any partial equilibrium can be completed to make a perfect Bayesian equilibrium. The proof works by constructing a skeptical belief for the receiver, which leads him to make a bad decision from the sender's viewpoint.

Lemma A1 *Let $(\mathcal{M}^*, D^*, M^*)$ be a partial equilibrium. Suppose that upon receiving any message $m \in \mathcal{M}^*$, the receiver takes a decision whose outcome is revealed in message m . Then, there exists a receiver strategy D with $D(m) = D^*(m)$ for every $m \in \mathcal{M}^*$, a sender strategy M with $M = M^*$, and a receiver belief function B , such that (B, D, M) is a perfect Bayesian equilibrium.*

Proof. We construct the perfect Bayesian equilibrium (B, D, M) as follows. First, we let $M(\theta) = M^*(\theta)$ for every state θ . Second, if $m \in \mathcal{M}^*$, we let $D(m) = D^*(m)$ and $B(m) = B^*(m)$, where, as above, $B^*(m)$ is the probability distribution over states θ given $M^*(\theta) = m$.

It remains to specify the decision of the receiver and his belief upon observing an off-path message. Consider such an off-path message $m \notin \mathcal{M}^*$. Let d^* be a decision whose outcome is revealed and that maximizes the receiver's utility among all decisions whose outcome is revealed. If more than one decision maximizes the receiver's utility, then we take d^* to be the largest among these decisions. If the message is empty, then we let $d^* = 0$. Let d^\dagger be the largest decision whose outcome is *not* revealed. There are two cases to consider.

- If $d^\dagger < d^*$ then let $B(m)$ be the belief that is consistent with m for all decisions whose outcomes are revealed in m , and otherwise associates outcome 0 to every decision not disclosed in m . Let $D(m) = d^*$.
- If $d^\dagger > d^*$ then let $B(m)$ be the belief that is consistent with m for all decisions whose outcomes are revealed in m , and otherwise associates outcome 0 to every undisclosed decision $d < d^\dagger$, and associates outcome b to decision d^\dagger . Let $D(m) = d^\dagger$.

In both cases, the receiver's decision is a best response given his beliefs, and his beliefs are always consistent with the message communicated.

We now show that the sender does not have a strictly profitable deviation. Let θ be any state, and m be any message consistent with θ . Let $m^o = M^*(\theta)$ and $d^o = D^*(m^o)$ be the message and decision that would respectively occur under the originally prescribed strategy for the sender and the receiver.

If $d^o < d^\dagger$, and $d^\dagger > d^*$, then the sender who deviates from message m^o to message m induces a larger decision and so a decrease of her utility: such a deviation is strictly suboptimal. If $d^o \leq d^*$, and $d^\dagger < d^*$, then the deviation is also suboptimal. Finally, the combination $d^o > d^\dagger$ and $d^o > d^*$ cannot occur. Indeed, if this combination was possible, the sender would be able to strictly increase her utility by revealing the entire decision-outcome path, by inducing the receiver to choose a decision smaller than d^o , which is not possible by definition of a partial equilibrium.

Overall, the triple (B, D, M) just defined satisfies the conditions required of a perfect Bayesian equilibrium. ■

General Message Spaces

We now consider the case in which the sender can provide more general messages than the messages of the set \mathcal{M} . Specifically, we assume that the sender can send, as message, any set of decision-outcome paths that includes the true decision-outcome path. This is the same as saying that the sender can communicate any set of states that includes the true state.

Let $\overline{\mathcal{M}}$ be this extended set of possible messages. A strategy for the sender is now a mapping M from Θ to $\overline{\mathcal{M}}$, a strategy for the receiver is now a mapping D from $\overline{\mathcal{M}}$ to \mathcal{D} , and a belief function which captures the receiver's belief on possible states upon reception of a message is now a mapping from $\overline{\mathcal{M}}$ to $\Delta(\Theta)$. The definition of perfect Bayesian equilibrium extends immediately to the case of more general message spaces. The notions of partial strategy profile and of partial equilibrium are analogous to the ones above, substituting $\overline{\mathcal{M}}$ for \mathcal{M} .

We provide below an analog of Lemma A1, for the case of general message spaces. Lemma A2 is more general than Lemma A1 and its proof shorter, however unlike Lemma A1 the proof is not constructive, the skeptical belief remains implicit.

Because the equilibria we construct in this paper systematically start from a partial equilibrium, and are then completed to a full, perfect Bayesian equilibrium applying Lemma A1, the more general Lemma A2 implies that the equilibria of this paper extend to the case of general message spaces.

Lemma A2 Assume the message space is $\overline{\mathcal{M}}$, and let $(\mathcal{M}^*, D^*, M^*)$ be a partial equilibrium with $\mathcal{M}^* \subseteq \mathcal{M}$. There exists a receiver strategy D defined over $\overline{\mathcal{M}}$ with $D(m) = D^*(m)$ for every $m \in \mathcal{M}^*$, a sender strategy M with $M = M^*$, and a receiver belief function B also defined over $\overline{\mathcal{M}}$, such that (B, D, M) is a perfect Bayesian equilibrium.

Proof. We construct the perfect Bayesian equilibrium (B, D, M) as follows.

Let B^* be the probability distribution over states θ given $M^*(\theta) = m$. As for the case of Lemma A1, we let $M(\theta) = M^*(\theta)$ for every state θ , and if $m \in \mathcal{M}^*$, we let $D(m) = D^*(m)$ and $B(m) = B^*(m)$.

Now let $m \notin \mathcal{M}^*$. For any decision-outcome path X , let $d^\dagger(X)$ be the decision that maximizes the receiver's utility for decision-outcome path X . If two or more such decisions exist, let $d^\dagger(X)$ be the largest one. Let

$$d^*(m) = \max_{X \in m} d^\dagger(X).$$

Because there are finitely many decisions, there exists X such that $d^\dagger(X) = d^*(m)$. Let $B(m)$ be the belief that assigns probability one to such decision-outcome path X , and let $D(m)$ be the decision $d^\dagger(X) = d^*(m)$.

Let us verify that the triple (B, D, M) just defined satisfies the conditions required of a perfect Bayesian equilibrium. By construction, belief function B is an adequate belief for the receiver. In addition, the receiver best responds according to his belief by following the prescribed strategy D , following the definition of the partial equilibrium for on-path messages, and following the definition of d^\dagger for off-path messages. Now let us check that the sender also best responds. Because $(\mathcal{M}^*, D^*, M^*)$ is a partial equilibrium, the sender is never strictly better off deviating to an on-path message. If, for state θ associated with decision-outcome path X , the sender deviates from sending message $M(\theta)$ to sending off-path message $m \notin \mathcal{M}^*$ compatible with X instead, the decision that the receiver makes is $D(m) = \max_{X' \in m} d^\dagger(X') \geq d^\dagger(X)$ which is larger than or equal to the decision that would occur had the sender disclosed the entire decision-outcome path X . By definition of a partial equilibrium, the sender is never strictly better off deviating to reveal the entire path X , even if X is not an on-path message. Hence, the sender is never strictly better off announcing an off-path message.

Therefore, (B, D, M) is a perfect Bayesian equilibrium. ■

B Omitted Proofs of Sections 3 and 4

B.1 Proof of Lemma 1

The first part of the lemma, the case $d \geq d'$, follows immediately from the law of motion that defines the distribution of $X(d)$. The second part of the lemma, the case $d \leq d'$, is proved by the projection formulas for jointly normal random variables. Let $x = X(d)$ and $x' = X(d')$. The pair (x, x') has a jointly normal distribution. In addition, $\text{Cov}[x, x'] = \sigma^2 d$ and $\text{Var}[x'] = \sigma^2 d'$. By the projection formulas for jointly normal random variables, we have

$$E[x | x'] = E[x] + \frac{\text{Cov}[x, x']}{\text{Var}[x']} (x' - E[x']) = \mu d + \frac{d}{d'} (x' - \mu d') = \frac{d}{d'} x',$$

and

$$\text{Var}[x | x'] = \text{Var}[x] - \frac{\text{Cov}[x, x']^2}{\text{Var}[x']} = \sigma^2 d - \sigma^2 \frac{d}{d'} d = \sigma^2 \frac{d}{d'} (d' - d).$$

B.2 Proof of Proposition 3

We prove a slightly more general result: if $x_C \in [\bar{x}, b]$, then there exists a conative equilibrium that, in the limit case $n \rightarrow \infty$, implements outcome x_C in almost all decision-outcome paths.

Fix $x_C \in [\bar{x}, b]$, and fix the number of available decisions $n + 1$. We focus on the construction of a partial equilibrium, which only accounts for on-path messages, and then apply Lemma A1 to complete the partial equilibrium to a full, perfect Bayesian equilibrium.

Let $\Delta = b - x_C$, and consider the following (partial) strategy profile:

- The sender reveals the outcome of the smallest decision that falls within the range $[b - \Delta, b + \Delta]$. If no such decision exists, the sender reveals the outcomes of all decisions.
- If the entire decision-outcome path is revealed, the receiver chooses the utility maximizing decision. If two or more decisions are optimal, the receiver chooses the smallest optimal decision. Otherwise, if the outcome of a single decision is revealed, the receiver chooses that decision.

Let us verify that this strategy profile is a partial equilibrium.

Sender Optimality: If the decision-outcome path does not include any outcome in the range $[b - \Delta, b + \Delta]$, then there exists only one on-path

message compatible with the decision-outcome path, and so the sender's message is trivially optimal among on-path messages. If the decision-outcome path includes any outcome in the range $[b - \Delta, b + \Delta]$, then for the message to be on path, the sender must communicate a decision paired with an outcome in the said range. The sender, who prefers smaller to larger decisions, is best off revealing the smallest such decision. Thus, the sender's message is optimal among on-path messages. In addition, the sender is never strictly better off revealing the entire decision-outcome path: doing so would yield a decision that is at least as large as the decision originally anticipated. Hence, the sender's strategy is optimal, in the partial equilibrium sense.

Receiver Optimality: The case of the sender revealing all decisions is immediate. We focus on the case in which the sender reveals the outcome $x^* \in [b - \Delta, b + \Delta]$ of only one decision $d^* > 0$.

The receiver is never better off choosing a decision strictly smaller than d^* , because the outcomes for every such decision fall outside of the interval $[b - \Delta, b + \Delta]$, for which the receiver's utility is strictly less than for a decision whose outcome is known to be inside this interval.

If $d > d^*$, the receiver believes the outcome $X(d)$ is distributed normally with mean $E[X(d) | X(d^*) = x^*] = x^* + (d - d^*)\mu$ and variance $\text{Var}[X(d) | X(d^*) = x^*] = (d - d^*)\sigma^2$. We then must consider two cases.

- If $x^* \geq b$, from the receiver's viewpoint, the average outcome from any decision strictly larger than d^* is further away from b than is x^* , and any decision strictly larger than d^* generates a positive amount of risk. Hence, the receiver is never strictly better off taking a decision strictly larger than d^* .
- If $x^* < b$, the receiver's expected utility when taking decision $d > d^*$ is

$$\begin{aligned} & -(b - E[X(d) | X(d^*) = x^*])^2 - \text{Var}[X(d) | X(d^*) = x^*] \\ & = -(b - x^* - (d - d^*)\mu)^2 - (d - d^*)\sigma^2 \end{aligned}$$

and his expected utility when taking decision d^* is $-(b - x^*)^2$. Hence, the receiver is better off choosing decision d^* exactly when

$$x^* \geq b - \frac{(d - d^*)\mu^2 + \sigma^2}{2\mu},$$

condition which is always satisfied if $x_C \geq \bar{x}$. So, as for the case $x^* \geq b$, the receiver is never strictly better off when taking a decision larger than d^* . Hence, the receiver's strategy is optimal for on-path messages.

The above analysis shows that the partial strategy profile just described is a partial equilibrium. Lemma A1 then guarantees that this partial strategy profile can be extended to a full strategy profile in which the receiver adopts a skeptical posture for off-path messages, and that this profile is a perfect Bayesian equilibrium.

In this equilibrium, in the limit $n \rightarrow \infty$, with probability 1, the decision-outcome path hits outcome x_C . The sender then reveals as message the pair (d^*, x_C) where d^* is the smallest decision for which $X(d^*) = x_C$. In addition, the outcomes of the decisions $d < d^*$ are less than x_C . Hence, this equilibrium is conative.

Note that there are conative equilibria which do not implement a fixed outcome in the limit case $n \rightarrow \infty$: for some conative equilibria, the outcome implemented can depend on the decision, and for any given equilibrium decision, it can also be a function of the realized decision-outcome path.

It is readily verified that, in the limit case $n \rightarrow \infty$, the sender's expected utility is maximized for the conative equilibrium which implements \bar{x} , this is the sender's preferred conative equilibrium. It is also readily verified that any conative equilibrium that, in the limit case $n \rightarrow \infty$, implements $x_C \in (\bar{x}, b]$, is strictly suboptimal for the sender, because for almost all decision-outcome paths, the decision is strictly smaller in the sender's preferred conative equilibrium.

B.3 Proof of Proposition 4

We will show that, in any conative equilibrium, as n grows large, the probability that the equilibrium outcome is outside the range $[\bar{x}, b]$ vanishes.

First observe that, if $x > b$, then as n grows large, the probability that the equilibrium outcome is greater than x vanishes. Indeed, if it was not the case, then with nonzero probability, for any sufficiently large number of decisions, revealing the full decision-outcome path would yield a smaller decision and thus would be a profitable deviation for the sender. It follows that the probability that the equilibrium outcome is greater than b vanishes as n grows large. In the remainder of the proof, we show that the equilibrium outcome is less than \bar{x} with vanishing probability.

For every number of decisions $n + 1$, let \mathcal{X}_n be the set of decision-outcome paths for which the sender communicates exactly one data point, and the

receiver chooses the decision communicated by the sender. By assumption, $\Pr(\mathcal{X}_n) \geq 1 - \epsilon_n$. Let $\mathcal{X}_n(d)$ be the subset of \mathcal{X}_n for which the sender reveals the outcome of decision d . Let $\mathcal{O}_n(d)$ be the set of possible outcomes $X(d)$ when $X \in \mathcal{X}_n(d)$. Note that $\mathcal{O}_n(d)$ may be empty for some values of d . Observe that if $x \in \mathcal{O}_n(d)$, then the receiver chooses decision d when the sender communicates (d, x) . If $X \in \mathcal{X}_n$, the sender's best response is thus to reveal the outcome x of the smallest decision d such that $x \in \mathcal{O}_n(d)$.

Suppose (d^*, x^*) is such that $x^* \in \mathcal{O}_n(d^*)$ and $d^* < d_n$. Let $d > d^*$. The receiver's belief on $X(d)$ upon receiving message (d^*, x^*) is characterized by the cumulative distribution

$$y \mapsto \Pr[X(d) \leq y \mid X(d^*) = x^*, X \in \mathcal{X}_n(d^*)].$$

Note that

$$\begin{aligned} & \Pr[X(d) \leq y \mid X(d^*) = x^*, X \in \mathcal{X}_n(d^*)] \\ &= \Pr[X(d) \leq y \mid X(d^*) = x^*, X \in \mathcal{X}_n, X(d') \notin \mathcal{O}_n(d') \forall d' < d^*] \\ &= \Pr[X(d) \leq y \mid X(d^*) = x^*, X \in \mathcal{X}_n], \end{aligned}$$

where the first equality owes to the sender's best response just mentioned, and the second equality is due to the Markov property of the decision-outcome path.

Because the probability that the decision-outcome path is in \mathcal{X}_n converges to 1, the value of $\Pr[X(d^*) \leq y \mid X(d^*) = x^*, X \in \mathcal{X}_n]$ becomes arbitrarily close to the value of $\Pr[X(d^*) \leq y \mid X(d^*) = x^*]$.⁴ Consequently, the expected utility of the receiver when taking a decision $d > d^*$, conditionally on $\{X(d^*) = x^* \text{ and } X \in \mathcal{X}_n\}$, becomes arbitrarily close to the expected utility of the receiver when taking a decision $d > d^*$, conditionally on $\{X(d^*) = x^*\}$. This expected utility is strictly larger than the utility of taking decision d^* if

$$x^* < b - \frac{(d - d^*)\mu^2 + \sigma^2}{2\mu},$$

and hence if $x^* < \bar{x}$, for n large enough, the receiver, upon observing message (d^*, x^*) , is strictly better off choosing a decision strictly larger than d^* . Hence, for any $x < \bar{x}$, as n grows large, for any state in which the sender communicates a single data point whose decision is not the largest available decision, the associated outcome cannot be less than x .

⁴In general, if, for every n , A_n and B_n are two nonnull events and $\Pr[B_n] \rightarrow 1$, then $|\Pr[A_n|B_n] - \Pr[A_n]| \rightarrow 0$.

In addition, as n grows large, with a probability that converges to 1, the sender strictly prefers to reveal the entire decision-outcome path than to reveal as only data point the outcome of the largest decision d_n , because by revealing the entire decision-outcome path, the receiver makes a decision which is strictly less than d_n for a set of states whose probability converges to 1. Therefore, as n grows large, with probability converging to 1, the sender reveals the outcome of a single decision which is not the largest available decision.

Putting the last two facts together, it follows that the probability that the equilibrium outcome is less than \bar{x} vanishes.

B.4 Proof of Proposition 5

If the sender does not provide any information, the receiver takes a decision d that maximizes his expected utility

$$-(E[X(d)] - b)^2 - \text{Var}[X(d)] = -(\mu d - b)^2 - \sigma^2 d.$$

As $n \rightarrow \infty$, the optimal decision converges to \bar{x}/μ . In the sender's preferred conative equilibrium, which implements outcome \bar{x} , in the limit case $n \rightarrow \infty$, the equilibrium decision is equal to the first "hitting time" (in the language of stochastic calculus) τ of the barrier \bar{x} for the decision-outcome path, which becomes a Brownian motion starting at 0, with drift μ and scale σ . This average hitting time is also equal to \bar{x}/μ (see, for example, Dixit 1993, p. 56).

B.5 Proof of Proposition 7

To begin, we show that the strategy profile is a partial equilibrium.

Sender Optimality: The strategy of the sender is optimal (in the partial equilibrium sense). If the sender reveals the outcomes of the range of decisions $\{d_x, \dots, d_y\}$, and that message is on path, then, by definition, $d_y \geq \bar{d}$, and so the receiver's decision is greater than or equal to d_R : the sender never strictly benefits by choosing $d_y > \bar{d}$. In addition, $d_x \geq d_R$ for the message to be on path, and if $d_x > d_R$, the receiver makes a decision greater than d_R which makes the sender worse off. The sender is also never strictly better off revealing the entire decision-outcome path—in the best possible case for the sender, the receiver would only make the same decision as in the originally prescribed sender strategy.

Receiver Optimality: Focusing on the on-path messages, if the receiver chooses a decision less than or equal to \bar{d} , decision d_R is, by definition of

d_R , optimal. If the receiver chooses a decision greater than \bar{d} , then by definition of \bar{d} , the receiver's expected utility is at most that of taking decision d_R . Hence, the receiver never strictly benefits from choosing a decision greater than \bar{d} .

A direct application of Lemma A1 guarantees that the partial equilibrium extends to a perfect Bayesian equilibrium, when the receiver adopts a skeptical posture when observing an off-path message. Note that, in this equilibrium, the beliefs of the receiver are neutral to the right of \bar{d} .

In the limit case $n \rightarrow \infty$, the decision-outcome path is the realization of a Brownian motion starting at 0, with drift μ and scale σ . If the receiver is given the outcome x of decision d , with $x \leq \bar{x}$, and chooses decision $d + \Delta$, with $\Delta \geq 0$, and if the receiver's beliefs are neutral to the right of d , then the receiver's expected utility is equal to

$$-(x + \mu\Delta - b)^2 - \Delta\sigma^2$$

whose maximum is reached for $\Delta \geq 0$ and is equal to

$$-\frac{\sigma^2}{\mu}(b - x) + \frac{\sigma^4}{4\mu^2}.$$

If $x = \bar{x}$, then the maximum is reached for $\Delta = 0$, and if $x > \bar{x}$, the maximum is reached for $\Delta > 0$. Besides, the utility of any decision d whose outcome is known to be x is $-(x - b)^2$. Therefore, in this limit case, the value of \bar{d} is the smallest value of d such that

$$b - X(d) = \frac{\mu}{\sigma^2} \min\{(b - X(d'))^2 : d' \in [0, d]\} + \frac{\sigma^2}{4\mu}.$$

Note that if d is the first hitting time of \bar{x} , then d satisfies this equation. Because this first hitting time is almost surely finite, in the limit case, $\bar{d} < \infty$ with probability 1. In turn, the decision d_R is the left-most minimizer of $d \mapsto (b - X(d))$ for $d \leq \bar{d}$. The fact that the equilibrium is referential is due to the fact that $d_R < \bar{d}$ with probability 1, which is a consequence of Proposition 8.

B.6 Proof of Proposition 8

Let us focus on the case $n \rightarrow \infty$, for which decision-outcome path is the realization of a Brownian motion starting at 0, with drift μ and scale σ . Given a decision-outcome path X , let us define τ_a, τ_b as follows.

- τ_a is the smallest decision d such that $X(d) = \bar{x}$.

- τ_b is the smallest decision d such that

$$b - X(d) = \frac{\mu}{\sigma^2} \min\{(b - X(d'))^2 : d' \in [0, d]\} + \frac{\sigma^2}{4\mu}.$$

Note that, in the interval equilibrium, \bar{d} is the minimum of τ_a and τ_b , while in the sender's preferred conative equilibrium, the implemented decision is equal to τ_a . The decision implemented in the interval equilibrium is, therefore, never greater than the decision implemented in any conative equilibrium (if the entire decision-outcome path is revealed, then the decisions are the same). We now prove that, with probability 1, $\tau_a > \tau_b$, implying that the decision implemented in the interval equilibrium is strictly smaller than the decision implemented in the conative equilibrium.

Note that, with probability one, $\tau_a < \infty$. The proof proceeds by contradiction. Let us suppose that $\tau_a < \tau_b$ with probability $\epsilon > 0$. Let $\delta > 0$ be such that $\bar{x} > \delta$. Let τ_δ be the stopping time defined by the smallest decision d such that $X(d)$ hits value $\bar{x} - \delta$. We then consider two stopping times based on τ_δ . First, a stopping time τ_L defined as the smallest decision $d > \tau_\delta$ such that

$$X(d) = \bar{x} - \delta - \frac{\mu}{\sigma^2} \delta^2.$$

Second, a stopping time τ_U defined as the smallest decision $d > \tau_\delta$ such that

$$X(d) = \bar{x}.$$

(Of course, $\tau_U = \tau_a$, but the context in which τ_U is used is different, so we find it convenient to define it independently.)

Note that τ_δ is finite with probability one, because $\mu > 0$. Thus, the probability that $\tau_U < \tau_L$ given that $\tau_\delta < \infty$ is the same as the probability that $\tau_U < \tau_L$. By standard results (see, for example, Dixit 1993, pp. 51–54), for any Brownian motion starting at zero, with drift μ and scale σ , and whose value at zero is $\bar{x} - \delta$, the probability of reaching the upper barrier $V^H \equiv \bar{x}$ before reaching the lower barrier $V^L \equiv \bar{x} - \delta - \frac{\mu}{\sigma^2} \delta^2$ is

$$\begin{aligned} & \frac{\exp(2V^H \mu/\sigma^2) - \exp(2(\alpha + \delta)\mu/\sigma^2)}{\exp(2V^H \mu/\sigma^2) - \exp(2V^L \mu/\sigma^2)} \\ &= \frac{\exp\left(2\left[\delta + \frac{\mu}{\sigma^2}\delta^2\right] \frac{\mu}{\sigma^2}\right) - \exp\left(2\delta \frac{\mu}{\sigma^2}\right)}{\exp\left(2\left[\delta + \frac{\mu}{\sigma^2}\delta^2\right] \frac{\mu}{\sigma^2}\right) - 1} = \frac{\mu}{\sigma^2} \delta + O(\delta^3), \end{aligned}$$

as $\delta \rightarrow 0$, where $\alpha = b - \bar{x}$.

Hence, the probability that $\tau_U < \tau_L$ converges to zero as $\delta \rightarrow 0$. We have $\tau_U = \tau_a$, and if $\tau_b > \tau_\delta$, then we also have $\tau_L \geq \tau_b$. As $\tau_a < \tau_b$ with probability ϵ and $\tau_\delta < \tau_a$ by definition, it follows that $\tau_U < \tau_L$ with probability ϵ . This contradicts the fact that the probability that $\tau_U < \tau_L$ converges to zero as $\delta \rightarrow 0$.

B.7 Proof of Proposition 9

As $\sigma \rightarrow 0$, all uncertainty vanishes, $\bar{x} \rightarrow b$, the receiver reaches his ideal outcome and the equilibrium decision of the interval equilibrium converges to b/μ , as does the decision under the no advice benchmark. And hence, as $\sigma \rightarrow 0$, the difference of sender's expected utilities converges to zero. As $\sigma \rightarrow \sqrt{2b\mu}$, $\bar{x} \rightarrow 0$, so the decision under the no advice benchmark converges to 0, as does the decision of the interval equilibrium. Thus, as $\sigma \rightarrow \sqrt{2b\mu}$, the difference of sender's expected utilities also converges to zero.

If $\sigma \in (0, \sqrt{2b\mu})$ then by Proposition 8, the sender's expected utility in the interval equilibrium is greater than the expected utility of the sender's preferred conative equilibrium. By Proposition 5, the sender's preferred conative equilibrium generates the same expected utility as in the no advice benchmark. Consequently, if $\sigma \in (0, \sqrt{2b\mu})$, the sender's expected utility in the interval equilibrium is greater than in no advice benchmark. Putting these facts together, the difference of sender's expected utilities is maximized for $\sigma \in (0, \sqrt{2b\mu})$.

B.8 Proof of Proposition 10

This proof requires the use of additional definitions. Throughout this section, let us call a *range equilibrium* a perfect Bayesian equilibrium in which, in every state of the world, the sender reveals the outcomes of consecutive decisions, and upon receiving an on-path message, the receiver chooses one the decisions whose outcomes have been disclosed in the message. We say that the range equilibrium (B, D, M) *state-by-state dominates* the range equilibrium (B', D', M') if for every state θ , the receiver's decision in the former equilibrium is smaller than or equal to the receiver's decision in the latter: $D(M(\theta)) \leq D'(M'(\theta))$. We say that a message is *complete* when for any decision d whose outcome is revealed by the message, the message also reveals the outcome of every decision smaller than d . Finally, given a partial equilibrium (\mathcal{M}^*, D, M) , we say that a message $m \in \mathcal{M}^*$ is *redundant* when there exists another message $m' \in \mathcal{M}^*$ such that the receiver makes the same decision upon observing m or m' , and either m includes all the data points of m' , or m' includes all the data points of m .

Lemma A3 *For any partial equilibrium $(\mathcal{M}^*, D^*, M^*)$, there exists a partial equilibrium $(\mathcal{M}^\dagger, D^\dagger, M^\dagger)$ which, in every state, implements the same decision as $(\mathcal{M}^*, D^*, M^*)$, and such that no message of \mathcal{M}^\dagger is redundant.*

Proof. Let $(\mathcal{M}^*, D^*, M^*)$ be a partial equilibrium. Note that the notion of redundancy defines a strict partial order on \mathcal{M}^* , where m is *redundant with respect to m'* when the receiver makes the same decision upon observing m or m' , and m includes all data points of m' .

Let \mathcal{M}^\dagger be the set of all the smallest elements of \mathcal{M}^* according to this partial order. We construct a partial equilibrium $(\mathcal{M}^\dagger, D^\dagger, M^\dagger)$ as follows.

Note that, by construction, one cannot find two messages of \mathcal{M}^\dagger such that one message is redundant with respect to the other. If $m \in \mathcal{M}^\dagger$, the receiver makes the same decision as in the original strategy: $D^\dagger(m) = D^*(m)$. For all states θ such that $M^*(\theta) \in \mathcal{M}^\dagger$, the sender continues to communicate the same message as in the original strategy: $M^\dagger(\theta) = M^*(\theta)$. For all states θ such that $M^*(\theta) \notin \mathcal{M}^\dagger$, there exists at least one message $m \in \mathcal{M}^\dagger$ such that $M^*(\theta)$ is redundant with respect to m , and we let $M^\dagger(\theta) = m$.

This defines a partial strategy profile. For all states, the receiver's decision in $(\mathcal{M}^\dagger, D^\dagger, M^\dagger)$ is the same as in $(\mathcal{M}^*, D^*, M^*)$. It is then immediate that the sender's strategy remains a best response, in the partial equilibrium sense. And for every $m \in \mathcal{M}^\dagger$, and for every $m' \in \mathcal{M}^*$ such that, in $(\mathcal{M}^*, D^*, M^*)$, m' is redundant with respect to m , the receiver's decision upon observing m' in $(\mathcal{M}^*, D^*, M^*)$ is the same as the receiver's decision upon observing m —call it d . Therefore, d is a best response for the receiver who only knows that θ is such that $M^*(\theta)$ is either equal to m or is redundant with respect to m . And hence, d is the best response to m when the sender follows strategy D^\dagger . This makes $(\mathcal{M}^\dagger, D^\dagger, M^\dagger)$ a partial equilibrium. ■

Lemma A4 *If a range equilibrium satisfies the property that the sender always reveals complete messages, then the interval equilibrium state-by-state dominates this range equilibrium.*

Proof. Note that the interval equilibrium is equivalent the equilibrium in which, as opposed to revealing the outcomes of the decisions $\{d_R, \dots, \bar{d}\}$, the sender reveals the outcomes of the decisions $\{d_0, \dots, \bar{d}\}$ (and the receiver continues to take decision d_R). Let us call this equilibrium the *completed interval equilibrium*. To prove the lemma, it is enough to show that any message of the completed interval equilibrium is a prefix of a message of the range equilibrium, and consequently the decision of the receiver in some state in the interval equilibrium is always less than or equal to the decision in the

range equilibrium. By contradiction, if a message of the range equilibrium is a strict prefix of a message of the interval equilibrium, then because the beliefs are neutral for the decisions not disclosed, the receiver is better off experimenting by taking an undisclosed decision: by definition of \bar{d} , in the completed interval equilibrium, the sender reveals the smallest number of decisions to avoid experimentation by the receiver. ■

Lemma A5 *If an equilibrium is a range equilibrium, then there exists a range equilibrium that state-by-state dominates, and in which the sender only reveals complete messages.*

Proof. Starting from a range equilibrium, let $(\mathcal{M}_0, D_0, M_0)$ be the induced partial equilibrium.

We construct a sequence of partial equilibria $(\mathcal{M}_k, D_k, M_k)$ for $k = 1, \dots, n$, where each equilibrium $(\mathcal{M}_k, D_k, M_k)$ satisfies the following properties:

- (i.) No message of \mathcal{M}_k is redundant.
- (ii.) For every $m \in \mathcal{M}_k$, if m discloses the outcomes of the decisions $\{d_a, \dots, d_b\}$ (and no more than these decisions) and if $d_b < d_k$, then $d_a = 0$.
- (iii.) For every $m \in \mathcal{M}_k$, if m discloses the outcomes of the decisions $\{d_a, \dots, d_b\}$ (and no more than these decisions) and if $d_b < d_k$, then for any decision-outcome path X that is compatible with message m , the sender reveals message m for such decision-outcome path, and upon observing m , the receiver takes the smallest decision whose outcome is revealed in m that maximizes the receiver's utility.
- (iv.) For every state, the decision implemented in $(\mathcal{M}_k, D_k, M_k)$ is less than or equal to the decision implemented in $(\mathcal{M}_0, D_0, M_0)$.

We construct the sequence iteratively. Starting from partial equilibrium $(\mathcal{M}_k, D_k, M_k)$, we construct partial equilibrium $(\mathcal{M}_{k+1}, D_{k+1}, M_{k+1})$ as follows.

1. For all $m \in \mathcal{M}_k$ such that either $D_k(m) > d_k$ or the largest decision whose outcome is revealed in m is less than d_k , $m \in \mathcal{M}_{k+1}$, $D_{k+1}(m) = D_k(m)$, and for every state θ such that $M_k(\theta) = m$, $M_{k+1}(\theta) = m$.
2. Let \mathcal{C} be the collection of the other messages, i.e., the messages m such that $D_k(m) \leq d_k$ and the largest decision whose outcome is revealed in m is less than or equal to d_k .

3. We define an equivalence relation in the set of decision-outcome paths as follows. We say that X is equivalent to X' when there exists a message $m \in \mathcal{C}$ which is both compatible with X and X' , and X and X' have the same outcomes for all decisions up to d_k . (The fact that this defines an equivalence relation is immediate.)
4. For every equivalence class C , let $m(C)$ be the longest message of the form $\{0, \dots, d_b\}$ that is compatible with every $X \in C$. We let $m(C) \in \mathcal{M}_{k+1}$, $D_{k+1}(m(C))$ be the smallest decision for the receiver which maximizes his utility among the decisions whose outcomes are revealed in $m(C)$, and for every state θ associated with $X \in C$, we let $M_{k+1}(\theta) = m(C)$.
5. The triple $(\mathcal{M}_{k+1}, D_{k+1}, M_{k+1})$ is easily verified to be a partial equilibrium. We finally apply Lemma A3 to get rid of redundant messages—this operation does not alter the structure of the equilibrium.

The partial strategy profile $(\mathcal{M}_n, D_n, M_n)$ defines a partial equilibrium, which is extended to a perfect Bayesian equilibrium by Lemma A1. This concludes the proof. ■

Proposition 10 then follows from the application of Lemmas A4 and A5.

C Formal Statements and Proofs of Section 5

C.1 Statements and Proofs for Section 5.1

Here we consider the case in which the sender prefers larger decisions as opposed to smaller decisions, letting her utility function be $u_S(d) = +d$. The receiver's utility function remains unchanged.

Let us briefly first revisit the case of nonstrategic advice, for the case when nature communicates (d^*, x^*) with $x^* > b$. Note that, the receiver's preference being unchanged, the case $x^* < b$ is already treated in the main text. As it turns out, the case $x^* > b$ —relevant when the sender prefers larger decisions—is more subtle. Letting

$$\beta(d) = \min \left\{ \frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} - b), b \right\},$$

we have the following:

- If $d^* < 4b^2/\sigma^2$, so that $\beta(d) = \frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} - b)$, then for $x^* < b + \beta(d^*)$, the receiver's optimal decision is d^* . As x^* increases, the optimal decision

decreases. In the limit as x^* grows to infinity, the optimal decision converges to 0.

- If $d^* > 4b^2/\sigma^2$, so that $\beta(d) = b$, then for $x^* < b + \beta(d^*) = 2b$, the receiver's optimal decision is d^* . For $x^* \in [2b, d^*\sigma^2/(2b)]$, the optimal decision is 0. Then, as x^* increases above $d^*\sigma^2/(2b)$, the optimal decision gradually increases towards d^* , and then decreases again towards 0. In the limit as $x^* \rightarrow \infty$, the optimal decision converges to zero.

The first result concerns the existence of a conative equilibrium.

Proposition A1 *If $x_C : \mathcal{D} \mapsto \mathbf{R}$ is nondecreasing with $x_C(d) \in [b, b + \beta(d)]$, then there exists an equilibrium that is conative and, in the limit as $n \rightarrow \infty$, almost surely implements outcome $x_C(d)$ when the equilibrium decision is d .*

Proof. Fix the number of available decisions $n + 1$. Let $\Delta(d) = x_C(d) - b$, and consider the following partial strategy profile.

- The sender reveals the outcome of the largest decision d that falls within the range $[b - \Delta(d), b + \Delta(d)]$. If no such decision exists, the sender reveals the outcomes of all decisions.
- If the entire decision-outcome path is revealed, the receiver chooses the utility maximizing decision. If two or more decisions are optimal, the receiver chooses the largest optimal decision. Otherwise, if the outcome of a single decision is revealed, the receiver chooses the revealed decision.

We now verify that this strategy profile is a partial equilibrium.

Sender Optimality: The optimality of the sender's decision is obtained by the same argument as in the proof of Proposition 3, and so is omitted.

Receiver Optimality: Optimality is immediate when the sender reveals the entire decision-outcome path. Let us focus on the case in which the sender reveals the outcome $x^* \in [b - \Delta(d), b + \Delta(d)]$ of exactly one decision d^* . The receiver is never better off choosing a decision strictly larger than d^* , because x_C is nondecreasing and so any such decision yields an outcome even further away from b than is x^* .

If $d < d^*$, the receiver believes that $X(d)$ is distributed normally with mean $E[X(d) | X(d^*)=x^*] = dx^*/d^*$ and variance $\text{Var}[X(d) | X(d^*)=x^*] = \sigma^2(d^* - d)d/d^*$.

We consider two cases.

- If $x^* \leq b$ then choosing decision $d < d^*$ yields an outcome which, on average, is further away from b than is x^* , and in addition increases the variance of the outcome. The receiver is better off choosing d^* .
- If $x^* > b$ then choosing decision $d < d^*$ yields the receiver's expected utility

$$\begin{aligned} - (b - \text{E}[X(d) \mid X(d^*)=x^*])^2 - \text{Var}[X(d) \mid X(d^*)=x^*] \\ = - \left(b - \frac{dx^*}{d^*} \right)^2 - \sigma^2 \frac{(d^* - d)d}{d^*} \end{aligned}$$

whereas the receiver's expected utility when taking decision d^* is $-(b - x^*)^2$. Hence, the receiver is better off choosing decision d^* when the difference of the two terms, which simplifies to

$$\frac{d^* - d}{(d^*)^2} ((d + d^*)x^* - dd^*\sigma^2 - 2bd^*x^*), \quad (\text{A1})$$

is nonpositive. We have $d^* - d \geq 0$, we remark that the term $(d+d^*)x^*-dd^*\sigma^2-2bd^*x^*$ is linear in d , so if it evaluates nonpositively at the two extremes $d = 0$ and $d = d^*$, it also evaluates nonpositively at all decisions between the two extremes. If $d = 0$, then

$$(d + d^*)x^* - dd^*\sigma^2 - 2bd^*x^* = (x^* - 2b)x^*d^* \leq 0$$

because $x_C(d^*) \leq b + \beta(d^*) \leq 2b$. If $d = d^*$ then

$$(d + d^*)x^* - dd^*\sigma^2 - 2bd^*x^* = d^*(2(x^* - b)x^* - d^*\sigma^2)$$

which is quadratic in x^* and is nonpositive if and only if $x^* \in [\frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} - b), \frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} + b)]$, and, so, is nonpositive because for all d , $x_C(d)$ is in the interval $[b, b + \beta(d)] \subset [\frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} - b), \frac{1}{2}(\sqrt{b^2 + 2d\sigma^2} + b)]$. Hence (A1) is nonpositive and the receiver's maximal expected utility is reached when choosing d^* .

Overall, the receiver's strategy is optimal in the partial equilibrium sense.

Therefore, the partial strategy profile just described is a partial equilibrium. In addition, an analog of Lemma A1 continues to hold for the sort of sender's preferences considered here, requiring only minor modifications. Thus, these partial strategies can be completed to regular, full strategies that make a perfect Bayesian equilibrium.

In the limit case $n \rightarrow \infty$, with probability 1, the decision-outcome path hits the frontier defined by the function $x_C(\cdot)$ at least once and at most finitely many times. The sender then reveals the pair $(d^*, x_C(d^*))$ where d^* is the largest decision for which $X(d^*) = x_C(d^*)$ (and $X(d) > x_C(d)$ if $d > d^*$). Therefore, the equilibrium is conative. ■

Lemma A6 *In any conative equilibrium, as n grows large, the probability that the equilibrium outcome of equilibrium decision d is above $b + \beta(d)$ vanishes.*

The proof of Lemma A6 analogous to the proof of Proposition 4 and is omitted. The lemma implies that the sender's preferred equilibrium, in the limit as the number of decisions goes to infinity, implements outcome $b + \beta(d)$ for equilibrium decision d .

The second result concerns the existence of a referential equilibrium that does at least as well for the sender as the sender's preferred conative equilibrium.

Proposition A2 *There exists an equilibrium that is referential and that, as the number of decisions goes to infinity, provides an expected utility for the sender that is greater than in any conative equilibrium.*

Proof. We build a referential equilibrium whose structure is similar to the interval equilibrium of the main model. Fix the number of decisions $n + 1$, and let \underline{d} be the largest decision that satisfies either one of these two properties: either

$$b \leq X(\underline{d}) \leq b + \beta(\underline{d})$$

or, $X(\underline{d}) > b + \beta(\underline{d})$ and

$$\max\{-(X(d) - b)^2 : d \geq \underline{d}\} \geq -b^2 \quad \text{if } \underline{d} \geq \frac{4b^2}{\sigma^2} \quad \text{and} \quad \frac{X(\underline{d})}{\underline{d}} \leq \frac{\sigma^2}{2b},$$

$$\begin{aligned} \max\{-(X(d) - b)^2 : d \geq \underline{d}\} &\geq \frac{\underline{d}\sigma^2(d\sigma^2 - 4b(X(\underline{d}) - b))}{4(X(\underline{d})^2 - \underline{d}\sigma^2)} \\ &\quad \text{if } \underline{d} < \frac{4b^2}{\sigma^2} \quad \text{or} \quad \frac{X(\underline{d})}{\underline{d}} > \frac{\sigma^2}{2b}. \end{aligned}$$

By convention, let $\underline{d} = \infty$ if no such decision exists. We consider the following partial strategy profile:

- If $\underline{d} = \infty$, the sender reveals the outcomes of all decisions. If $\underline{d} < \infty$, let $\underline{d}^+ = \underline{d}$, and let d_R^+ be the right-most decision among all decisions greater

than or equal to \underline{d} which maximizes the receiver's utility given knowledge of the outcomes. The sender then reveals all the decisions of the interval $\{\underline{d}^+, \dots, d_R^+\}$.

- The receiver chooses the decision that maximizes his utility among the collection of decisions revealed by the sender, and if two or more decisions maximize the receiver's utility, the receiver chooses the largest decision.

There are two important facts. First, $-b^2$ is the maximum possible expected utility the receiver can get if he chooses a decision d less than \underline{d}^+ , for the case when $\underline{d}^+ \geq 4b^2/\sigma^2$ and when $X(\underline{d}^+)/\underline{d}^+ \leq \sigma^2/(2b)$. Second, it can be shown that

$$\frac{\underline{d}^+ \sigma^2 (\underline{d}^+ \sigma^2 - 4b(X(\underline{d}^+) - b))}{4(X(\underline{d}^+)^2 - \underline{d}^+ \sigma^2)}$$

is an upper bound on the maximum possible utility that the sender can obtain by choosing a decision d less than \underline{d}^+ if either $\underline{d}^+ < 4b^2/\sigma^2$ or $X(\underline{d}^+)/\underline{d}^+ > \sigma^2/(2b)$ (this upper bound becomes tight as the number of decisions available becomes infinite). These two facts follow from the above description regarding the receiver's optimal decision in the case of a nonstrategic sender, and from solving the relevant maximization problems. The calculations are tedious but straightforward and so are omitted.

These two facts combined imply that the sender's and the receiver's strategies are a partial equilibrium—the argument is as in Proposition B.5. Then, as in the proof of Proposition A1, we can extend these strategies to make a full strategy profile in which the receiver holds a skeptical posture, and show with an analog to Lemma A1 that these strategies are a perfect Bayesian equilibrium.

In the limit case $n \rightarrow \infty$, we get a compact characterization of the equilibrium: the decision \underline{d}^+ is the largest decision that satisfies either

$$X(\underline{d}^+) = b + \beta(\underline{d}^+),$$

or, if $\underline{d}^+ < 4b^2/\sigma^2$,

$$\max\{-(X(d) - b)^2 : d \geq \underline{d}^+\} \geq \frac{\underline{d}^+ \sigma^2 (\underline{d}^+ \sigma^2 - 4b(X(\underline{d}^+) - b))}{4(X(\underline{d}^+)^2 - \underline{d}^+ \sigma^2)}.$$

With probability 1, such a decision \underline{d}^+ exists. The above-mentioned thresholds that determine the decision \underline{d}^+ imply that the equilibrium decision of the sender's preferred conative equilibrium is never greater than the equilibrium decision of this referential equilibrium, and that with positive probability, it is

strictly smaller. Finally, with positive probability, the message communicated by the sender discloses the outcomes of more than one decision (strictly speaking, to make the equilibrium referential so that, with probability 1, the sender discloses the outcomes of more than one decision, we can have the sender disclose the outcomes of all the decisions greater than or equal to \underline{d}^+ —doing so does not affect the receiver’s decision). ■

C.2 Statements and Proofs for Section 5.2

We now consider the modification of the main model in which the receiver has a preference for larger decisions, letting his utility function be $u_R(x, d) = -(x - b)^2 + c \times d$, where c is a positive constant, and $c < \sigma^2$. The coefficient c captures the degree of preference of larger decisions, the larger the coefficient, the greater the bias. The sender’s utility function remains as in the main model.

We start with the existence of conative equilibria.

Proposition A3 *If $x_C \in [\bar{x} + c/(2\mu), b]$, there exists a conative equilibrium that, as the number of decisions goes to infinity, almost always implements outcome x_C . In addition, in any conative equilibrium, as the number of decisions goes to infinity, the equilibrium decision is almost always in the interval $[\bar{x} + c/(2\mu), b]$.*

Proof. The second part of the proposition follows by the same arguments as in Proposition 4. We focus on the first part of the proposition.

The idea is similar to the proof of Proposition 3. Fix $x_C \in [\bar{x} + c/(2\mu), b]$, and fix the number of available decisions $n+1$. Letting $\Delta = b - x_C$, we consider the same partial strategy profile as in the proof of Proposition 3, and verify that this strategy profile is a partial equilibrium.

Sender Optimality: The argument for the sender optimality is the same as in the proof of Proposition 3, and is omitted.

Receiver Optimality: Let us focus on the case in which the sender reveals the outcome $x^* \in [b - \Delta, b + \Delta]$ of only one decision $d^* > 0$. The receiver is never better off choosing a decision strictly smaller than d^* , because the outcomes for every such decision fall outside of the interval $[b - \Delta, b + \Delta]$ and the outcome is further away from the receiver’s ideal outcome, and also because the decision is smaller but the receiver prefers larger decisions.

If $d > d^*$, the receiver believes the outcome $X(d)$ is distributed normally with mean $E[X(d) | X(d^*)=x^*] = x^* + (d - d^*)\mu$ and variance $\text{Var}[X(d) | X(d^*)=x^*] = (d - d^*)\sigma^2$.

The receiver's expected utility when taking decision $d > d^*$ is

$$\begin{aligned} & -(b - E[X(d) | X(d^*)=x^*])^2 - \text{Var}[X(d) | X(d^*)=x^*] + cd \\ & = -(b - x^* - (d - d^*)\mu)^2 - (d - d^*)\sigma^2 + cd, \end{aligned}$$

and his expected utility when taking decision d^* is $-(b - x^*)^2 + cd^*$. Hence, the receiver is better off choosing decision d^* exactly when

$$x^* \geq b + \frac{c}{2\mu} - \frac{(d - d^*)\mu^2 + \sigma^2}{2\mu},$$

condition which is always satisfied if $x_C \geq \bar{x} + c/(2\mu)$. So the receiver is never strictly better off when taking a decision strictly larger than d^* . The receiver's strategy is optimal in the partial equilibrium sense.

The partial strategy profile just describes is thus a partial equilibrium. The logic of Lemma A1 continues to hold for this sort of receiver's preferences, guaranteeing that these partial strategies can be extended to full strategies that make a perfect Bayesian equilibrium, in which the receiver has a skeptical posture when observing off-path messages. In the limit as $n \rightarrow \infty$, the decision-outcome path hits outcome x_C with probability 1, and so the equilibrium is conative. ■

The sender's preferred conative equilibrium corresponds to $x_C = \bar{x} + c/(2\mu)$: as c increases, the sender is worse off.

We now discuss the existence of a referential equilibrium that is similar to the interval equilibrium of the main model.

Proposition A4 *There exists a referential equilibrium for which, as the number of decisions goes to infinity, the sender's expected utility is greater than her expected utility in any conative equilibrium.*

Proof. Fix the number of available decisions $n + 1$ and consider the following partial strategy profile. To describe the sender's strategy, analogously to the main model, we let \bar{d} be the smallest decision such that, for all decisions to the right of \bar{d} , denoted $\bar{d} + \Delta$, we have:

$$\max_{d \leq \bar{d}} u_R(X(d)) \geq E[u_R(X(\bar{d} + \Delta)) | X(\bar{d})]. \quad (\text{A2})$$

If no such decision exists, let \bar{d} be the greatest possible decision d_n . We then let d_R be the decision that maximizes the receiver's utility among all the decisions less than or equal to \bar{d} , and, if two or more optimal decisions exist, set d_R to be the smallest one. The sender reveals the outcomes of all decisions $\{d_R, \dots, \bar{d}\}$. The receiver's strategy is to take decision d_R upon receiving a message that discloses the outcomes of decisions $\{d_R, \dots, \bar{d}\}$.

By the same argument as in the proof of Proposition 7, this partial strategy profile is a partial equilibrium, and Lemma A1 can be used to complete the strategy profile and make the completed profile a perfect Bayesian equilibrium.

In the limit as the number of decisions goes to infinity, \bar{d} is the smallest decision for which (A2) becomes an equality, and so, solving for this equality in the same way as in the proof of Proposition 7, we obtain that \bar{d} is the smallest decision that satisfies

$$b - X(\bar{d}) = \frac{\mu}{\sigma^2 - c} \min\{(b - X(d))^2 + c(\bar{d} - d) : d \in [0, \bar{d}]\} + \frac{\sigma^2 - c}{4\mu}.$$

With positive probability, the value of $(b - X(d))^2 + c(\bar{d} - d)$ is minimized for $d < \bar{d}$ for a positive mass of decision-outcome paths. As in Appendix C.1, to ensure that the sender communicates the outcomes of two or more decisions with probability 1, we can have the sender reveal the outcomes of all the decisions up to decision \bar{d} —this does not affect the receiver's response.

Also note that outcome $x = \bar{x} + c/(2\mu)$ satisfies

$$b - x = \frac{\mu}{\sigma^2 - c} (b - x)^2 + \frac{\sigma^2 - c}{4\mu},$$

which implies that the smallest decision whose outcome is equal to $\bar{x} + c/(2\mu)$ is never less than \bar{d} . And, of course, with positive probability, $X(\bar{d}) < \bar{x} + c/(2\mu)$. By our remark above, this implies that this referential equilibrium generates a strictly larger expected utility for the sender than does the sender's preferred conative equilibrium. ■

C.3 Statements and Proofs for Section 5.3

Finally, we consider the extension of the main model in which the sender is imperfectly informed about the state of the world. The utilities for the sender and the receiver are as in the main model. Fixing the number of decisions $n + 1$, the outcome of decision d continues to be denoted $X(d)$ and distributed as in the main model. However, instead of observing $X(d)$ for every decision d , the sender now observes $Y(d)$ for every decision d , with $Y(d_0) = X(x_0) = 0$

and

$$Y(d_i) = Y(d_{i-1}) + X(d_i) - X(d_{i-1}) + \sqrt{\frac{\varepsilon^2}{\sqrt{n}}} \xi_i,$$

for $i = 1, \dots, n$, where ξ_i is independently drawn from the standard normal distribution, and $\varepsilon > 0$ captures the amount of noisy in the signals the sender gets to observe. Remark that, when n grows to infinity, X becomes a Brownian motion with drift μ and scale σ , and $Y(d) = X(d) + \varepsilon W(d)$, for W a standard Brownian motion independent of X , so Y is a Brownian motion with drift μ and scale $\sqrt{\sigma^2 + \varepsilon^2}$.

Let $Z(d)$ denote the best estimate of $X(d)$ (a minimizer of the mean-squared error) given the sender's information:

$$Z(d) = \mathbb{E}[X(d) \mid Y(d'), \forall d'].$$

Let $\gamma = \varepsilon^2 / (\sigma^2 + \varepsilon^2)$. The project formula for jointly normal random variables implies that

$$Z(d) = (1 - \gamma)Y(d) + \gamma\mu d, \quad \text{and} \quad \text{Var}[X(d) \mid Y(d'), \forall d'] = \gamma\sigma^2 d.$$

In particular, as n grows to infinity, Z is distributed as a Brownian motion with drift μ and scale $(1 - \gamma)\sigma$. So, compared to *true* decision-outcome path, the drift of the *estimated* decision-outcome path is the same, but the scale is reduced by a factor of $1 - \gamma$, which captures the informativeness of the sender's signals. As $\gamma \rightarrow 0$, the signal becomes perfectly informative and the estimated decision-outcome path becomes confounded with the true decision-outcome path, and as $\gamma \rightarrow 1$, the signal becomes perfectly uninformative and the estimated decision-outcome path becomes a straight line equal to the unconditional expected outcome. It is worth pointing out that the value of the signal $Y(d)$ for decision d is a sufficient statistic to compute the distribution of $X(d)$ conditional on all of the sender's information:

$$\begin{aligned} \mathbb{E}[X(d) \mid Y(d)] &= \mathbb{E}[X(d) \mid Y(d'), \forall d'], \\ \text{Var}[X(d) \mid Y(d)] &= \text{Var}[X(d) \mid Y(d'), \forall d']. \end{aligned}$$

In addition, if $d \geq d'$,

$$\mathbb{E}[X(d) \mid Y(d')] = \mathbb{E}[X(d) \mid Y(d''), \forall d'' \leq d'] = Z(d') + (d - d')\mu, \quad (\text{A3})$$

and

$$\text{Var}[X(d) | Y(d')] = \text{Var}[X(d) | Y(d''), \forall d'' \leq d'] = \gamma d' \sigma^2 + (d - d') \sigma^2. \quad (\text{A4})$$

The expected utility of the receiver who takes decision d given knowledge of $Y(d)$ or (equivalently) $Z(d)$, is

$$-(Z(d) - b)^2 - \gamma \sigma^2 d. \quad (\text{A5})$$

Compared to the receiver's expected utility conditional on $X(d)$, $-(X(d) - b)^2$, notice the presence of the second term $-\gamma \sigma^2 d$. This term captures the disutility the receiver gets for making larger decisions, due to the compounded noise in the sender's signals.

We now turn to the analysis of conative equilibria.

Proposition A5 *If the sender is imperfectly informed about the state, then there does not exist a conative equilibrium.*

Proof. Let $\bar{z} = b - \sigma^2/(2\mu)$, which is also equal to \bar{x} . If nature communicates to the receiver the value $Y(d)$ or $Z(d)$ of decision d , and if $Z(d) \geq \bar{z}$, then the receiver's expected utility is no greater when making decision $d' > d$ than when making decision d , by the same argument as in the construction of the conative equilibria in Section 4. And, again by these same arguments, if n is sufficiently large and $Z(d) < \bar{z}$, then the receiver gets greater expected utility by making some decision $d' > d$. These facts imply, by the same arguments as in Proposition 4, that when the number of decisions goes to infinity, a conative equilibrium cannot implement a decision whose outcome is less than \bar{z} .

Since, as the number of decisions goes to infinity, the estimated decision-outcome path is a Brownian motion starting at 0 with drift μ and scale $(1-\gamma)\sigma$, for every decision d_T , there is a positive probability that the outcomes of all the decisions less than d_T are below \bar{z} . And, because the receiver incurs a disutility linear in the decision, as shown in (A5), if d_T is chosen large enough, the receiver would rather make decision 0, whose outcome is known. Hence, a conative equilibrium does not exist if $\gamma > 0$. ■

However, there exist equilibria in which the sender either reveals the outcome of a single decision (and the receiver then rubber-stamps the sender's recommendation), or reveals the outcomes of all decisions—an *one-or-all* equilibrium.

For example, consider the following strategy profile. Let

$$\Delta(d) = \sqrt{\frac{\sigma^4}{4\mu^2} - \gamma \sigma^2 d}, \quad \text{and} \quad d_M = \frac{\sigma^2}{4\gamma\mu^2}.$$

- The sender reveals the outcome of the smallest decision $d \leq d_M$ that falls within the range $[b - \Delta(d), b + \Delta(d)]$. If no such decision exists, then the sender reveals the outcomes of all decisions.
- If the entire decision-outcome path is revealed, the receiver makes the utility-maximizing decision, choosing the smallest decision in case of ties. If the outcome of a single decision is revealed, the receiver makes that decision. The receiver adopts a skeptical posture for off-path messages.

Proposition A6 *The one-or-all strategy profile described above is a perfect Bayesian equilibrium.*

Proof. The optimality of the sender's strategy is immediate, by the same arguments as in Proposition 3. The optimality of the receiver's strategy follows from the following facts.

First, note that for all $d \leq d_M$, the range of outcomes $[b - \Delta(d), b + \Delta(d)]$ is included in the range of outcomes $[b - \sigma^2/(2\mu), b + \sigma^2/(2\mu)]$. If the sender communicates the outcome of only one decision d^* , applying the receiver's beliefs given by (A3) and (A4) and using the same logic as in Proposition 3, the receiver is never strictly better off choosing a decision to the right of the disclosed decision.

Second, note that, if the sender discloses the outcome of a single decision d^* such that the outcome is on the boundary of the range $[b - \Delta(d^*), b + \Delta(d^*)]$, the receiver's expected utility, when taking decision d^* , is independent of d^* and is equal to

$$-\left(b - b \pm \sqrt{\frac{\sigma^4}{4\mu^2} - \gamma\sigma^2d^*}\right) - \gamma\sigma^2d^* = -\frac{\sigma^4}{4\mu^2}.$$

If the receiver makes a decision $d \leq d^*$ for which $Z(d) \notin [b - \Delta(d), b + \Delta(d)]$, then the receiver's expected utility, given $Z(d)$, is less than $-\sigma^4/(4\mu^2)$. Hence, the receiver is never better off deviating to the left when the sender reveals the signal of only one decision, because according to the sender's strategy, the sender's estimated outcomes of all the decisions d to the left of the decision whose outcome is revealed fall outside the range $[b - \Delta(d), b + \Delta(d)]$.

The proposition then follows from the application of Lemma A1. ■

Note that, in this one-or-all equilibrium, in the limit as the number of decisions goes to infinity, and as $\gamma \rightarrow 0$, we have that $d_M \rightarrow \infty$, so the sender almost always discloses the outcome of only one decision, and we have $\Delta(d) \rightarrow b - \bar{x}$, so the equilibrium converges to the sender's preferred conative equilibrium of Section 4.

Finally, we turn to the analysis of referential equilibria. We consider the following strategy profile, similar to the interval equilibrium of Section 4.

- To define the sender's strategy, let \bar{d} be the smallest decision such that

$$\max_{d \leq \bar{d}} E[u_R(X(d)) | Y(d'), \forall d' \leq \bar{d}] \geq \max_{d > \bar{d}} E[u_R(X(d)) | Z(\bar{d})],$$

and $\bar{d} = d_n$ if no such decision exists. Let d_R be the smallest decision less than or equal to \bar{d} that maximizes the receiver's expected utility given knowledge of $Y(d_R)$. The sender reveals the outcomes of all the decisions $\{d_R, \dots, \bar{d}\}$.

- When the receiver observes the signals of a range of decisions $\{d_R, \dots, \bar{d}\}$, the receiver takes decision d_R . For off-path messages, the receiver adopts a skeptical posture.

Simple but tedious calculations show that, as the number of decisions goes to infinity, the sender chooses, as \bar{d} , the smallest decision d whose estimated outcome hits either one of two thresholds, \bar{z} as defined in Proposition A5, and $\underline{z}(d)$, defined as

$$b - \frac{\sigma^2}{4\mu} - \frac{\mu}{\sigma^2} (Z(\hat{d}) - b)^2,$$

where \hat{d} is the smallest decision less than or equal to d that maximizes the receiver's expected utility conditionally on the sender's information. And, if $\gamma \rightarrow 0$, this strategy profile converges to the interval equilibrium defined in Section 4 (the upper threshold then becomes irrelevant, by Proposition 8).

Proposition A7 *The strategy profile just described is a perfect Bayesian equilibrium, it is referential and, in the limit as the number of decisions goes to infinity, generates strictly more expected utility to the sender than any one-or-all equilibrium.*

Proof. The fact that the strategy profile just described is a perfect Bayesian equilibrium follows from arguments analogous to those of Proposition 7. In this equilibrium, the sender discloses the outcomes of two or more decisions with positive probability, and as in the other two extensions, in this equilibrium the sender's strategy can be extended to reveal the outcomes of all the decisions less than \bar{d} without changing the receiver's response, thereby guaranteeing that the outcomes of two or more decisions are revealed with probability 1.

Finally, by the same logic as in Proposition 4, in any one-or-all equilibrium, if the sender reveals the outcome of only one decision and in the limit as the

number of decisions goes to infinity, it must be the case that the outcome of the revealed decision is no greater than \bar{z} . This fact implies that, in this limit case, the referential equilibrium defined above implements a decision as least as good for the sender as the decision implemented in any one-or-all equilibrium, for all signal realizations. In addition, since, with positive probability, the estimated decision-outcome path hits the lower threshold before hitting the upper threshold, the sender's expected utility is strictly greater in this referential equilibrium than in any one-or-all equilibrium. ■

References

- Ben-Porath, E., E. Dekel, and B. L. Lipman (2017). Mechanisms with evidence: Commitment and robustness. Boston University Department of Economics Working Papers Series WP2017-001.
- Bendor, J., S. Taylor, and R. Van Gaalen (1985). Bureaucratic expertise versus legislative authority: A model of deception and monitoring in budgeting. *The American Political Science Review* 79(4), 1041–1060.
- Brocas, I. and J. D. Carrillo (2007). Influence through ignorance. *The RAND Journal of Economics* 38(4), 931–947.
- Callander, S. (2008). A theory of policy expertise. *Quarterly Journal of Political Science* 3(2), 123–140.
- Callander, S. (2011). Searching and learning by trial and error. *The American Economic Review* 101(6), 2277–2308.
- Chandra, A., D. Cutler, and Z. Song (2012). Who ordered that? The economics of treatment choices in medical care. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, Chapter Six, pp. 397–432. Elsevier.
- Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica* 50(6), 1431–1451.
- Dixit, A. (1993). *The Art of Smooth Pasting*. Routledge.
- Dye, R. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research* 23(1), 123–145.
- Dziuda, W. (2011). Strategic argumentation. *Journal of Economic Theory* 146(4), 1362–1397.
- Garfagnini, U. and B. Strulovici (2016). Social experimentation with interdependent and expanding technologies. *The Review of Economic Studies* 83(4), 1579–1613.
- Gibbons, R., N. Matouschek, and J. Roberts (2013). Decisions in organizations. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 373–431. Princeton University Press.

- Gilligan, T. and K. Krehbiel (1987). Collective decision making and standing committees: An informational rationale for restrictive amendment procedures. *Journal of Law, Economics, and Organization* 3(2), 287–335.
- Glazer, J. and A. Rubinstein (2004). On optimal rules of persuasion. *Econometrica* 72(6), 1715–1736.
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics* 24(3), 461–483.
- Gruber, J. and M. Owings (1996). Physician financial incentives and cesarean section delivery. *The RAND Journal of Economics* 27(1), 99–123.
- Hart, S., I. Kremer, and M. Perry (2017). Evidence games: Truth and commitment. *The American Economic Review* 107(3), 690–713.
- Hendel, I., A. Nevo, and F. Ortalo-Magne (2009). The relative performance of real estate marketing platforms: MLS versus FSBOMadison.com. *The American Economic Review* 99(5), 1878–1898.
- Jakobson, R. (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in Language*, pp. 350–377. MIT Press.
- Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *The American Economic Review* 101(6), 2590–2615.
- Levitt, S. D. and C. Syverson (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics* 90(4), 599–611.
- Matthews, S. A. and A. Postlewaite (1985). Quality testing and disclosure. *The RAND Journal of Economics* 16(3), 328–340.
- Meyer, M. (2017). Communication with self-interested experts. Lecture notes, Nuffield College, Oxford University.
- Milgrom, P. and J. Roberts (1986). Relying on the information of interested parties. *The RAND Journal of Economics* 17(1), 18–32.
- Milgrom, P. and J. Roberts (1988). Economic theories of the firm: Past, present, and future. *The Canadian Journal of Economics* 21(3), 444–458.

- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* 12(2), 380–391.
- Rapoport, D. (2017). Evidence and skepticism in verifiable disclosure games. *Working paper, available at SSRN, abstract No 2978288*.
- Rayo, L. and I. Segal (2010). Optimal information disclosure. *Journal of Political Economy* 118(5), 949–987.
- Sher, I. (2011). Credibility and determinism in a game of persuasion. *Games and Economic Behavior* 71(2), 409–419.
- Shin, H. (2003). Disclosures and asset returns. *Econometrica* 71(1), 105–133.
- Sobel, J. (2013). Giving and receiving advice. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress*, pp. 373–431. Princeton University Press.
- Weber, M. (1958). *From Max Weber: Essays in Sociology*. Oxford University Press: New York.