Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Title: Effective Persuasion

By Ying Chen and Wojciech Olszewski

No.
1310

This paper has been accepted for publication in International Economic Review. It is also available as an SSRN working paper (doi:10.2139/ssrn.1888608)

Effective Persuasion*

Ying Chen[†] and Wojciech Olszewski[‡]

February 3, 2011

Abstract

Do elementary statistics or equilibrium theory deliver any rules of thumb regarding how we should argue in debates? We suggest a framework for normative analysis of debates.

In our framework, each discussant wants the audience to believe that the actual state coincides with the discussant's favorite state. We show that if the discussants' payoff functions in the audience's posterior are concave above the prior, convex below the prior, and exhibit some form of loss aversion, then the discussant who begins the debate should first present weaker arguments rather than stronger arguments, and the discussant who speaks second should respond with weak arguments to weak arguments, and with strong arguments to strong arguments.

We derive similar rules of thumb regarding the choice between presenting evidence that is independent of the opponent's evidence versus presenting evidence that is potentially correlated.

1 Introduction

The way we argue is important for achieving our goals in a debate. After every debate in presidential or parliamentary elections, we often ask who "won" the debate. Other examples abound, ranging from deliberation contests in colleges to our everyday experience. Everyone who engages in a public debate faces the problem of how to persuade effectively: Which points to raise and in which order? Which of them to emphasize and which of them to disregard? How to respond to the points raised by opponents?

*We thank Mikhail Safronov for excellent research assistance; Wojciech Olszewski also thanks the National Science Foundation for research support (CAREER award SES-0644930).

[†]Department of Economics, Arizona State University, P.O. Box 873806, Tempe, AZ 85287-3806

[‡]Department of Economics, Northwestern University, 2001 Sheridan Rd. Evanston IL 60208-2600

Traditional models of communication, such as Crawford and Sobel (1982), focus on information transmission and do not discuss how the agents argue and what constitute good deliberation skills. Indeed, these issues seem to be complex. Deliberation depends on what the discussants know about the audience and the opponents, as well as on the discussants' experience acquired in the process of learning by doing. Discussants may also attempt to sway the audience by stirring emotions and by taking advantage of other psychological effects.

In this paper, we propose a normative framework which assumes away the psychological side of debates. Our goal is to see whether formal models, based on Bayes' rule and equilibrium concepts, deliver any rules of thumb regarding the way we should argue in debates.

The optimal way to persuade is very much driven by the audience's expectations. Indeed, suppose that we plan to make a certain point or present a certain piece of evidence, but the audience expects us to comment on another aspect, or presenting another piece of evidence. Then, when we make the point as intended, the audience may begin to think that we have little to say on the other aspect, or we lack the evidence they anticipated us to reveal, thus undermining the force of our argument. So, how effective a point really is depends on what we lose by not following the audience's expectation.

It is largely an empirical question what an audience's expectations are in a particular situation. Based on our theoretical study, we are able to say little about them. But we are able to say more about what the discussants want these expectations to be. We take the perspective that the discussants can shape the audience's expectations, or build reputation for arguing (presenting their evidence) in a certain manner. So, the question addressed in this paper is what reputation the discussants may, or should want to build. In this sense, the analysis is more applicable to repeated than to one-shot interactions.

More precisely, we study a model with two discussants and an audience. The discussants have conflicting objectives: each of them wants to convince the audience that the state of the world is the one she likes most, independent of the actual state. Each discussant has a finite number of pieces of hard evidence. Discussants move sequentially, and each of them can reveal at most one piece of evidence at a time. A positive probability of termination makes it possible that discussants will not manage to reveal all the evidence they have. Therefore, the order in which they reveal their evidence is important. Each discussant can commit to a strategy, i.e., which piece of evidence she will reveal, contingent on the evidence she has and the evidence that has been revealed earlier.

We restrict attention to situations in which each discussant has at most two pieces of hard evidence and the debate terminates after each discussant presents at most one piece. We make this simplifying assumption partly for tractability. More importantly, it allows us to minimize the controversies regarding the predictive power of equilibrium analysis since we will study only simple games. Indeed, the optimal strategy of the discussant who moves second, called *the follower*, is derived from Bayes' rule, and deriving the optimal strategy of the discussant who moves first, called *the leader*, in addition require predicting the response of the follower.

Within this simple model, we address two questions. Should discussants always present their strongest evidence first? And should they lean towards presenting the evidence that is independent or the evidence that is correlated with the potential evidence of their opponents?

The answers depend on the discussants' payoffs, which are functions of the audience's posterior belief. For example, suppose the follower's payoff function is concave above the audience's prior belief, convex below the prior and exhibits loss aversion (Kahneman and Tversky, 1979). Then the follower should respond with weaker evidence to weaker evidence presented by the leader, and with stronger evidence to stronger evidence. If the leader's payoff function is also concave above the prior, convex below the prior and exhibits loss aversion (call this the Kahneman and Tversky preference), then she should begin arguing with weaker evidence.

As to whether the discussants should present independent or correlated evidence, we find that if the players have Kahneman and Tversky preferences, then the follower should respond with evidence correlated with what was presented by the leader if the leader's strategy is to present potentially correlated evidence first. On the other hand, the follower should lean towards presenting independent evidence if the leader's strategy is to avoid presenting potentially correlated evidence. Anticipating this response, the leader prefers presenting independent evidence first, at least over presenting evidence that is highly correlated with the evidence of the follower.

Related literature

Earlier work on persuasion games focuses on characterizing conditions under which self-interested parties reveal or fail to reveal all of the verifiable information they have (for example, Milgrom, 1981; Milgrom and Roberts, 1986; and Shin, 1994). We instead consider situations in which the discussants are constrained to reveal a limited amount of evidence, and try to characterize how the discussants should argue.

One inspiration for our paper is the recent work of Glazer and Rubinstein (2001, 2004, 2006). These authors are also interested in optimal rules of persuasion. However, they view a debate as a mechanism by which an uninformed decision maker extracts information from the informed discussants. We restrict attention to a particular game form, one that in our opinion resembles most debates, but allow players to commit to their debate strategies (or to build reputation for debating in a particular manner).

Slightly more remotely related are papers by Dziuda (2007), Gentzkow and Kamenica (2009), Olszewski (2004), and Sher (2008, 2009). These papers also study the choice of arguments or questions in the context of persuasion or eliciting information from self-interested parties. One aspect in Gentzkow and Kamenica (2010) that is similar to our paper is the importance of the curvature of the sender's utility in the receiver's belief, but the two paper study different models (our model has two adversarial sides engaging in debate whereas there is a single sender in Gentzkow and Kamenica) and address different questions (we are interested in finding the best way to argue given an information structure whereas Gentzkow and Kamenica find the optimal information structure for a sender who tries to persuade a receiver).

2 Basic Model

There are two (a priori) equally likely states of the world, $\omega = a$ or b; two agents (discussants), A and B,[1] and an audience. The agents argue in front of the audience that the state is a or b, respectively. Each agent is equipped with at most two signals, or pieces of "hard" evidence, in favor of her claim: s_I and t_I, where $I = A$ or B. Agents move sequentially, presenting one argument at a time. Agent A (the leader) moves first, presenting one piece of evidence available to her (if she has any) according to her choice; agent B (the follower) moves second, also presenting one piece of evidence (if she has any) according to her choice. We assume that the discussion ends here.[2] The agents are not allowed to be silent when they have an argument. That is, whenever an agent presents no evidence in favor of her claim, it is interpreted as the complete lack of arguments. This assumption simplifies the analysis and we believe that allowing for strategic hiding of signals would not affect any of our results.

The audience forms a posterior belief μ about the state of the world. This belief is contingent on the presented arguments, and the strategies of the two agents which the audience correctly anticipates. The agents' preferences are monotonic functions of the audience's posterior. That is, the utility of agent A, denoted by u_A, is an increasing function of the probability assigned by belief μ

[1] We use the pronoun "she" for player A and "he" for player B.

[2] The results would not be affected if we assumed that the discussion ends only with a positive probability, and with the complementary probability the agents who has two pieces of evidence would have a chance to present the second piece.

to state a, denoted by μ_a, and the utility of agent B, denoted by u_B, is an increasing function of the probability assigned by belief μ to state b, denoted by μ_b. This approach is inspired by Geanakoplos, Pearce and Stacchetti (1989). Agents are expected-utility maximizers.

One may argue that the agents' utility should depend only indirectly on the audience's beliefs, through the audience's actions. One example of a utility function u_I $(I = A, B)$ derived from actions is given by: $u_I(\mu_i) = (\mu_i)^3 + 3(\mu_i)^2(1-\mu_i)$. This arises in a situation when the audience consists of three members who have private information on their thresholds for choosing one alternative over the other and the outcome is determined by majority voting.

Indeed, suppose a member votes for alternative a if and only if she believes the probability that the state is a exceeds the threshold t and the agents' prior over each member's threshold t is uniformly distributed on $[0, 1]$ and independent of the others' thresholds. Then, for agent A, when the posterior that the state is a is μ_a, the probability that alternative a is chosen by the three-member audience is $(\mu_a)^3 + 3(\mu_a)^2(1-\mu_a)$. If agent A's utility is linear in the probability that a is chosen, then we can represent it by $u_A(\mu_a) = (\mu_a)^3 + 3(\mu_a)^2(1-\mu_a)$. Note that it is convex on $[0, 1/2]$ and concave on $[1/2, 1]$.[3] (Because of symmetry, B has a similar utility function.)

However, the way in which a debate affects actions in many settings (e.g., political debates) is difficult to determine, even though the discussants often seem to focus on "winning the debate," assuming that this will have desirable impact on the audience's actions. Our model can be viewed as the reduced form of a model in which the members of the audience take actions after the debate.

2.1 Information Structure

The following table exhibits the prior distribution over signals: s_A, s_B, contingent on $\omega = a$.

	$\neg s_B$	s_B
s_A	$(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)$	$(1-\rho)\varepsilon(1-\varepsilon)$
$\neg s_A$	$(1-\rho)\varepsilon(1-\varepsilon)$	$\varepsilon^2+\rho\varepsilon(1-\varepsilon)$

[3] More generally, if the audience consists of an odd number of n members and they vote by majority rule, then the utility function is $u_A(\mu_a) = \sum_{i=0,\ldots,\frac{n+1}{2}} C_n^i (\mu_a)^{n-i}(1-\mu_a)^i$. It is straightforward, although somewhat tedious (and hence omitted), to show that this function is convex on $[0, 1/2]$ and concave on $[1/2, 1]$.

where $0 < \varepsilon < 1/2$, and $0 \leq \rho < 1$. The prior is symmetric contingent on $\omega = b$. That is, if the state is a, then it is more likely that agent A has signal s_A (which in the table is denoted simply by s_A) but agent B does not have signal s_B (this is denoted by $\neg s_B$) than that agent A does not have signal s_A but agent B has signal s_B. And if $\rho \neq 0$, signals s_A and $\neg s_B$ are (positively) correlated conditionally on the state of the world. Contingent on the state being a, the odds that agent B has signal s_B are ε, but contingent in addition on agent A having signal s_A, they are $(1-\rho)\varepsilon$.

Agent $I = A$ or B obtains signal t_I with probability $1-\delta$, where $\delta \leq 1/2$, contingent on $\omega = i$, and she obtains t_I only with probability δ, contingent on $\omega \neq i$. Signals t_A and t_B are conditionally independent, and they are conditionally independent of signals s_A and s_B.

The model easily generalizes to any finite number of agents, any finite number of signals, and more general prior probability distributions over signals. We believe that this extension is a promising subject for future research. However, this simpler version of the model will be sufficient for our analysis.

2.2 Strategies and Equilibria

Each agent I has only one decision to make: whether to present s_I or t_I, when she has both signals at hand. Agent B's decision may depend on the signal that has been revealed by agent A (or the lack thereof).

We assume that agents can precommit to their strategies. In equilibrium, the strategy of the leader must be a best response to the strategy of the follower, and the action of the follower, contingent on each information node (i.e., contingent on s_A, t_A, or no signal having been revealed) must be optimal, given the strategy of the leader. For simplicity, we restrict attention to pure strategies. In Appendix A, we show that many of the results extend to the case when agents are allowed to be playing mixed strategies.[4]

The commitment assumption can be motivated on the grounds of reputation effects. Agents can build reputation for arguing in a specific manner, or revealing their arguments in a specific order. We refer to reputation effects in the intuitive, informal sense, rather than to any specific, existing model. However, we conjecture that our model can be viewed as a reduced form of the following repeated game:

Suppose that agents A and B are long-run players who play repeatedly against a sequence of

[4] Intuitively, mixed strategies seem relevant. One can, for strategic reasons, pretend to be skeptical about strategic arguing in debates, making the pose of revealing arguments in random order.

short-run audiences. At the end of each period, the audiences observe the arguments each agent had in that period. Suppose that the repeated game is one of incomplete information, in which for each pure strategy of each player, there is a small but positive probability of a commitment type who plays this strategy. With the complementary probability, the agents are strategic. Finally, there is a large population (a continuum) of agents A and a large population of agents B, and every period the members of the two populations are matched randomly.

Another motivation for the commitment assumption is costly information acquisition. Imagine that agents must acquire signals, just before presenting them. The cost of acquiring one signal is negligible, but the cost of acquiring any additional signal is prohibitively large. There is a chance that an agent I searching for signal s_I or t_I will fail, i.e., signal $\neg s_I$ or $\neg t_I$, respectively, will be obtained. The audience observes which signal each agent tries to acquire as well as the evidence presented. In Appendix C, we show that the results in this alternative model of costly information acquisition are similar to the results presented in the main text.[5]

3 Advice for the Follower

3.1 Always Respond with your Strongest Argument?

Our first application within the simple persuasion model addresses the following question: Should the follower respond with weak arguments to weak arguments, and with strong arguments to strong arguments? Or, should she always respond with her strongest argument? For example, suppose the opponent gives an argument that does not sound convincing. Should one counterargue (or build a reputation for counterarguing) decisively, or rather disregard the opponent's argument, trying to make the impression that she could give a powerful response, but does not want to get involved in a discussion of low quality?

Formally, suppose that $\rho = 0$, and $\varepsilon < \delta = 1/2$. That is, there is no conditional correlation between different arguments, and arguments s_A and s_B are stronger, i.e., more informative about the state of the world, than arguments t_A and t_B. For simplicity, we assume that t_A and t_B convey no information about the state of the world. (In Appendix B, we extend the model to incorporate the case that the weaker signal is also informative, i.e., $\varepsilon < \delta < 1/2$.)

[5] Notice, however, that the two models are not equivalent. For example, agent A who has both signals s_A and s_B, and contemplates which signal to reveal, has different information regarding agent B's signals than agent A who is deciding which signal to search for.

Suppose first that the leader's strategy is to present the weak signal when she has both and she brings the weak piece of evidence t_A in favor of her claim. The following table exhibits μ_b, the posterior belief of the audience that $\omega = b$ under each strategy of the follower, given the signals at the follower's disposal:

	only t_B	only s_B	both s_B & t_B
str. t_B	$\mu_b = 1/2$	$\mu_b = 1 - \varepsilon$	$\mu_b = 1/2$
str. s_B	$\mu_b = \varepsilon$	$\mu_b = 1 - \varepsilon$	$\mu_b = 1 - \varepsilon$
	1/8	1/8	1/8

Table 1. The leader who plays the strategy to present t_A when having both, presented t_A.

The columns correspond to the events that the follower has: only signal t_B, only signal s_B, and both s_B and t_B, respectively. The event that the follower has no signal is not included because the followers strategy is irrelevant for the audience's posterior contingent on this event. The first two rows correspond to the two strategies of the follower: str. t_B is an abbreviation for the strategy of presenting t_B when he has both signals and str. s_B is an abbreviation for the strategy of presenting s_B when he has both signals. The last row exhibits the ex ante probability of each event.

We call a function $u : [0, 1] \to R$ *concave at* 1/2 if

$$\frac{1}{2}u\left(\frac{1}{2}+x\right)+\frac{1}{2}u\left(\frac{1}{2}-x\right)\leq u\left(\frac{1}{2}\right), \text{ for every } x \in (0, 1/2];$$

and *convex at* 1/2 if

$$\frac{1}{2}u\left(\frac{1}{2}+x\right)+\frac{1}{2}u\left(\frac{1}{2}-x\right)\geq u\left(\frac{1}{2}\right), \text{ for every } x \in (0, 1/2].$$

Proposition 1. (i) *Suppose the leader plays the strategy of presenting the weaker signal in the case of having both signals. If the follower's utility is concave at* 1/2*, she should respond with the weak signal to the weak signal. If the follower's utility is convex at* 1/2*, she should respond with the strong signal to the weak signal.*

Proof. The only two events in which the follower obtains different utilities under different strategies are when she has only signal t_B, and when she has both signals s_B and t_B. Contingent on the union of the two events, the follower's expected utility is

$$\frac{1}{2}u_B(\varepsilon)+\frac{1}{2}u_B(1-\varepsilon),$$

when she plays the strategy of responding with the stronger signal, and

$$u_B(1/2),$$

when she plays the strategy of responding with the weaker signal. Since $\varepsilon/2 + (1-\varepsilon)/2 = 1/2$, the concavity (convexity) of u_B makes the former expression larger (smaller) than the latter expression.■

The proof of Proposition 1 is straightforward. Nevertheless, we find it helpful to explain the argument verbally and intuitively. Notice first that the process of information revelation has the following (general) martingale property: the audience's belief regarding the state of the world, at any point in time, is determined by the strategies of players that have already moved and the signals that have been revealed, and are independent of the strategies of the players who will move in the future.

Thus, the follower's strategy does not affect the expected beliefs of the audience, and affects only the dispersion of the beliefs. This dispersion is higher when the follower plays the strategy of responding with a strong signal, because when she presents a strong signal the audience gets convinced that the state is b, but when she presents a weak signal, the audience infers that she lacks a strong signal, and gets convinced that the state is a. When the follower plays the strategy of responding with a weak signal, the audience does not infer much about the state when this weak signal is presented, so less dispersion is observed.

Similarly, we obtain the tables exhibiting μ_b, the posterior belief of the audience that $\omega = b$ under each strategy of the follower, contingent on other strategies and signals revealed by the leader:

	only t_B	only s_B	both s_B& t_B
str. t_B	$\mu_b = 1-\varepsilon$	$\mu_b = \frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$	$\mu_b = 1-\varepsilon$
str. s_B	$\mu_b = 1/2$	$\mu_b = \frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$	$\mu_b = \frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$
	$\frac{\varepsilon(1-\varepsilon)}{4}$	$\frac{(1-\varepsilon)^2}{8} + \frac{\varepsilon^2}{8}$	$\frac{(1-\varepsilon)^2}{8} + \frac{\varepsilon^2}{8}$

Table 2. The leader who plays the strategy to present s_A when having both, presented t_A.

	only t_B	only s_B	both s_B & t_B
str. t_B	$\mu_b = \varepsilon$	$\mu_b = 1/2$	$\mu_b = \varepsilon$
str. s_B	$\mu_b = \frac{\varepsilon^2}{(1-\varepsilon)^2 + \varepsilon^2}$	$\mu_b = 1/2$	$\mu_b = 1/2$
	$\frac{(1-\varepsilon)^2}{4} + \frac{\varepsilon^2}{4}$	$\frac{\varepsilon(1-\varepsilon)}{2}$	$\frac{\varepsilon(1-\varepsilon)}{2}$

Table 3. The leader who plays the strategy to present s_A when having both, presented s_A.

We omit the table for the case in which the leader plays the strategy to present t_A when having both signals and has presented signal s_A. This because for $\delta = 1/2$, the entries in the first two rows of this table are the same as those in Table 3 and the entries in the third row of this table is equal to those in Table 3 multiplied by 1/2.

Proposition 1 (continued). (ii) *Suppose the leader plays the strategy of presenting the stronger signal in the case of having both signals. If the follower's utility is concave on* $[1/2, 1]$*, she should respond with the weak signal to the weak signal. If the follower's utility is convex on* $[1/2, 1]$*, she should respond with the strong signal to the weak signal.*

(iii) *Independent of the leader's strategy: if the follower's utility is concave on* $[0, 1/2]$*, she should respond with the weak signal to the strong signal; and if the follower's utility is convex on* $[0, 1/2]$*, she should respond with the strong signal to the strong signal.*

The proof and intuition are similar to that for part (i), and so will be omitted.

Since our results, and therefore the advice for the follower, depends on concavity or convexity of her utility function, one would like to know which properties seem to be a better modelling choice. The answer to this question depends, of course, on the way the audience uses the information conveyed by the discussants, or in other words, on the subsequent actions of the audience.

Nevertheless, we conjecture that in many situations of interest the utility functions may be concave on the interval $[1/2, 1]$, convex on the interval $[0, 1/2]$, and have a kink, so that they are concave at 1/2. The example of the utility function $u_I(\mu_i) = (\mu_i)^3 + 3(\mu_i)^2(1-\mu_i)$ illustrates how concavity on $[1/2, 1]$ and convexity on $[0, 1/2]$ arise naturally in some situations. This conjecture is also supported by the intuition that the fifty-fifty belief is “pivotal” in certain applications, and there seems to be a bigger difference between the posterior being equal to .49 and .51 than between .01 and .03, or .97 and .99. Furthermore, given the fifty-fifty prior, the discussants may have some form of loss aversion, analogous to that experimentally demonstrated by Kahneman and Tversky

(1979). Figure 1 contains an example of a utility function with these properties.

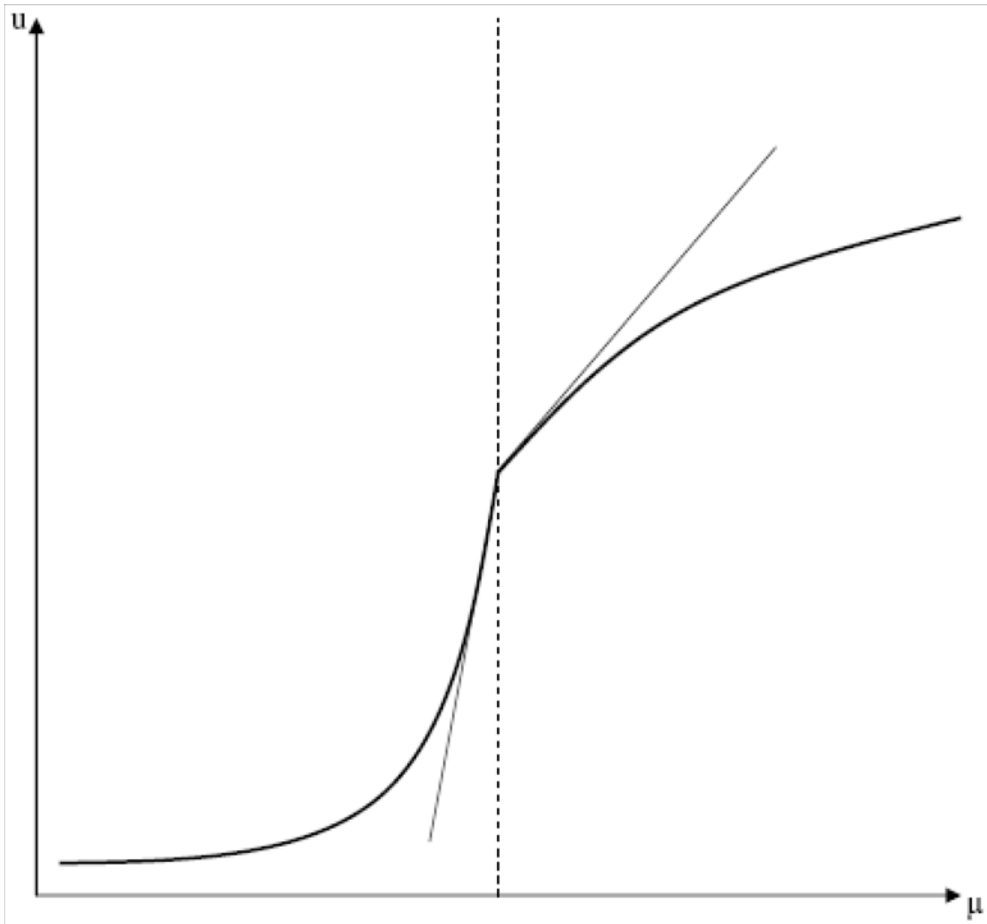

Figure 1. The utility function u, depicted in bold, has a kink at $1/2$.

Under these assumptions on the utility function, the follower should respond by, or try to build a reputation for, presenting weak arguments in response to weak arguments and presenting strong arguments in response to strong arguments, independent of the leader's strategy. It is especially striking that the follower may want to develop a reputation of presenting weak argument as it is less informative. Of course, in a one-shot interaction, responding with a weak argument in case of having both arguments is not an equilibrium strategy without commitment. Even if the audience expects the follower to present the weak argument, it is still better to present the strong argument when it is available as it changes the audience's posterior favorably (and the implicit inference that the weak evidence is absent does not matter for posterior when $\delta = 1/2$). In the less extreme case where δ, is lower than $1/2$, it can be an equilibrium strategy to present the weaker argument even without commitment. As long as t_B is not too weak compared to s_B, presenting the strong argument when the audience expects the weaker one is damaging because the audience will infer that the weaker evidence does not exist. In contrast, if the weaker argument is presented, the audience does not draw any inference about whether the stronger evidence exists. (See Appendix B for details.)

3.2 Correlated Versus Independent Evidence

To illustrate the question we address in this section, recall the following problem discussed in Glazer and Rubinstein (2001): Suppose you are trying to convince the audience that in most capital cities, the level of education has risen recently. Your opponent brings hard evidence showing that the level of education in Bangkok has fallen. Should you respond by bringing similar evidence showing that the level of education has risen from Manila or similar evidence from Mexico City (or perhaps other more distant cities)?

Glazer and Rubinstein argue, and they present evidence from questionnaires, that most people would recommend bringing the evidence from Manila, which seems more similar or related to Bangkok. Similarly, most people would recommend bringing the evidence from Brussels to counter the evidence from Amsterdam. Glazer and Rubinstein also argue that this phenomenon is not confined to cases in which people have implicit beliefs about some correlation between arguments.

Although we believe that correlation between arguments plays an important role in debates, the results of this section show that an explanation of Glazer and Rubinstein's experiment based on correlation pattern is not as straightforward as one might expect, and it requires several qualifications.

Suppose first that the leader's strategy is to present signal s_A when she has both and she brings evidence s_A in favor of her claim. The following table exhibits the posterior belief of the audience, μ_b, under each strategy of the follower, given the signals at the follower's disposal. (Again, we omit the event that the follower has no evidence because the follower's strategy does not affect the posterior in this case.)

	only s_B	only t_B	both s_B & t_B
str. s_B	$\frac{1}{2}$	$\frac{\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right](1-\delta)}{\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right](1-\delta)+\left[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)\right]\delta}$	$\frac{1}{2}$
str. t_B	δ	$\frac{(1-\delta)\varepsilon}{(1-\delta)\varepsilon+\delta(1-\varepsilon)}$	$\frac{(1-\delta)\varepsilon}{(1-\delta)\varepsilon+\delta(1-\varepsilon)}$
	$\frac{1}{2}(1-\rho)\varepsilon(1-\varepsilon)$	$\frac{1}{2}\left[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)\right]\delta+\frac{1}{2}\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right](1-\delta)$	$\frac{1}{2}(1-\rho)\varepsilon(1-\varepsilon)$

Table 4. The leader who plays the strategy to present s_A when having both, presented s_A.

Suppose that $\rho \neq 0$, and $\varepsilon = \delta < 1/2$. That is, arguments s_A and $\neg s_B$ are conditionally correlated, but any single argument is equally informative about the state of the world.

Proposition 2. (i) *Suppose the leader plays the strategy of presenting signal s_A in the case of having both signals, and presents signal s_A. Then, the follower should respond with signal s_B* when having both signals if his utility function is concave on $[0, 1/2]$; and *the follower should respond with signal t_B* when having both signals if his utility function is convex on $[0, 1/2]$.

Proof. By the martingale property of the process of information revelation, or by direct computations, the expected posterior beliefs are equal under the two strategies of the follower,

$$E^{\text{str. } s_B}(\mu_b) = E^{\text{str. } t_B}(\mu_b).$$

Thus, the proposition follows from the fact that for $\delta = \varepsilon$ we have

$$\frac{(1-\delta)\varepsilon}{(1-\delta)\varepsilon + \delta(1-\varepsilon)} = 1/2$$

and

$$\frac{\left[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)\right](1-\delta)}{\left[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)\right](1-\delta) + \left[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)\right]\delta} \in (\delta, 1/2)\,.$$

■

To understand this result, notice that when the leader who plays the strategy s_A (i.e., presenting s_A in the case of having both signals) has presented signal s_A, the audience's posterior μ_b can be no higher than 1/2. If the follower plays strategy s_B (t_B) and presents the signal s_B (t_B), then μ_b is equal to 1/2. But if the follower plays strategy s_B (t_B) but presents the signal t_B (s_B), then μ_b is lower than 1/2. Because s_A and $\neg s_B$ are positively correlated, μ_b is higher if the follower plays strategy s_B but presents the signal t_B than if the follower plays strategy t_B but presents the signal s_B. So, there is more dispersion in the audience's belief if the follower plays the strategy t_B than if he plays strategy s_B. Hence he is better off playing strategy t_B if his utility function is convex on $[0, 1/2]$ and he is better off playing strategy s_B if his utility function is concave on $[0, 1/2]$.

Similarly, we can derive the optimal strategy for the follower, contingent on the other strategy and the signals revealed by the leader. To do this, we obtain the following tables exhibiting the posteriors of the audience μ_b for different strategies and signals of the follower and the leader.

Let

$$\pi = \frac{1}{2}\delta^2\left[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)\right]+\frac{1}{2}(1-\delta)^2\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]$$

$$\pi' = \frac{1}{2}\delta(1-\delta)[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]+\frac{1}{2}(1-\delta)\delta\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]$$

$$\pi'' = \frac{1}{2}[\delta^2+(1-\delta)^2](1-\rho)\varepsilon(1-\varepsilon)$$

and

$$\mu_1 = \frac{\delta[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]}{\delta[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]+(1-\delta)\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]}$$

$$\mu_2 = \frac{\delta^2[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]}{\delta^2[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]+(1-\delta)^2\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]}.$$

	only s_B	only t_B	both s_B & t_B
str. s_B	$1-\delta$	$1-\mu_2$	$1-\delta$
str. t_B	$\frac{1}{2}$	$\frac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon+\delta^2(1-\varepsilon)}$	$\frac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon+\delta^2(1-\varepsilon)}$
	$\delta(1-\delta)(1-\rho)\varepsilon(1-\varepsilon)$	π	π''

Table 5. The leader who plays the strategy to present t_A when having both, presented s_A.

	only s_B	only t_B	both s_B & t_B
str. s_B	$\frac{\delta(1-\varepsilon)}{\delta(1-\varepsilon)+(1-\delta)\varepsilon}$	ε	$\frac{\delta(1-\varepsilon)}{\delta(1-\varepsilon)+(1-\delta)\varepsilon}$
str. t_B	$\frac{\delta^2(1-\varepsilon)}{\delta^2(1-\varepsilon)+(1-\delta)^2\varepsilon}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}\delta(1-\delta)$	$\frac{1}{2}\delta(1-\delta)$	$\frac{1}{2}\delta(1-\delta)$

Table 6. The leader who plays the strategy to present t_A when having both, presented t_A

	only s_B	only t_B	both s_B & t_B
str. s_B	μ_1	$1/2$	μ_1
str. t_B	μ_2	$1-\varepsilon$	$1-\varepsilon$
	π	$(1-\rho)\varepsilon(1-\varepsilon)\delta(1-\delta)$	π'

Table 7. The leader who plays the strategy to present s_A when having both, presented t_A.

Proposition 2 (continued). (ii) *Suppose the leader plays the strategy of presenting signal t_A when having both signals, and presents signal s_A. Then, the follower should respond with signal*

s_B when having both signals if her utility function is concave on $[1/2, 1]$; and the follower should respond with signal t_B when having both signals if her utility function is convex on $[1/2, 1]$.

(iii) *Suppose the leader plays the strategy of presenting signal t_A when having both signals, and presents signal t_A. Then, the follower is indifferent between the two possible strategies.*

(iv) *Suppose the leader plays the strategy of presenting signal s_A when of having both signals, and presents signal t_A. Then, the follower should respond with signal s_B when having both signals if her utility function is concave on $[1/2, 1]$ and concave at $1/2$, and the follower should respond with signal t_B when having both signals if her utility function is convex on $[1/2, 1]$ and convex at $1/2$.*

Proof. In order to obtain part (ii), notice that when $\delta = \varepsilon$ the entries in Table 5 become

$$\frac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon + \delta^2(1-\varepsilon)} = 1 - \delta$$

and

$$\begin{aligned} 1 - \mu_2 &= \frac{(1-\delta)^2[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)]}{(1-\delta)^2[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)] + \delta^2[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)]} \\ &= \frac{(1-\delta)[\delta + \rho(1-\delta)]}{(1-\delta)[\delta + \rho(1-\delta)] + \delta[(1-\delta) + \rho\delta]} \in (1/2, 1-\delta)\,. \end{aligned}$$

This together with the property that

$$E^{\text{str. } s_B}\ (\mu_b) = E^{\text{str. } t_B}\ (\mu_b)$$

yield (ii).

Part (iii) follows immediately from Table 6 for $\delta = \varepsilon$.

In order to obtain part (iv), notice that when $\delta = \varepsilon$, the entries in Table 7 become

$$\begin{aligned} \mu_1 &= \frac{\delta[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)]}{\delta[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)] + (1-\delta)\,[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)]} \\ &= \frac{1 - \varepsilon + \rho\varepsilon}{1+\rho} \in (1/2, 1-\varepsilon) \end{aligned}$$

and

$$\begin{aligned} \mu_2 &= \frac{\delta^2[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)]}{\delta^2[(1-\varepsilon)^2 + \rho\varepsilon(1-\varepsilon)] + (1-\delta)^2\,[\varepsilon^2 + \rho\varepsilon(1-\varepsilon)]} \\ &= \frac{\varepsilon(1-\varepsilon) + \rho\varepsilon^2}{2\varepsilon(1-\varepsilon) + \rho[(1-\varepsilon)^2 + \varepsilon^2]} < 1/2. \end{aligned}$$

Recall that

$$E^{\text{str. } s_B}\ (\mu_b) = E^{\text{str. } t_B}\ (\mu_b).$$

So we compare two lotteries.[6] The lottery generated by strategy s_B has outcomes $1/2$ and μ_1; and the lottery generated by strategy t_B has outcomes μ_2 and $\mu_3 = 1 - \varepsilon$, where

$$\mu_2 < 1/2 < \mu_1 < \mu_3.$$

We believe that a graphical argument is the simplest. This argument is illustrated in Figure 2. By concavity of the utility function on $[1/2, 1]$, the line passing through $(1/2, u_B(1/2))$ and $(\mu_1, u_B(\mu_1))$ is steeper than the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$. By concavity at $1/2$, the line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$ is steeper than the line passing through $(1/2, u_B(1/2))$ and $(1 - \mu_2, u_B(1 - \mu_2))$.

Direct calculations show that

$$1 - \mu_2 < \mu_3.$$

Thus, by concavity on $[1/2, 1]$, the line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$ is steeper than the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$.

If $(\mu_2, u_B(\mu_2))$ belonged to the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$, then $(E(\mu)$, $E^{\text{str. } t_B}\ u_B(\mu))$ would belong to that line as well. However, since the line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$ is steeper than the line passing though $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$, it must be the case that $(E(\mu_b)$, $E^{\text{str. } t_B}\ u_B(\mu_b))$ lies below the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$.

On the other hand, $(E(\mu_b)$, $E^{\text{str. } s_B}\ u_B(\mu_b))$ lies above the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$, because it belongs to the steeper line passing through $(1/2, u_B(1/2))$ and $(\mu_1, u_B(\mu_1))$. This yields that

$$E^{\text{str. } t_B}\ u_B(\mu_b) \leq E^{\text{str. } s_B}\ u_B(\mu_b).$$

■

[6] Strictly speaking these are not lotteries because the ex ante probabilities do not add up to 1, but a simple normalization transforms these into lotteries and the argument is not affected. For expositional convenience, we call them lotteries here and in the proofs of other propositions as well.

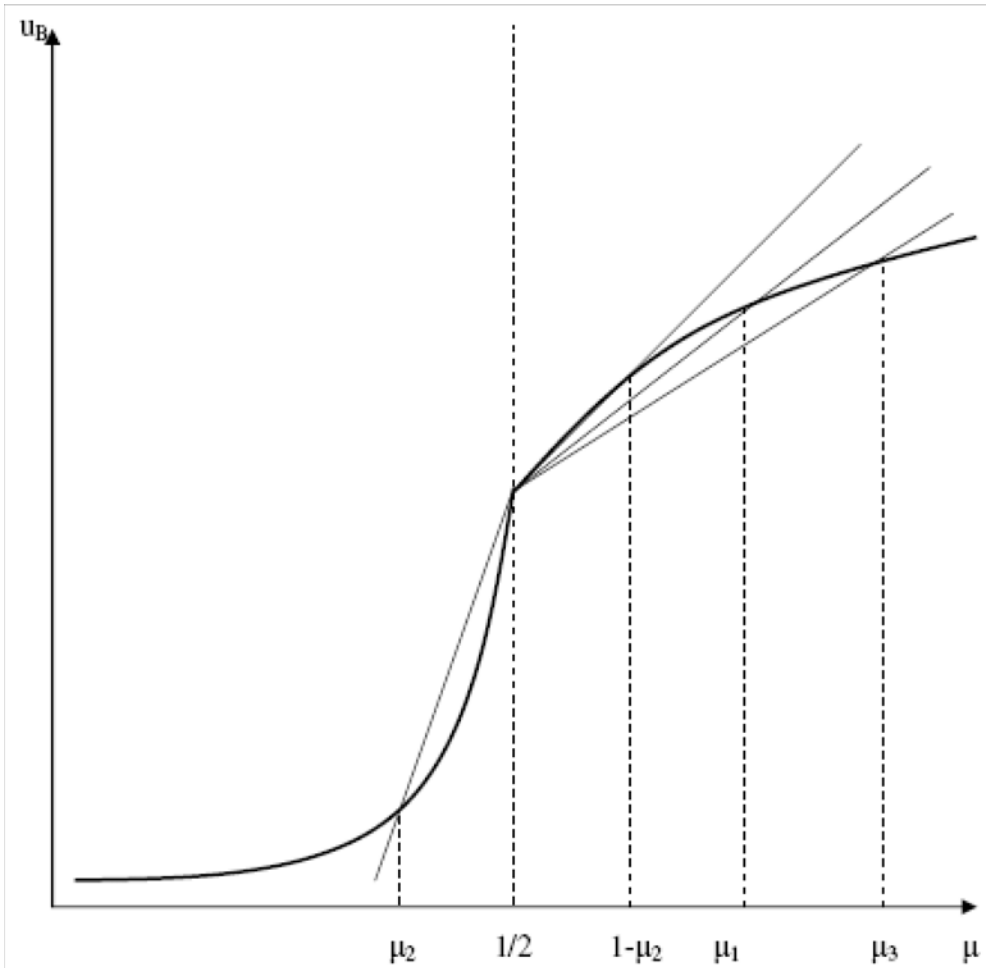

Figure 2. The line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$ is steeper than the line passing through $(1/2, u_B(1/2))$ and $(1-\mu_2, u_B(1-\mu_2))$, which in turn is steeper than the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$.
Also, the line passing through $(1/2, u_B(1/2))$ and $(\mu_1, u_B(\mu_1))$ is steeper than the line passing through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$.

The intuition for part (ii) is similar to that for part (i): if the leader plays the strategy t_A but presents signal s_A, then the audience's posterior μ_b is at least 1/2 as long as the follower shows some evidence. If the follower's strategy is t_B, the posterior μ_b is above 1/2 if she presents t_B and it is equal to 1/2 if she presents s_B. If the follower's strategy is s_B, the posterior μ_b is above 1/2 if she presents s_B and it is still above 1/2 even when she presents t_B. This is because s_A and $\neg s_B$ are positively correlated. Since the leader already reveals that she has evidence s_A but no t_A, even if the follower reveals that she only has evidence t_B and does not have evidence s_B, the audience's posterior is still above 1/2. In short, because s_A and $\neg s_B$ are positively correlated, the follower's strategy s_B induces a lower dispersion in the audience's belief, and hence it is the optimal strategy if her utility function is concave on $[1/2, 1]$, while the strategy t_B is optimal if her utility function is convex on $[1/2, 1]$.

Part (iii) is straightforward: if the leader's strategy is to present the conditionally independent signal t_A in case she has both and she presents t_A, then the distribution of posteriors is the same no matter what strategy the follower plays.

Part (iv) deals with the optimal strategy for the follower if the leader plays the strategy s_A but presents signal t_A, i.e., the leader reveals that she has evidence t_A, but not s_A. If the follower responds with strategy s_B, then the posterior μ_b is higher than 1/2 if she presents s_B and is equal to 1/2 if she presents t_B. If the follower responds with strategy t_B, then the posterior μ_b is higher than 1/2 if she presents t_B, but it is lower than 1/2 if she presents s_B. Because of the positive correlation between $\neg s_A$ and s_B, the revelation that the follower has s_B but no t_B does not sway the posterior favorably enough for it to exceed 1/2. So if the follower's utility exhibits loss aversion at 1/2 and concavity on $[1/2, 1]$, then she should play the strategy s_B. Similarly, if the follower's utility is convex at 1/2 and convex on $[1/2, 1]$, then she should play the strategy t_B.

Although in deriving the results we have assumed commitment to strategies, we would like to point out that the strategies we have identified for the follower are equilibrium strategies even without commitment. For example, if the audience expects that the follower will respond to the leader's signal s_A with strategy t_B, then indeed the follower should optimally choose the strategy t_B, because with this expectation the audience believes that the follower does not have signal t_B when s_B is revealed but believes that the follower may also have s_B when t_B is revealed.

4 Advice for the Leader

4.1 Always Raise Your Strongest Argument?

We now turn to the problem of the leader. What argument should the leader raise in anticipation of the follower's response? In particular, is it always wise for the leader to raise her strongest argument?

In the following analysis, we make the same assumptions as in section 3.1, i.e., $\rho = 0$ and $\varepsilon < \delta = \frac{1}{2}$. We also assume that the follower's utility function is convex on $[0, 1/2]$ and concave on $[1/2, 1]$ and it is concave at 1/2. (Similar results can be derived if we make alternative assumptions on the convexity or concavity of the players' utility functions.) Then, according to Proposition 1, the follower responds with a weak signal to a weak signal, and responds with a strong signal to a strong signal.

The following table exhibits the follower's response to each strategy of the leader, given the signals at the leader's disposal. We omit the case in which the leader has only s_A because according to Proposition 1 (iii), the follower's best response is independent of the leader's strategy in this case and therefore the resulting distributions of posteriors are the same under the different strategies of the leader. We also omit the case in which the leader had neither s_A or t_A, again because the resulting distributions of posteriors are the same under the different strategies of the leader.

	only t_A	both s_A & t_A
str. t_A	follower responds with str. t_B, see Table 1′	follower responds with str. t_B, see Table 1′
str. s_A	follower responds with str. t_B, see Table 2′	follower responds with str. s_B, see Table 3′

.

Tables $1', 2'$ and $3'$ below exhibit μ_a, the audience's posterior that $\omega = a$, contingent on the strategies and signals of the leader and incorporating the response of the follower. They contain the relevant rows from Tables 1, 2 and 3.

	only t_B	only s_B	both s_B & t_B	neither s_B nor t_B
str. t_B	$\mu_a = 1/2$	$\mu_a = \varepsilon$	$\mu_a = 1/2$	$\mu_a = 1 - \varepsilon$
	1/8	1/8	1/8	1/8

Table 1′. The leader who plays strategy t_A has either t_A or both t_A and s_A.

So she presents t_A and the follower responds with strategy t_B.

	only t_B	only s_B	both s_B & t_B	neither s_B nor t_B
str. t_B	$\mu_a = \varepsilon$	$\mu_a = \dfrac{\varepsilon^2}{(1-\varepsilon)^2 + \varepsilon^2}$	$\mu_a = \varepsilon$	$\mu_a = 1/2$
	$\dfrac{\varepsilon(1-\varepsilon)}{4}$	$\dfrac{(1-\varepsilon)^2}{8} + \dfrac{\varepsilon^2}{8}$	$\dfrac{(1-\varepsilon)^2}{8} + \dfrac{\varepsilon^2}{8}$	$\dfrac{\varepsilon(1-\varepsilon)}{4}$

.

Table 2′. The leader who plays strategy s_A has only t_A. So she presents

signal t_A and the follower responds with strategy t_B.

	only t_B	only s_B	both s_B & t_B	neither s_B or t_B
str. s_B	$\mu_a = \dfrac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$	$\mu_a = 1/2$	$\mu_a = 1/2$	$\mu_a = \dfrac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$
	$\dfrac{(1-\varepsilon)^2}{8} + \dfrac{\varepsilon^2}{8}$	$\dfrac{\varepsilon(1-\varepsilon)}{4}$	$\dfrac{\varepsilon(1-\varepsilon)}{4}$	$\dfrac{(1-\varepsilon)^2}{8} + \dfrac{\varepsilon^2}{8}$

Table 3′. The leader who plays strategy s_A has both s_A and t_A.

So she present s_A and the follower responds with strategy s_B.

(The last row in Tables $1' - 3'$ shows the ex ante probabilities of the events, and it is equal to the last row of Table 1-3, respectively, except in Table $3'$ where it is multiplied by 1/2 because Table 3 contains also the event that the leader has only s_A.)

Proposition 3. *Suppose both players' utility functions are convex on* $[0, 1/2]$*, concave on* $[1/2, 1]$ *and concave at* $1/2$*. Then the leader should play strategy* t_A*, i.e., present the weak signal* t_A *when she has both signals.*

Proof. Since the distributions of posteriors are the same in the event "the leader only has s_A" and the event "the leader has neither s_A nor t_A," we only need to consider the events "the leader only has t_A" and the event "the leader has both s_A and t_A."

To prove the proposition, we need to show that

$$\frac{1}{8}u_A(\varepsilon) + \frac{1}{4}u_A(1/2) + \frac{1}{8}u_A(1-\varepsilon)$$

$$\geq \left(\frac{\varepsilon(1-\varepsilon)}{4} + \frac{(1-\varepsilon)^2+\varepsilon^2}{8}\right)u_A(\varepsilon) + \frac{(1-\varepsilon)^2+\varepsilon^2}{8}u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2+\varepsilon^2}\right) +$$
$$+\frac{3\varepsilon(1-\varepsilon)}{4}u_A(1/2) + \frac{(1-\varepsilon)^2+\varepsilon^2}{4}u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}\right).$$

Since $\frac{\varepsilon(1-\varepsilon)}{4} + \frac{(1-\varepsilon)^2+\varepsilon^2}{8} = \frac{1}{8}$, we only need to show that

$$\frac{1}{4}u_A(1/2) + \frac{1}{8}u_A(1-\varepsilon)$$

$$\geq \frac{(1-\varepsilon)^2+\varepsilon^2}{8}u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2+\varepsilon^2}\right) + \frac{3\varepsilon(1-\varepsilon)}{4}u_A(1/2) + \frac{(1-\varepsilon)^2+\varepsilon^2}{4}u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}\right).$$

Concavity of u_A at 1/2 implies that

$$\frac{(1-\varepsilon)^2+\varepsilon^2}{8}u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2+\varepsilon^2}\right) + \frac{(1-\varepsilon)^2+\varepsilon^2}{8}u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}\right) \leq \frac{(1-\varepsilon)^2+\varepsilon^2}{4}u_A(1/2).$$

Since $1/2 < (1-\varepsilon) < \frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}$, concavity of u_A on $[1/2,1]$ implies that

$$\frac{\left((1-\varepsilon)^2+\varepsilon^2\right)}{8} u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}\right) + \frac{(2\varepsilon(1-\varepsilon))}{8} u_A(1/2) \leq \frac{1}{8} u_A(1-\varepsilon).$$

Hence

$$\begin{aligned} & \frac{(1-\varepsilon)^2+\varepsilon^2}{8} u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2+\varepsilon^2}\right) + \frac{3\varepsilon(1-\varepsilon)}{4} u_A(1/2) + \frac{(1-\varepsilon)^2+\varepsilon^2}{4} u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}\right) \\ \leq & \left(\frac{(1-\varepsilon)^2+\varepsilon^2}{4} + \frac{\varepsilon(1-\varepsilon)}{2}\right) u_A(1/2) + \frac{1}{8} u_A(1-\varepsilon) = \frac{1}{4} u_A(1/2) + \frac{1}{8} u_A(1-\varepsilon). \end{aligned}$$

So the leader should play strategy t_A.■

To gain some intuition for this result, note that there are two events in which the leader's strategy matters: the leader only has the weak signal; the leader has both the weak and the strong signals.

If the leader plays the strategy of presenting the weak signal when she has both, then in either one of these two events, she presents the weak signal and the follower responds with the weak strategy, i.e., presenting the weak signal when he has both. So, $\mu_a = 1/2$ if the follower presents a weak signal, $\mu_a = \varepsilon$ if the follower presents a strong signal, and $\mu_a = (1-\varepsilon)$ if the follower presents neither.

If the leader plays the strategy of presenting the strong signal when she has both, then the follower's strategy depends on whether the leader reveals the strong signal or the weak signal.

In particular, if the leader reveals the strong signal, then the follower responds with the strong strategy, i.e., presenting the strong signal when he has both. The posterior is $\mu_a = 1/2$ if the follower presents a strong signal. If the follower fails to present a strong signal, μ_a is above $1/2$; in fact, μ_a in this case is higher than $(1-\varepsilon)$, because there are now two informative signals in favor of state a.

If the leader reveals the weak signal, then the follower responds with a weak strategy, i.e., presenting the weak signal when he has both. The posterior μ_a is equal to ε if the follower presents a weak signal, is lower than ε if the follower presents a strong signal, and is equal to $1/2$ if the follower presents neither signal.

To summarize, if the leader plays the strong strategy, the posteriors induced have more dispersion on $[1/2,1]$ and also around $1/2$ than if she plays the weak strategy. So if the leader's utility function is concave on $[1/2,1]$ and concave at $1/2$, she should play the weak strategy.

4.2 Correlated or Independent Evidence?

From Proposition 2, we have the following table that summarizes the follower's best response contingent on the leader's strategy and the signal that has been presented.

	only s_A	both s_A and t_A	only t_A
str. s_A	follower responds with str. t_B	follower responds with str. t_B	follower responds with str. s_B
str. t_A	follower responds with str. s_B	follower is indifferent	follower is indifferent

Recall that

$$\begin{aligned}\mu_1 &= \frac{\delta[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]}{\delta[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]+(1-\delta)\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]}\\ \mu_2 &= \frac{\delta^2[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]}{\delta^2[(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)]+(1-\delta)^2\left[\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right]}.\end{aligned}$$

Let

$$\begin{aligned}\mu_3 &= \frac{(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)}{(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)+\varepsilon^2}\\ \mu_4 &= \frac{\left((1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)\right)(1-\delta)}{\left((1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)\right)(1-\delta)+\left(\varepsilon^2+\rho\varepsilon(1-\varepsilon)\right)\delta}.\end{aligned}$$

Tables $4', 5', 6'$ and $7'$ below exhibit the audience's posterior μ_a. They contain the relevant rows from Tables 4, 5, 6 and 7.

	only s_B	only t_B	both s_B & t_B	neither s_B or t_B
str. t_B	$1-\delta$	$\frac{\delta(1-\varepsilon)}{(1-\delta)\varepsilon+(1-\varepsilon)\delta}$	$\frac{\delta(1-\varepsilon)}{(1-\delta)\varepsilon+(1-\varepsilon)\delta}$	μ_4
	$\frac{1}{2}(1-\rho)\varepsilon(1-\varepsilon)$	$\frac{1}{2}(1-\varepsilon)^2\delta+\frac{1}{2}\varepsilon^2(1-\delta)+\frac{1}{2}\rho\varepsilon(1-\varepsilon)$	$\frac{1}{2}(1-\rho)\varepsilon(1-\varepsilon)$	$\frac{1}{2}(1-\varepsilon)^2(1-\delta)+\frac{1}{2}\varepsilon^2\delta+\frac{1}{2}\rho\varepsilon(1-\varepsilon)$

Table 4′. The leader who plays the strategy s_A has either s_A or both t_A and s_A.

So she presents s_A and the follower responds by strategy t_B.

	only s_B	only t_B	both s_B & t_B	neither s_B nor t_B
str. s_B	δ	μ_2	δ	μ_3
	$(1-\rho)\varepsilon(1-\varepsilon)\delta(1-\delta)$	π	π''	π'

Table 5′. The leader who plays the strategy t_A has only s_A.

So she presents s_A and the follower responds with strategy s_B.

	only s_B	only t_B	both s_B & t_B	neither s_B nor t_B
str. s_B	$\frac{(1-\delta)\varepsilon}{\delta(1-\varepsilon)+(1-\delta)\varepsilon}$	$1-\varepsilon$	$\frac{(1-\delta)\varepsilon}{\delta(1-\varepsilon)+(1-\delta)\varepsilon}$	$\frac{(1-\delta)^2(1-\varepsilon)}{(1-\delta)^2(1-\varepsilon)+\delta^2\varepsilon}$
	$\frac{1}{2}\delta(1-\delta)$	$\frac{1}{2}\delta(1-\delta)$	$\frac{1}{2}\delta(1-\delta)$	$\frac{1}{2}(1-\delta)^2(1-\varepsilon)+\frac{1}{2}\delta^2\varepsilon$

Table 6′. The leader who plays the strategy t_A has either t_A or both t_A and s_A.

So she presents t_A and the follower is indifferent. Suppose she responds with s_B.

	only s_B	only t_B	both s_B & t_B	neither s_B nor t_B
str. s_B	$1-\mu_1$	$\frac{1}{2}$	$1-\mu_1$	$\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}$
	π	$(1-\rho)\varepsilon(1-\varepsilon)\delta(1-\delta)$	π'	π''

Table 7′. The leader who plays the strategy s_A only has t_A.

So she presents t_A and the follower responds by strategy s_B.

Assume $\delta=\varepsilon$, as in section 3.2. In Table 8 and Table 9, we summarize the audience's posterior μ_a induced by the leader's strategies, incorporating the follower's best responses, and the corresponding ex ante probabilities.

str. s_A	$1-\mu_1$	$\frac{1}{2}$	$1-\delta$	$\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}$	μ_4
	$\pi+\pi'$	$(1-\delta)\,\delta+$ $(1-\rho)\,\delta^2\,(1-\delta)^2$	$\frac{1}{2}(1-\rho)\,\delta\,(1-\delta)$	π''	$\frac{1}{2}(1-\delta)^3+\frac{1}{2}\delta^3$ $+\frac{1}{2}\rho\delta\,(1-\delta)$

Table 8. The leader plays strategy s_A.

str. t_A	δ	μ_2	$\frac{1}{2}$	$1-\delta$	μ_3	$\frac{(1-\delta)^3}{(1-\delta)^3+\delta^3}$
	$(1-\rho)\,\delta^2\,(1-\delta)^2+\pi''$	π	$\delta\,(1-\delta)$	$\frac{1}{2}\delta\,(1-\delta)$	π'	$\frac{1}{2}\left((1-\delta)^3+\delta^3\right)$

Table 9. The leader plays strategy t_A.

We call a function $u:[0,1]\to R$ *strictly concave at* $1/2$ if

$$\frac{1}{2}u\left(\frac{1}{2}+x\right)+\frac{1}{2}u\left(\frac{1}{2}-x\right)<u\left(\frac{1}{2}\right),\text{ for every }x\in(0,1/2];$$

Proposition 4. *Suppose both players' utility functions are continuous on* $[0,1]$*, concave on* $[1/2,1]$ *and strictly concave at* $1/2$*. The leader should play strategy* t_A*, i.e., present the independent signal when she has both, if* ρ *is sufficiently high.*

Proof. Decompose the lottery induced by each strategy into two: one corresponding to the entries of the last two columns of Tables 8 and 9, and the other corresponding to the entries of the first three columns. Consider first the lotteries corresponding to the entries of the last two columns. Since

$$\pi''\left(\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}\right)+\frac{1}{2}\left((1-\delta)^3+\delta^3+\rho\delta(1-\delta)\right)\mu_4=\pi'\mu_3+\left(\frac{1}{2}\left((1-\delta)^3+\delta^3\right)\right)\frac{(1-\delta)^3}{(1-\delta)^3+\delta^3},$$

these lotteries have the same mean.

Since $\frac{1}{2}<\mu_3<\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}<\mu_4<\frac{(1-\delta)^3}{(1-\delta)^3+\delta^3}$ for $\rho\in(0,1)$, it follows that if u_A is concave on $[1/2,1]$, then

$$\begin{aligned}&\pi''u_A\left(\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}\right)+\frac{1}{2}\left((1-\delta)^3+\delta^3+\rho\delta(1-\delta)\right)u_A(\mu_4) \qquad (1)\\ \geq\ &\pi'u_A(\mu_3)+\frac{(1-\delta)^3}{(1-\delta)^3+\delta^3}u_A\left(\frac{1}{2}\left((1-\delta)^3+\delta^3\right)\right).\end{aligned}$$

That is, for the last two columns of Tables 8 and 9, the leader prefers the lottery induced by strategy s_A than that induced by strategy t_A if u_A is concave on $[1/2,1]$.

The other columns of Tables 9 and 10 correspond to lotteries with mean 1/2, since

$$(\pi+\pi')(1-\mu_1)+\left(\frac{1}{2}(1-\rho)\delta(1-\delta)\right)(1-\delta)=\left(\pi+\pi'+\frac{1}{2}(1-\rho)\delta(1-\delta)\right)\frac{1}{2},$$

and

$$\left((1-\rho)\delta^2(1-\delta)^2+\pi''\right)\delta+\pi\mu_2+\frac{1}{2}\delta(1-\delta)(1-\delta)=\left((1-\rho)\delta^2(1-\delta)^2+\pi''+\pi+\frac{1}{2}\delta(1-\delta)\right)\frac{1}{2}.$$

So, whether the leader prefers strategy s_A over t_A depends also on her preference over the two lotteries with mean 1/2. From inequality (1), we have

$$\begin{aligned}E^{\text{str}.s_A}u_A(\mu_a)-E^{\text{str}.t_A}u_A(\mu_a)\geq{}&(\pi+\pi')u_A(1-\mu_1)+\left((1-\delta)\delta+(1-\rho)\delta^2(1-\delta)^2\right)u_A(1/2)\\&+\left(\frac{1}{2}(1-\rho)\delta(1-\delta)\right)u_A(1-\delta)-\left((1-\rho)\delta^2(1-\delta)^2+\pi''\right)u_A(\delta)\\&-\pi u_A(\mu_2)-\delta(1-\delta)u_A(1/2)-\frac{1}{2}\delta(1-\delta)u_A((1-\delta))\end{aligned}$$

Since $\lim_{\rho\to1}\mu_1=\frac{1}{2}$, $\lim_{\rho\to1}\mu_2=\delta$, $\lim_{\rho\to1}\pi=\frac{1}{2}\delta(1-\delta)$, $\lim_{\rho\to1}\pi'=\frac{1}{2}\delta(1-\delta)$, $\lim_{\rho\to1}\pi''=0$, u_A is continuous in μ, and π,π',π'' are continuous in ρ, we have

$$\lim_{\rho \to 1} \left(E^{\text{str}.s_A} u_A(\mu) - E^{\text{str}.t_A} u_A(\mu) \right) \geq$$
$$\delta (1-\delta) u_A (1/2) - \frac{1}{2} \delta (1-\delta) u_A (\delta) - \frac{1}{2} \delta (1-\delta) u_A ((1-\delta)) .$$

If u_A is strictly concave at 1/2, then

$$\delta (1-\delta) u_A (1/2) - \frac{1}{2} \delta (1-\delta) u_A (\delta) - \frac{1}{2} \delta (1-\delta) u_A ((1-\delta)) > 0.$$

Given the continuity of the leader's payoff in ρ, the leader should play strategy s_A when ρ is sufficiently close to 1.■

The proof shows that the leader's preference over the two strategies depends on the concavity or convexity of her utility function on $[1/2, 1]$, and her preference over the lotteries from the first three columns of Table 8 and Table 9, which have mean 1/2. However, it can be shown that her preference over the two lotteries with mean 1/2 is in general not determined by the concavity or convexity of the leader's utility function at 1/2, and we are only able to establish the leader's preference over these two lotteries when the correlation between s_A and $\neg s_B$ is sufficiently high.

To gain some intuition for this result, let us take a closer look at the two pairs of lotteries in turn. First, compare the lottery $\left(\mu_3, \frac{(1-\delta)^3}{(1-\delta)^3+\delta^3}\right)$ induced by strategy t_A and the lottery $\left(\frac{(1-\delta)^2}{\delta^2+(1-\delta)^2}, \mu_4\right)$ induced by strategy s_A. The one induced by strategy t_A has more dispersion because (conditionally) independent evidence in favor of the same state sways the audience belief more than (conditionally and positively) correlated evidence. Because strategy t_A reveals independent evidence and strategy s_A reveals correlated evidence, strategy t_A induces more extreme posteriors than strategy s_A when the leader has favorable evidence and the follower has no evidence in his favor at all. So if the leader's payoff is concave on $[1/2, 1]$, she prefers the lottery induced by strategy s_A.

Next, compare the lotteries that have mean 1/2. It is perhaps easiest to understand the result when $\rho = 1$, i.e., s_A and $\neg s_B$ are perfectly correlated. In this case, the lottery induced by strategy s_A is degenerate, containing only posterior equal to 1/2 whereas the lottery induced by strategy t_A contains posteriors both below and above 1/2. If the leader's preference exhibits loss aversion, the leader prefers the lottery induced by strategy s_A. By continuity the result holds when s_A and s_B are sufficiently correlated.

We assumed strict concavity 1/2 for Proposition 4, but a similar result holds if we instead assume strictly concavity on $[1/2, 1]$. The following example illustrates this.

Example 1: Suppose $\varepsilon = \delta$, and the payoff function of player A is $u_A(\mu_A) = \mu_a^3 + 3\mu_a^2(1-\mu_a)$ and the payoff function of player B is $u_B(\mu_B) = \mu_b^3 + 3\mu_b^2(1-\mu_b)$. So both players' payoff functions are strictly convex on $[0, 1/2]$ and strictly concave on $[1/2, 1]$.

Simple calculation shows that $\lim_{\rho\to1}\mu_1 = \frac{1}{2}$, $\lim_{\rho\to1}\mu_2 = \delta$,

$$\lim_{\rho\to1}\mu_3 = \frac{(1-\delta)^2 + \delta(1-\delta)}{(1-\delta)^2 + 2\delta(1-\delta) + \delta^2} = 1-\delta,$$

$$\lim_{\rho\to1}\mu_4 = \frac{\left((1-\delta)^2 + \delta(1-\delta)\right)(1-\delta)}{\left((1-\delta)^2 + \delta(1-\delta)\right)(1-\delta) + \left(\delta^2 + \delta(1-\delta)\right)\delta} = \frac{(1-\delta)^2}{(1-\delta)^2 + \delta^2},$$

$\lim_{\rho\to1}\pi = \frac{1}{2}\delta(1-\delta)$, $\lim_{\rho\to1}\pi' = \frac{1}{2}\delta(1-\delta)$, $\lim_{\rho\to1}\pi'' = 0$. So as ρ goes to 1, the leader's expected payoff when playing strategy s_A is

$$\begin{aligned}\lim_{\rho\to1} E^{\text{str}.s_A}u_A(\mu_a) &= 2\delta(1-\delta)u_A(1/2) + \left(\delta^2 - \delta + 1/2\right)u_A\left(\frac{(1-\delta)^2}{(1-\delta)^2+\delta^2}\right)\\ &= \delta(1-\delta) + \left(\delta^2-\delta+1/2\right)\left(2\delta^2-2\delta+1\right)^{-3}(\delta-1)^4\left(4\delta^2-2\delta+1\right).\end{aligned}$$

And the leader's expected payoff when playing strategy t_A is

$$\lim_{\rho\to1} E^{\text{str}.t_A}u_A(\mu_a) = \frac{1}{2}\delta(1-\delta)u_A(\delta) + \delta(1-\delta)u_A(1/2) + \delta(1-\delta)u_A(1-\delta)$$
$$+\frac{1}{2}\left((1-\delta)^3+\delta^3\right)u_A\left(\frac{(1-\delta)^3}{(1-\delta)^3+\delta^3}\right) = \frac{1}{2}\delta(1-\delta)\left(2\delta^3-3\delta^2+3\right)$$
$$+\frac{1}{2}\left(3\delta^2-3\delta+1\right)^{-2}(\delta-1)^6\left(3\delta^2-3\delta+2\delta^3+1\right)$$

The difference in the leader's expected payoff between strategy s_A and strategy t_A is

$$\lim_{\rho\to1}\left(E^{\text{str}.s_A}u_A(\mu_a) - E^{\text{str}.t_A}u_A(\mu_a)\right)$$
$$= \frac{1}{2}\left(2\delta^2-2\delta+1\right)^{-2}\left(3\delta^2-3\delta+1\right)^{-2}(1-\delta)^3(1-2\delta)^3\left(\delta^2-\delta+1\right)\left(8\delta^2-8\delta+3\right)\delta^3.$$

Since $\delta < 1/2$, $\delta^2-\delta+1 = (\delta-1/2)^2 + 3/4 > 0$, $8\delta^2-8\delta+3 = 8(\delta-1/2)^2+1 > 0$, we have $\lim_{\rho\to1}(E^{\text{str}.s_A}u_A(\mu_a) - E^{\text{str}.t_A}u_A(\mu_a)) > 0$. So if the correlation between the two pieces of evidence is sufficiently high, the leader should present signal s_A when she has both signals at her disposal.

■

Returning to the Glazer and Rubinstein's (2001) experiment, it indeed follows from Propositions 2 (ii) and 4 that the follower should respond (or try to build the reputation for responding) with signal s_B to signal s_A, as the experimental evidence from Glazer and Rubinstein (2001) suggests, at least when agents have preferences which are concave above the prior, convex below the prior, exhibit some form of loss aversion, and the correlation coefficient is high.

5 Conclusion

We provide a normative framework for the analysis of arguing in public debates. The discussant's payoff depends on the audience's posterior belief in his or her case. We focus on what we find to be the most reasonable case: the payoff is concave above the prior, convex below the prior, and concave at the prior. In this case, the model suggests that discussants should disregard arguments made by their opponents that do not seem relevant or convincing, and respond with strong evidence to strong evidence. Moreover, they should begin a debate with presenting weaker rather than stronger evidence, although, as we show in Appendix B, these conclusions rely on the assumption that weak evidence is sufficiently week. The model also offers an explanation of the experimental findings of Glazer and Rubinstein (2001) based on the assumption that the correlation between the education levels of neighboring or similar cities is high.

We have studied only two applications of persuasion in this paper, but the model provides a framework for exploring other applications as well. One example is whether a discussant should build a reputation for speaking first or wait until the opponent has made an argument. Another example is whether a discussant should build a reputation for presenting more standard versus more original evidence. That is, if the probability of obtaining signals s_A and s_B is relatively low in both states of the world; and the probability of obtaining signals t_A and t_B is higher in both states of the world, should the discussant reveal the more standard evidence s_I or the more original evidence t_I?

6 Appendix A: Mixed Strategies

6.1 Weak versus Strong Evidence

We show that in the "weak vs. strong" case, our results generalize to mixed strategies, at least as far as the payoffs are concave above 1/2, convex below 1/2, and concave at 1/2. Suppose that the leader plays the strategy of presenting signal s_A with probability γ and signal t_A with probability $1-\gamma$ when having both signals, and the follower plays the strategy of presenting s_B with probability λ and t_B with probability $1-\lambda$ when having both signals. Suppose the leader presented signal t_A and the follower presented signal t_B. This happens in state a with probability

$$\begin{aligned}
&[\Pr(t_A, \neg s_A \mid a) + (1-\gamma)\Pr(t_A, s_A \mid a)] \cdot [\Pr(t_B, \neg s_B \mid a) + (1-\lambda)\Pr(t_B, s_B \mid a)] \\
&= \left[\frac{1}{2}\varepsilon + (1-\gamma)\frac{1}{2}(1-\varepsilon)\right] \cdot \left[\frac{1}{2}(1-\varepsilon) + (1-\lambda)\frac{1}{2}\varepsilon\right],
\end{aligned}$$

and in state b with probability

$$\begin{aligned}
&[\Pr(t_A, \neg s_A \mid b) + (1-\gamma)\Pr(t_A, s_A \mid b)] \cdot [\Pr(t_B, \neg s_B \mid b) + (1-\lambda)\Pr(t_B, s_B \mid b)] \\
&= \left[\frac{1}{2}(1-\varepsilon) + (1-\gamma)\frac{1}{2}\varepsilon\right] \cdot \left[\frac{1}{2}\varepsilon + (1-\lambda)\frac{1}{2}(1-\varepsilon)\right].
\end{aligned}$$

Thus, the audience posterior that the state is b, μ_b, when hearing t_A and t_B presented is

$$\mu_b(t_A, t_B) = \frac{(1-\gamma\varepsilon)(1-\lambda(1-\varepsilon))}{2-\gamma-\lambda+2\gamma\lambda\varepsilon(1-\varepsilon)}.$$

Similarly,

$$\mu_b(t_A, s_B) = \frac{(1-\gamma\varepsilon)(1-\varepsilon)}{1-2\gamma\varepsilon(1-\varepsilon)}.$$

The probability that the leader presents signal t_A and the follower has only signal t_B is

$$\begin{aligned}
\Pr(t_A\&\text{ only } t_B) &= \frac{1}{2}\Pr(t_A\&\text{ only } t_B \mid a) + \frac{1}{2}\Pr(t_A\&\text{ only } t_B \mid b) \\
&= \frac{1}{2}[\Pr(t_A, \neg s_A \mid a) + (1-\gamma)\Pr(t_A, s_A \mid a)] \cdot \Pr(t_B, \neg s_B \mid a) \\
&\quad + \frac{1}{2}[\Pr(t_A, \neg s_A \mid b) + (1-\gamma)\Pr(t_A, s_A \mid b)] \cdot \Pr(t_B, \neg s_B \mid b) \\
&= \frac{1}{8}\{1-\gamma[\varepsilon^2 + (1-\varepsilon)^2]\}.
\end{aligned}$$

Similarly, the probability that the leader presents the signal t_A and the follower has only signal s_B, and the probability that the leader presents the signal t_A and the follower has both signal s_B and signal t_B are both equal to

$$\frac{1}{8}[1-2\gamma\varepsilon(1-\varepsilon)].$$

Notice that the audience's posterior μ_b does not depend on λ when the audience is presented signals t_A and s_B and it is higher than 1/2 for any λ. When the audience is presented with signals t_A and t_B, μ_b decreases in λ and it is higher than 1/2 when $\gamma > \lambda$ and lower than 1/2 when $\gamma < \lambda$. By the concavity of the utility function on $[1/2, 1]$ and concavity at 1/2, the follower's best response must be either $\lambda = 0$ or $\lambda = 1$. Therefore, it is optimal to respond with weak signal to weak signal by Proposition 1 (i) and (ii).

Suppose now that the leader has presented signal s_A. If the follower presents signal s_B, the audience posterior is $\mu_b(s_A, s_B) = 1/2$ independent of the strategies. When the audience is presented with signals s_A and t_B,

$$\mu_b(s_A, t_B) = \frac{\varepsilon(1 - \lambda(1 - \varepsilon))}{1 - 2\lambda\varepsilon(1 - \varepsilon)},$$

which is less than 1/2, and decreases in λ. The probability of the latter event decreases with λ. Since ex ante the expected posterior must be equal across all values of λ, it follows from the convexity of the follower's utility function on $[0, 1/2]$ that the follower should respond with signal s_B in the case of having both signals.

Proposition 5. *Suppose the follower's utility is concave on* $[1/2, 1]$*, concave at* $1/2$*, and convex on* $[0, 1/2]$. Independent of the leader's strategy*, the follower should respond to a weak signal with a weak signal, and to a strong signal with a strong signal.*

We now turn to the optimal strategy of the leader. Suppose that the leader, when having both signals, presents signal s_A with probability γ and presents signal t_A with probability $1-\gamma$, and the follower plays the best response strategy, i.e. responds with signal t_B to signal t_A, and with signal s_B to signal s_A.

The audience's posterior belief that the state is a, μ_a, contingent on six possible events are derived as follows.

When the audience is presented with t_A and t_B, which happens with probability $(2-\gamma)/8$, (call this event t_A, t_B), we have

$$\mu_a(t_A, t_B) = \frac{1 - \gamma + \gamma\varepsilon}{2 - \gamma}.$$

When the audience is presented with t_A and s_B, which happens with probability $[1-2\gamma\varepsilon(1-\varepsilon)]/8$, (call this event t_A, s_B), we have

$$\mu_a(t_A, s_B) = \frac{\varepsilon - \gamma\varepsilon + \gamma\varepsilon^2}{1 - 2\gamma\varepsilon + 2\gamma\varepsilon^2}.$$

When the audience is presented with s_A and t_B or with s_A and no evidence, which happens with probability $(1+\gamma)[(1-\varepsilon)^2+\varepsilon^2]/4$ (call this joint event s_A, t_B or s_A, $\varnothing$), we have

$$\mu_a\left(s_A, t_B\right) = \mu\left(s_A, \varnothing\right) = \frac{(1-\varepsilon)^2}{(1-\varepsilon)^2+\varepsilon^2}.$$

When the audience is presented with s_A and s_B, which happens with probability $(1+\gamma)\varepsilon(1-\varepsilon)/2$ (call this event s_A, s_B), we have

$$\mu_a\left(s_A, s_B\right) = 1/2.$$

When the audience is presented with t_A and no evidence, which happens with probability $\{1-\gamma[(1-\varepsilon)^2+\varepsilon^2]\}/8$ (call this event t_A, $\varnothing$), we have

$$\mu_a\left(t_A, \varnothing\right) = \frac{[1-\gamma(1-\varepsilon)](1-\varepsilon)}{1-\gamma[(1-\varepsilon)^2+\varepsilon^2]}.$$

The ex ante expected posterior belief must be equal across all values of γ. We will show that the ex ante expected utility for the leader for any γ is no larger than that for $\gamma=0$. When $\gamma=0$, the audience's posterior is equal to $1/2$ with probability $1/4$, is equal to ε with probability $1/8$, and is equal to $1-\varepsilon$ with probability $1/8$. It can also be equal to $(1-\varepsilon)^2/[(1-\varepsilon)^2+\varepsilon^2]$ with probability $[(1-\varepsilon)^2+\varepsilon^2]/4$, and to $1/2$ with probability $\varepsilon(1-\varepsilon)/2$, but we will disregard the last two terms as they are common for all values of γ.

Claim 1:

$$\frac{1}{8}u_A(1-\varepsilon) \geq \Pr(t_A, \varnothing)\cdot u_A(\mu_a(t_A, \varnothing)) + \frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2}\cdot u_A(\mu_a(s_A, t_B \text{ or } s_A, \varnothing))$$

The superscript star at the symbol Pr indicates that the term $[(1-\varepsilon)^2+\varepsilon^2]/4$, common for all values of γ, was disregarded. The claim follows from the fact that

$$\frac{1}{8}(1-\varepsilon) = \Pr(t_A, \varnothing)\cdot \mu_a(t_A, \varnothing) + \frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2}\cdot \mu_a(s_A, t_B \text{ or } s_A, \varnothing),$$

$$\frac{1}{2} \leq \mu_a(t_A, \varnothing) \leq 1-\varepsilon \leq \mu_a(s_A, t_B \text{ or } s_A, \varnothing),$$

and the concavity of u_A on $[1/2, 1]$.

Represent $\mu_a(t_A, t_B)$ as $\alpha\varepsilon + (1-\alpha)(1/2)$ and $\mu_a(t_A, s_B)$ as $\beta\varepsilon + (1-\beta)\{\varepsilon^2/[\varepsilon^2+(1-\varepsilon)^2]\}$. It turns out that for every γ, we have

$$\alpha \Pr(t_A, t_B) + \beta \Pr(t_A, s_B) = \frac{1}{8} \tag{2}$$

and

$$(1-\beta)\Pr(t_A, s_B) = \frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2}. \tag{3}$$

It now follows from (2) and (3) that

Claim 2:

$$\Pr(t_A, t_B) \cdot u_A(\mu_a(t_A, t_B)) + \Pr(t_A, s_B) \cdot u_A(\mu_a(t_A, s_B)$$
$$\leq \frac{1}{8} \cdot u_A(\varepsilon) + \frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2} \cdot u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2 + \varepsilon^2}\right) + (1-\alpha)\Pr(t_A, t_B) \cdot u_A(1/2)$$

Claim 3: Due to the concavity of u at $1/2$,

$$\frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2} \cdot u_A\left(\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2 + \varepsilon^2}\right) + \frac{\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)}{2} \cdot u_A\left(\frac{\varepsilon^2}{(1-\varepsilon)^2 + \varepsilon^2}\right)$$
$$\leq \overset{*}{\Pr}(s_A, t_B \text{ or } s_A, \varnothing) \cdot u_A(1/2);$$

recall that

$$\frac{(1-\varepsilon)^2}{(1-\varepsilon)^2 + \varepsilon^2} = \mu_a(s_A, t_B \text{ or } s_A, \varnothing).$$

Claim 4:

$$(1-\alpha)\Pr(t_A, t_B) + \Pr{}^*(s_A, t_B \text{ or } s_A, \varnothing) + \Pr(s_A, s_B) = 1/4$$

Indeed, $\Pr(t_A, t_B) + \Pr(t_A, s_B) + \Pr^*(s_A, t_B \text{ or } s_A, \varnothing) + \Pr(s_A, s_B) + \Pr(t_A, \varnothing) = 1/2$. Thus, we obtain the claim from (2), (3) and the fact that $\Pr(t_A, \varnothing) + \Pr^*(s_A, t_B \text{ or } s_A, \varnothing)/2 = 1/8$.

Now, we have that $\Pr(t_A, t_B) \cdot u_A(\mu_a(t_A, t_B)) + \Pr(t_A, s_B) \cdot u_A(\mu_a(t_A, s_B)) + \Pr(t_A, \varnothing) \cdot u_A(\mu_a(t_A, \varnothing)) + \Pr^*(s_A, t_B \text{ or } s_A, \varnothing) \cdot u_A(\mu_a(s_A, t_B \text{ or } s_A, \varnothing)) + \Pr(s_A, s_B) \cdot u_A(\mu_a(s_A, s_B)) \leq \Pr(t_A, t_B) \cdot u_A(\mu_a(t_A, t_B)) + \Pr(t_A, s_B) \cdot u_A(\mu_a(t_A, s_B)) + (1/8) \cdot u_A(1-\varepsilon) + [\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)/2] \cdot u_A(\mu_a(s_A, t_B \text{ or } s_A, \varnothing)) + \Pr(s_A, s_B) \cdot u_A(\mu_a(s_A, s_B)) \leq (1/8) \cdot u_A(\varepsilon) + [\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)/2] \cdot u_A(\varepsilon^2/[(1-\varepsilon)^2 + \varepsilon^2]) + (1-\alpha)\Pr(t_A, t_B) \cdot u_A(1/2) + (1/8) \cdot u_A(1-\varepsilon) + [\Pr^*(s_A, t_B \text{ or } s_A, \varnothing)/2] \cdot u_A(\mu_a(s_A, t_B \text{ or } s_A, \varnothing)) + \Pr(s_A, s_B) \cdot u_A(\mu_a(s_A, s_B)) \leq (1/8) \cdot u_A(\varepsilon) + (1-\alpha)\Pr(t_A, t_B) \cdot u_A(1/2) + (1/8) \cdot u_A(1-\varepsilon) + \Pr^*(s_A, t_B \text{ or } s_A, \varnothing) \cdot u_A(1/2) + \Pr(s_A, s_B) \cdot u_A(\mu_a(s_A, s_B)) \leq (1/8) \cdot u_A(\varepsilon) + (1/8) \cdot u_A(1-\varepsilon) + (1/4) \cdot u_A(1/2)$.

The $i-$th inequality follows from Claim i, for $i = 1, 2, 3, 4$.

Proposition 6. *Suppose the utility functions of both agents are concave on* $[1/2, 1]$*, concave at* $1/2$*, and convex on* $[0, 1/2]$*. The the leader should play strategy* t_A*, i.e., present the weak signal* t_A *when she has both signals.*

6.2 Correlated versus Independent Evidence

This application turned out to be much less tractable when we allow for mixed strategies. We were unable to check whether the model predictions are consistent with the experimental findings of Glazer and Rubinstein even in the case of high ρ. Although we have managed to find out the best response of the follower to any mixed strategy of the leader (some of the best responses turned out indeed to be in mixed strategies), the computation of the optimal strategy of the leader happened to be quite messy. However, the derivation of the leader's optimal strategy is important for explaining Glazer and Rubinstein's findings even in the pure-strategy case.

7 Appendix B: Strong versus Weak, but Informative, Evidence

In this appendix, we discuss how our analysis of weak vs. strong evidence extends to the case in which the weak evidence is also informative about the state of the world. In particular, we assume $\rho = 0$ and $\varepsilon < \delta < 1/2$. By continuity, our results from the main text generalize to the case when δ is sufficiently close to 1/2. So, we are particularly interested in situations in which δ is not too close to 1/2. We restrict attention to the advice for the follower; the leader's case turns out to be much less tractable.

Suppose first that the leader's strategy is to present the weak signal in the case she has both, and she brings the weak piece of evidence t_A in favor of her claim. The follower has at most two pieces of evidence: a strong signal s_B and a weak signal t_B. The following table exhibits the posterior belief of the audience that $\omega = b$ under each strategy of the follower, given the signals at the follower's disposal:

	only t_B	only s_B	both s_B and t_B
str. t_B	$\mu_b = 1/2$	$\mu_b = \dfrac{(1-\varepsilon)\,\delta^2}{(1-\varepsilon)\,\delta^2 + \varepsilon\,(1-\delta)^2}$	$\mu_b = 1/2$
str. s_B	$\mu_b = \varepsilon$	$\mu_b = \dfrac{(1-\varepsilon)\,\delta}{(1-\varepsilon)\,\delta + \varepsilon\,(1-\delta)}$	$\mu_b = \dfrac{(1-\varepsilon)\,\delta}{(1-\varepsilon)\,\delta + \varepsilon\,(1-\delta)}$
	$\dfrac{1}{2}\delta\,(1-\delta)$	$\frac{1}{2}\left((1-\delta)^2\,\varepsilon + \delta^2\,(1-\varepsilon)\right)$	$\dfrac{1}{2}\delta\,(1-\delta)$

Table 10. The leader who plays the strategy to present t_A when having both, presented t_A

Proposition 7. **(i)** *Suppose the leader plays the strategy of presenting the weaker signal in the case of having both signals. If* $\left(\frac{\delta}{1-\delta}\right)^2 \leq \left(\frac{\varepsilon}{1-\varepsilon}\right)$ and *the follower's utility is convex at* 1/2*, convex* *on* $[0,1/2]$ *and concave on* $[1/2,1]$, *then the follower should respond to a weak signal with a strong signal. Moreover, if the follower's utility is strictly convex on* $[0,1/2]$ *or strictly concave on* $[1/2,1]$, then *the follower should strictly prefer responding to a weak signal with a strong signal.*

Proof: As shown in Table 10, to compare the follower's utilities under strategies t_B and strategy s_B, we only need to compare two lotteries with the same mean. The lottery generated by strategy t_B has outcomes 1/2 and $\mu_1 = \frac{(1-\varepsilon)\delta^2}{(1-\varepsilon)\delta^2+\varepsilon(1-\delta)^2}$, which is lower than 1/2 if $\left(\frac{\delta}{1-\delta}\right)^2 \leq \left(\frac{\varepsilon}{1-\varepsilon}\right)$; and the lottery generated by strategy s_B has outcomes $\mu_2 = \varepsilon$ and $\mu_3 = \frac{(1-\varepsilon)\delta}{(1-\varepsilon)\delta+\varepsilon(1-\delta)}$, where

$$\mu_2 = \varepsilon < \mu_1 \leq 1/2 < \mu_3 < 1 - \mu_2$$

Since $1/2 < \mu_3 < 1-\mu_2$, by the concavity of u_B at 1/2 and its concavity on $[1/2,1]$, the line passing through $(\mu_2, u_B(\mu_2))$ and $(\mu_3, u_B(\mu_3))$ is above the line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$. By convexity on $[0,1/2]$, the line passing through $(\mu_1, u_B(\mu_1))$ and $(1/2, u_B(1/2))$ is below than the line passing through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$. This yields that

$$E^{\text{str. } t_B}\ u_B(\mu_b) \leq E^{\text{str. } s_B}\ u_B(\mu_b).$$

Similarly, one obtains the strict inequality when u_B is strictly convex on $[0,1/2]$ or strictly concave on $[1/2,1]$.■

The advice for the follower in the case when δ is "small" is therefore different from that in the case when δ is close 1/2. Indeed, It follows from Proposition 1 (i) that the follower should be indifferent between the two possible responses, when $\delta = 1/2$ and u_B is symmetric around 1/2, i.e.,

$$\frac{1}{2}u\left(\frac{1}{2}+x\right) + \frac{1}{2}u\left(\frac{1}{2}-x\right) = u\left(\frac{1}{2}\right), \text{ for every } x \in (0, 1/2],$$

independently of the shape of u_B on intervals $[0,1/2]$ and $[1/2,1]$. In contrast, the follower should strictly prefer responding to a weak signal with a strong signal when δ is close to ε, and the follower's utility is strictly convex on $[0,1/2]$ or strictly concave on $[1/2,1]$, even when u_B is symmetric around 1/2.

Intuitively, an important difference between the two cases comes from the fact that $\mu_1 \leq 1/2$ when $\left(\frac{\delta}{1-\delta}\right)^2 \leq \left(\frac{\varepsilon}{1-\varepsilon}\right)$ while $\mu_1 > 1/2$ for $\delta = 1/2$; recall that strategy t_B generates a lottery with outcomes 1/2 and μ_1. Thus, different features of the shape of u_B matter for Propositions 1(i) and 7(i).

Note finally that the advice for the follower becomes ambiguous when the follower's utility is concave at 1/2. Informally speaking, there is a trade-off between the effects described in Propositions

1(i) and 7(i), and the advice for the follower depends on the convexity on $[0, 1/2]$ and the concavity on $[1/2, 1]$ compared to the concavity at $1/2$.

Now, suppose the that the leader's strategy is still to present the weak signal in the case she has both, but she brings the strong piece of evidence s_A in favor of her claim. The following table exhibits the posterior belief of the audience that $\omega = b$ under each strategy of the follower, given the signals at the follower's disposal:

	only t_B	only s_B	both s_B and t_B
str. t_B	$\mu_b = \dfrac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon + \delta^2(1-\varepsilon)}$	$\mu_b = 1/2$	$\mu_b = \dfrac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon + \delta^2(1-\varepsilon)}$
str. s_B	$\mu_b = \dfrac{(1-\delta)^2\varepsilon^2}{(1-\delta)^2\varepsilon^2 + \delta^2(1-\varepsilon)^2}$	$\mu_b = 1-\delta$	$\mu_b = 1-\delta$
	$\frac{1}{2}(1-\varepsilon)^2\delta^2 + \frac{1}{2}\varepsilon^2(1-\delta)^2$	$\delta(1-\delta)\varepsilon(1-\varepsilon)$	$\frac{1}{2}\varepsilon(1-\varepsilon)\left(\delta^2 + (1-\delta)^2\right)$

Table 11. The leader who plays the strategy to present t_A when having both, presented s_A.

Proposition 7. (continued) (ii) *Suppose the leader plays the strategy of presenting the weaker signal in the case of having both signals and the follower's utility is concave at* $1/2$*, convex on* $[0, 1/2]$*, concave on* $[1/2, 1]$*. If* δ is sufficiently close to ε*, then the follower should respond to a strong signal with a weak signal.*

Proof: As shown in Table 11, to compare the follower's utilities under strategies t_B and strategy s_B, we only need to compare two lotteries with the same mean. The lottery generated by strategy t_B has outcomes $1/2$ and $\mu_1 = \frac{(1-\delta)^2\varepsilon}{(1-\delta)^2\varepsilon+\delta^2(1-\varepsilon)}$, and the lottery generated by strategy s_B has outcomes $\mu_2 = \frac{(1-\delta)^2\varepsilon^2}{(1-\delta)^2\varepsilon^2+\delta^2(1-\varepsilon)^2}$ and $\mu_3 = 1-\delta$.

When δ is sufficiently close to ε, we have

$$1 - \mu_3 = \delta \leq \mu_2 < 1/2 < \mu_1 < \mu_3.$$

Since u_B is convex on $[0, 1/2]$ and $1 - \mu_3 < \mu_2$, the line that goes through $(\mu_2, u_B(\mu_2))$ and $(1/2, u_B(1/2))$ is steeper than the line that goes through $(1 - \mu_3, u_B(\mu_3))$ and $(1/2, u_B(1/2))$. Since u_B is concave at $1/2$, it follows that the line that goes through $(\mu_2, u_B(\mu_2))$ and $(\mu_3, u_B(\mu_3))$ must be below the line that goes through $(1/2, u_B(1/2))$ and $(\mu_3, u_B(\mu_3))$. Also, since u_B is concave on $[1/2, 1]$ and $\mu_1 < \mu_3$, the line that goes through $(1/2, u_B(1/2))$ and is $(\mu_1, u_B(\mu_1))$ is steeper than the line that goes through $(1/2, u_B(1/2))$ and is $(\mu_3, u_B(\mu_3))$. Hence, the line that goes through $(1/2, u_B(1/2))$ and is $(\mu_1, u_B(\mu_1))$ is above the line that goes through $(\mu_2, u_B(\mu_2))$ and $(\mu_3, u_B(\mu_3))$. This yields that

$$E^{\text{str. } t_B}\ u_B(\mu_b) \geq E^{\text{str. } s_B}\ u_B(\mu_b).$$

■

If the leader's strategy is to present the strong evidence when she has both, then we get similar results as in Proposition 1 (ii) and (iii). That is, the results on the follower's optimal strategy generalize to the case in which the weak piece of evidence is informative if the leader's strategy is to present the strong evidence when she has both.

8 Appendix C: Costly Information Acquisition

In this appendix, we consider an alternative but closely related model of costly information acquisition. We assume that the cost of acquiring one signal is negligible, but the cost of acquiring another signal is prohibitively large. So, each agent acquires just one signal. Agent I searching for signal s_I or t_I may fail, in which case signal $\neg s_I$ or $\neg t_I$, respectively, will be obtained. The signal obtained by the agents are publicly observed. Each agent decides what signal to acquire just before his turn to speak. In particular, the follower observes what the leader obtains before deciding what signal to acquire. The audience observes what signal each agent tries to acquire.

The model of costly information acquisition provides some robustness analysis for the results obtained within the main text. In particular, one feature of the model studied in the main text is that the audience must make inference about the strategies of the agents. The audience's posterior depends not only on the arguments presented by the agents, but also on the audience belief about their strategies. In contrast, the audience's posterior depends only on the presented arguments in the model of this section.

We show that the results of the two models are consistent. However, we also find differences, e.g., ones regarding Glazer and Rubinstein's findings.

8.1 Weak versus Strong Evidence

As in section 3.1, assume that $\rho = 0$ and $\varepsilon < \delta = 1/2$. Suppose the leader's strategy is to acquire the weak signal t_A. Since $\delta = 1/2$, the audience's posterior is independent of what signal the leader obtains. If the follower decides to acquire the weak signal, then the posterior μ_b is $1/2$, no matter what signal the follower obtains. If the follower decides to acquire the strong signal s_B, then the posterior μ_b is ε if the follower obtains $\neg s_B$ and the posterior μ_b is $(1-\varepsilon)$ if the follower obtains s_B.

Proposition 8. **(i)** *Suppose the leader plays the strategy of acquiring the weak signal. Independent of what the leader obtains, if the follower's utility is concave at* $1/2$*, she should acquire the weak signal; if the follower's utility is convex at* $1/2$*, she should acquire the strong signal.*

Suppose the leader's strategy is to acquire the strong signal s_A and she obtains s_A. If the follower decides to acquire the weak signal, then the posterior μ_b is ε, no matter what signal the the follower obtains. If the follower decides to acquire the strong signal, then the posterior μ_b is $\left(\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)^2}\right)$ if the follower obtains $\neg s_B$ and the posterior μ_b is $1/2$ if the follower obtains s_B.

Proposition 8. (continued) (ii) *Suppose the leader plays the strategy of acquiring the strong signal* s_A *and obtains* s_A*. If the follower's utility is convex on* $[0, 1/2]$*, she should acquire the strong signal; if the follower's utility is concave on* $[0, 1/2]$*, she should acquire the weak signal.*

Suppose the leader's strategy is to acquire the strong signal s_A and she obtains $\neg s_A$. If the follower decides to acquire the weak signal, then the posterior μ_b is $(1-\varepsilon)$, no matter what signal the the follower obtains. If the follower decides to acquire the strong signal, then the posterior μ_b is $1/2$ if the follower obtains $\neg s_B$ and the posterior μ_b is $\left(\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}\right)$ if the follower obtains s_B.

Proposition 8. (continued) (iii) *Suppose the leader plays the strategy of acquiring the strong signal and obtains* $\neg s_A$*. If the follower's utility is concave on* $[1/2, 1]$*, she should acquire the weak signal; if the follower's utility is convex on* $[1/2, 1]$*, she should acquire the strong signal.*

We now turn to the problem of the leader. As in section 4.1, we assume that the players' utility functions are convex on $[0, 1/2]$, concave on $[1/2, 1]$ and concave at $1/2$.

Suppose the leader plays the strategy of acquiring the weak signal. Then, as shown in Proposition 8 (i), the follower responds by acquiring the weak signal. So, the posterior μ_a is $1/2$ no matter what signals the players obtain.

Suppose the leader plays the strategy of acquiring the strong signal s_A. If the leader obtains s_A, then, as shown in Proposition 8 (ii), the follower responds by acquiring the strong signal and the posterior μ_a is $1/2$ if the follower obtains s_B and the posterior μ_a is $\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}$ if the follower obtains $\neg s_B$. If the leader obtains $\neg s_A$, then, as shown in Proposition 8 (iii), the follower responds by acquiring the weak signal and hence the posterior μ_a is ε no matter what signal the follower obtains. Since $\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2} > 1-\varepsilon$ and the leader's utility function is convex on $[0, 1/2]$, concave on $[1/2, 1]$ and concave at $1/2$, the line that goes through $(\varepsilon, u_A(\varepsilon))$ and $\left(\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}, u_A\left(\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}\right)\right)$ is below $(1/2, u_A(1/2))$. So we have the following result.

Proposition 9. *Suppose both players' utility functions are convex on* $[0, 1/2]$*, concave on* $[1/2, 1]$ *and concave at* $1/2$*. Then the leader should play the strategy of acquiring the weak signal* t_A*.*

8.2 Correlated versus Independent Evidence

As in section 3.2, assume that $\varepsilon = \delta < 1/2$ and $\rho > 0$.

Suppose the leader's strategy is to acquire the independent signal t_A. Then, no matter what signal she obtains, the follower is indifferent between acquiring t_B and acquiring s_B.

Suppose the leader's strategy is to acquire the correlated signal s_A and she obtains s_A. If the follower decides to acquire the correlated signal s_B, then the posterior μ_b is $\left(\frac{\varepsilon^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)}\right)$ if the follower obtains $\neg s_B$ and the posterior μ_b is $1/2$ if the follower obtains s_B. If the follower decides to acquire the independent signal t_B, then the posterior μ_b is $\left(\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)^2}\right)$ if the follower obtains $\neg t_B$ and the posterior μ_b is $1/2$ if the follower obtains t_B. Since $\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)^2} < \frac{\varepsilon^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)} < 1/2$, there is more dispersion in the audience's posterior if the follower acquires the independent signal.

Proposition 10. (i) *Suppose the leader plays the strategy of acquiring the correlated signal* s_A *and obtains* s_A*. If the follower's utility is convex on* $[0, 1/2]$*, then she should acquire the independent signal* t_B*. If the follower's utility is concave on* $[0, 1/2]$*, then she should acquire the correlated signal* s_B*.*

Suppose the leader's strategy is to acquire the correlated signal s_A and she obtains $\neg s_A$. If the follower decides to acquire the correlated signal s_B, then the posterior μ_b is $1/2$ if the follower obtains $\neg s_B$ and the posterior μ_b is $\left(\frac{(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)}\right)$ if the follower obtains s_B. If the follower decides to acquire the independent signal t_B, then the posterior μ_b is $1/2$ if the follower obtains $\neg t_B$ and the posterior μ_b is $\left(\frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}\right)$ if the follower obtains t_B. Since $1/2 < \frac{(1-\varepsilon)^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)} < \frac{(1-\varepsilon)^2}{\varepsilon^2+(1-\varepsilon)^2}$, there is more dispersion in the audience's posterior if the follower acquires the independent signal.

Proposition 10. (continued) (ii) *Suppose the leader plays the strategy of acquiring the correlated signal* s_A *and obtains* $\neg s_A$*. If the follower's utility is convex on* $[1/2, 1]$*, then she should acquire the independent signal* t_B*. If the follower's utility is concave on* $[1/2, 1]$*, then she should acquire the correlated signal* s_B*.*

We now turn to the problem of the leader. As in section 4.2, assume that each agent's utility is convex on $[0, 1/2]$ and concave on $[1/2, 1]$. As shown in Proposition 10 (i), if the leader plays the strategy of acquiring the signal s_A and obtains s_A, then the follower responds by acquiring the independent signal t_B. If the leader plays the strategy of acquiring the signal t_A and obtains t_A, then the follower is indifferent between acquiring s_B and acquiring t_B. We can therefore without loss of generality assume that the follower responds by acquiring the signal t_B. Hence, the leader faces the same lotteries if she successfully obtains the signal that she decided to acquire, no matter what her strategy is. On the other hand, if the leader plays the strategy of acquiring the signal s_A and obtains $\neg s_A$, then, as shown in Proposition 10 (ii), the follower responds by acquiring the correlated signal s_B. The posterior μ_a is 1/2 if the follower obtains $\neg s_B$ and the posterior μ_a is $\left(\frac{\varepsilon^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)}\right)$ if the follower obtains s_B. If the leader plays the strategy of acquiring the signal t_A and obtains $\neg t_A$, the follower is indifferent between acquiring s_B and acquiring t_B. Again without loss of generality, we assume that the follower responds by acquiring the signal s_B. The posterior μ_a is 1/2 if the follower obtains $\neg s_B$ and the posterior μ_a is $\left(\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)^2}\right)$ if the follower obtains s_B. Since $\frac{\varepsilon^2}{\varepsilon^2+(1-\varepsilon)^2} < \frac{\varepsilon^2+\rho\varepsilon(1-\varepsilon)}{\varepsilon^2+(1-\varepsilon)^2+2\rho\varepsilon(1-\varepsilon)} < 1/2$, the posterior μ_a has more dispersion on $[0, 1/2]$ if the leader's strategy is to acquire the independent signal t_A.

Proposition 11. *Suppose both players' utility functions are convex on* $[0, 1/2]$*, concave on* $[1/2, 1]$*. Then the leader should play the strategy of acquiring signal* t_A*.*

How can the Glazer-Rubinstein findings be reconciled with the model of costly information acquisition? One way is to consider a hypothetical scenario in which the leader happened to play the strategy of acquiring the correlated signal s_A, and obtains signal s_A. This scenario indeed seems to be suggested to the subjects in the experiment. Then, Proposition 10 (i) says that the strategy s_B is preferred by an agent who is risk averse with respect to the audience posterior beliefs on interval $[0, 1/2]$. This is a qualitatively different assumption compared to ones needed to explain the findings within the model from the main text.

9 References

Crawford, V., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, **50**, 1431-1452.

Dziuda, W. (2007): "Strategic Argumentation," Unpublished Manuscript.

Geanakoplos, J., D. Pearce and E. Stacchetti (1989), "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, **1**, 60-79.

Gentzkow, M., and E. Kamenica (2009): "Bayesian Persuasion," *American Economic Review*, forthcoming.

Glazer, J., and A. Rubinstein (2001): "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, **36**, 158-173.

Glazer, J., and A. Rubinstein (2004): "On the Optimal Rules of Persuasion," *Econometrica*, **72**, 1715-1736.

Glazer, J., and A. Rubinstein (2006): "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, **1**, 395-410.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, **47**, 263-291.

Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, **12**, 350-391.

Milgrom, P., and J. Roberts (1986): "Relying on the Information of Interested Parties," *Rand Journal of Economics*, **17**, 18-32.

Olszewski, W. (2004): "Informal Communication," *Journal of Economic Theory* **117**, 180-200.

Sher, I. (2008): "Credibility and Determinism in a Game of Persuasion," Unpublished Manuscript.

Sher, I. (2009): "Persuasion and Limited Communication," Unpublished Manuscript.

Shin, H-S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, **64**, 253-264.