Online Appendix for

The Selective Disclosure of Evidence: An Experiment

Agata Farina — University of Chicago
Guillaume R. Fréchette — New York University
Alessandro Ispano — Université Paris Dauphine
Alessandro Lizzeri — Princeton University
Jacopo Perego — Columbia University

Contents

C Additional Figures and Results

C.1 The Dynamics of Subjects Behavior

We present summary figures of the evolution of behavior, two for senders and two for receivers. In all cases, the variable of interest is plotted for each treatment against blocks of five rounds.

On the sender side, Figure C.1 plots the evolution of the gap between the GPA of the equilibrium message and that of the actual message. In all but one treatment, the final gap is smaller than the starting one: consistent with messages being more selected and less concealed. However, this evolution is substantial only in the treatments where selection is the main force.

Figure C.2 displays the fraction of messages that exactly correspond to the equilibrium prediction. Across all treatments, the majority of messages are consistent with the equilibrium; and by the end more than 75% of the messages in four of the six treatments correspond to equilibrium. Again, the evolution is most noticeable in treatments where selection is the dominant force. By the end, the treatment with the lowest rate of equilibrium messages is K_3, N_3. The deviations in that treatment are driven by some senders concealing lower signals.

On the receiver side, Figure C.3 shows the average belief for messages with a GPA of 4 (overall the most common GPA corresponding to 46% of the sample). The figure illustrates the fact that in most treatments, receivers become more skeptical of messages that have the same GPA. In some treatments the magnitude of these changes is large.

Figure C.4 shows that receivers are learning to better guess the probability of the red urn with experience. The y-axis is the absolute difference between the probability of a red urn, given the GPA of a message, and the guess. Accuracy tends to increase with experience, but overall subjects seem more accurate in treatments with lower N.

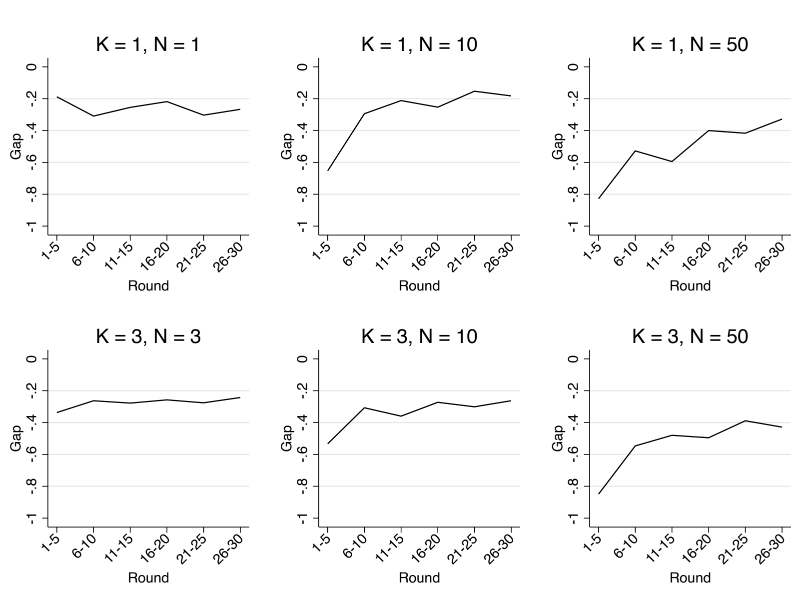


Figure C.1: Average Gap Between GPA of Equilibrium Message and GPA of the Actual Message Over Rounds

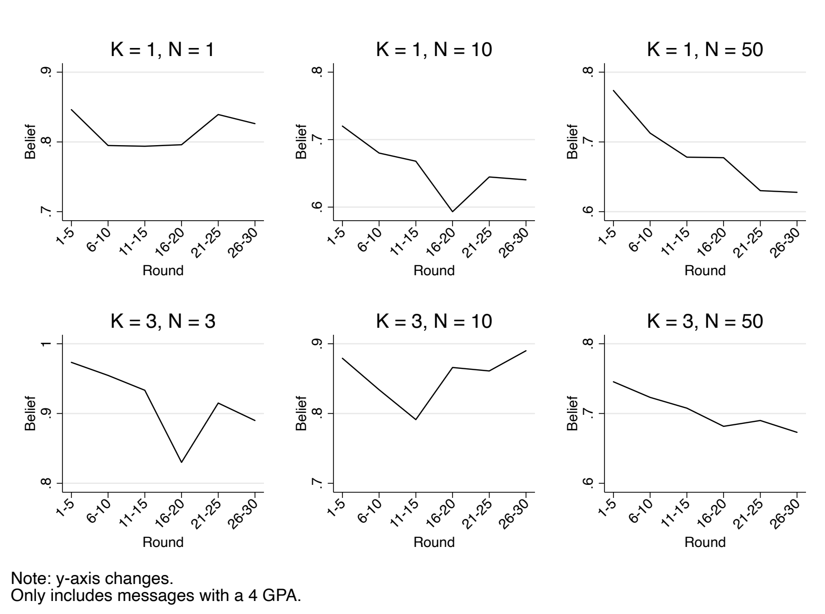


Figure C.3: Average Belief Over Rounds

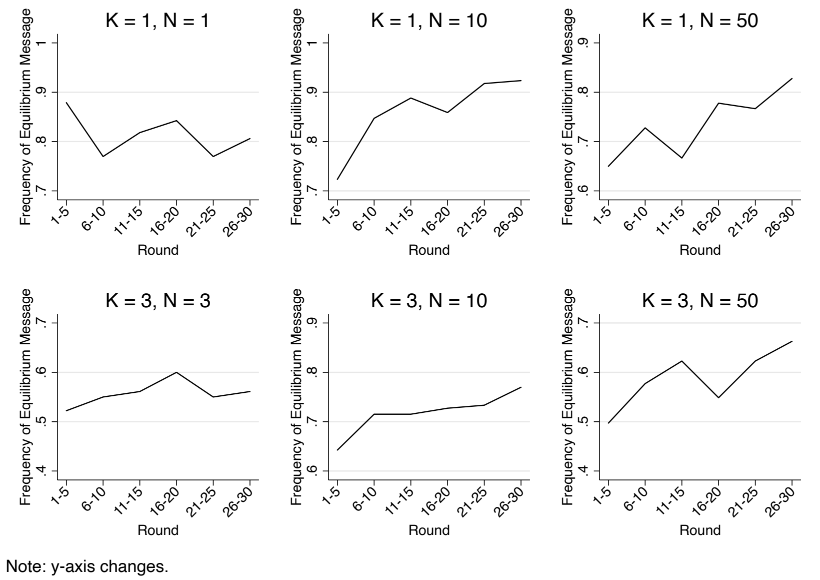


Figure C.2: Equilibrium Messages Over Rounds

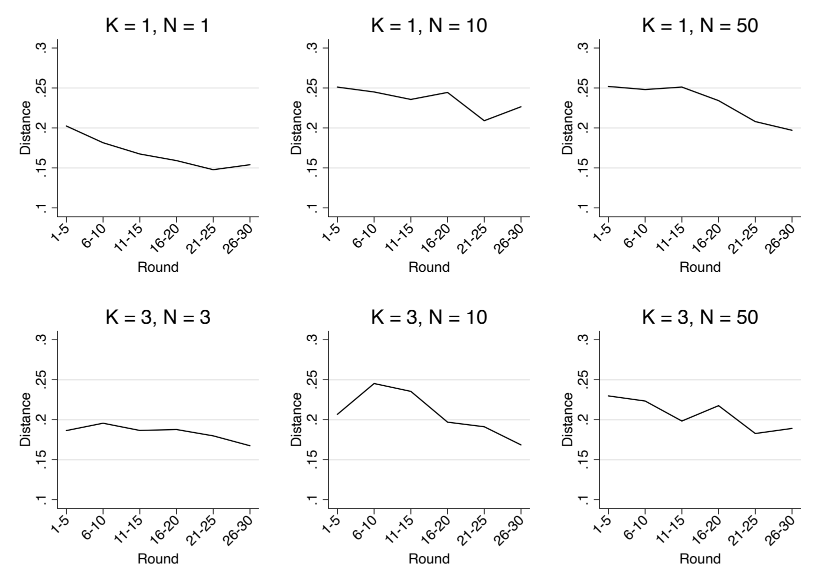
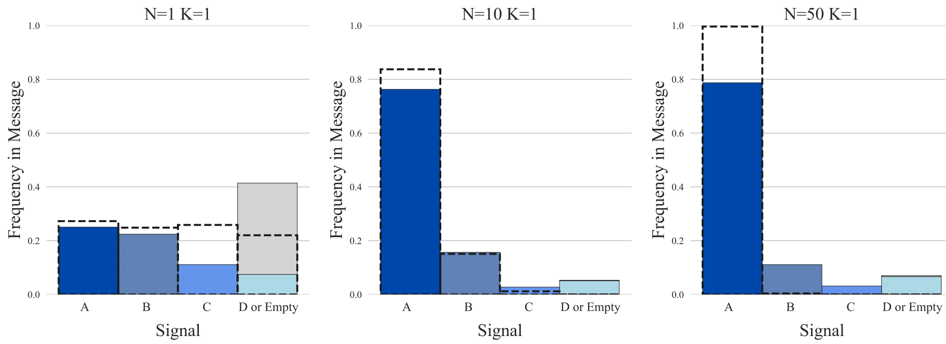
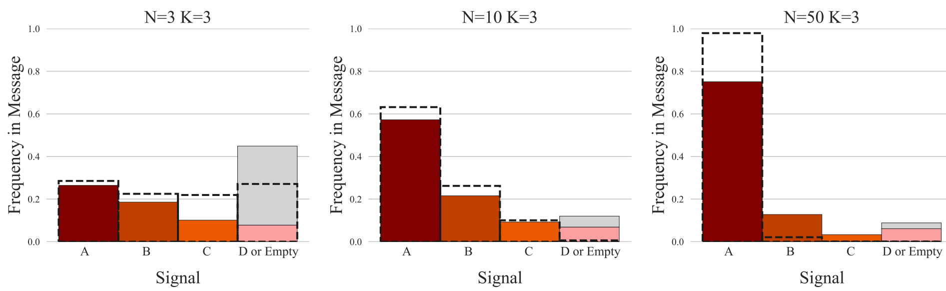


Figure C.4: Accuracy Over Rounds

C.2 Additional Results for Senders



(a) Signal Distribution for $K = 1$



(b) Signal Distribution for $K = 3$

Figure C.5: Signal Distributions

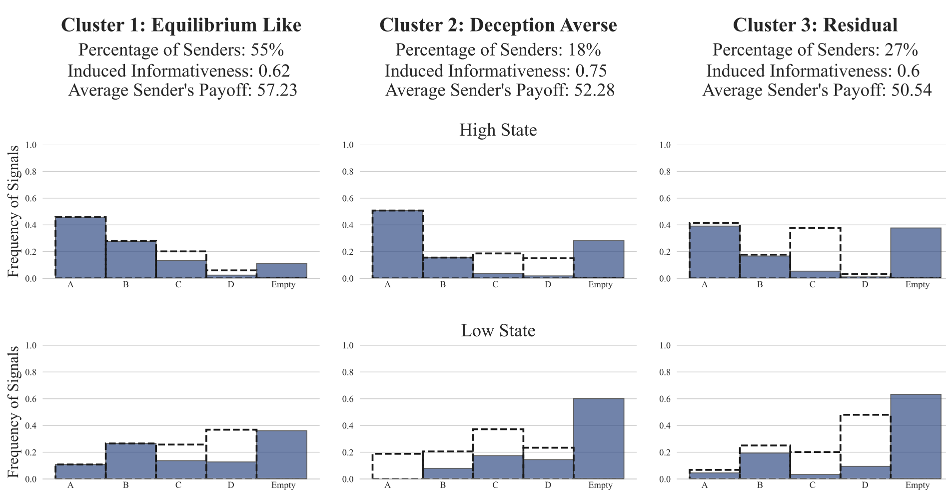


Figure C.6: Sender's Clustering for the Treatment (K_1, N_1)

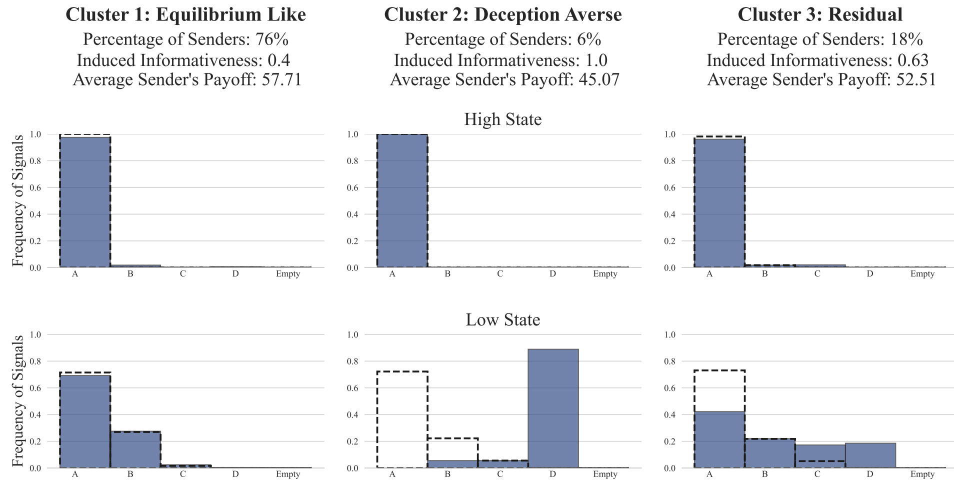


Figure C.7: Sender's Clustering for the Treatment (K_1, N_{10}).

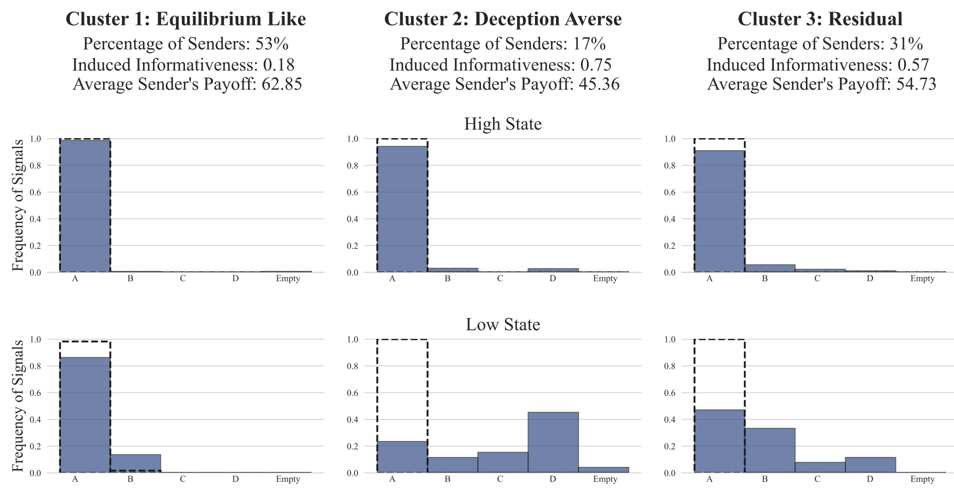


Figure C.8: Sender's Clustering for the Treatment (K_1, N_{50})

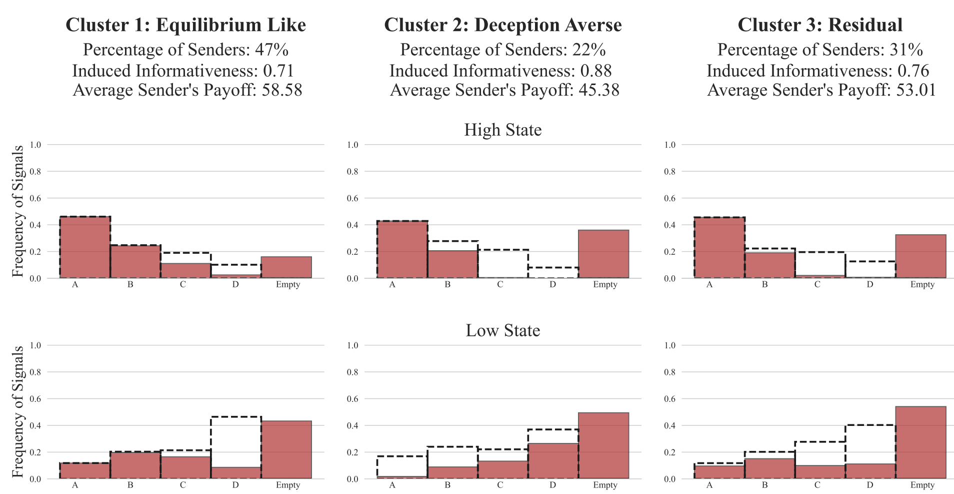


Figure C.9: Sender's Clustering for the Treatment (K_3, N_3)

C.3 Additional Results for Receivers

We study how, on average, the guesses made by the receivers respond to the message GPA. Our theoretical predictions suggest that keeping fixed a value of the GPA, receivers should become more skeptical as N increases, leading to lower guesses for any given GPA. Indeed, a higher value of N allows for more selection on the part of the sender, making favorable messages less informative about the type being high and unfavorable messages more informative about the type being low. In Figures C.10 and C.11, we plot polynomial fits of the actual receivers' guesses and of the guesses of an idealized Bayesian receiver as a function of the message GPA.

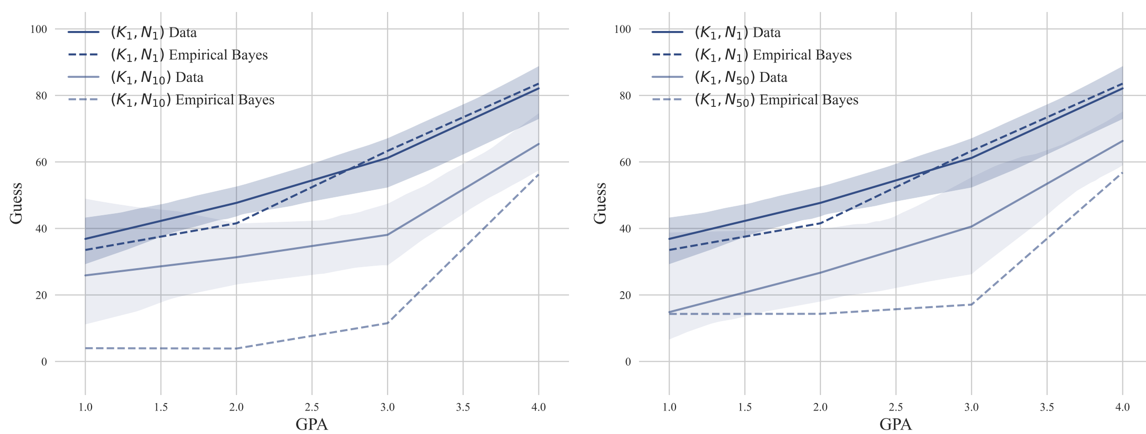


Figure C.10: Receivers' Average Guesses for $K = 1$

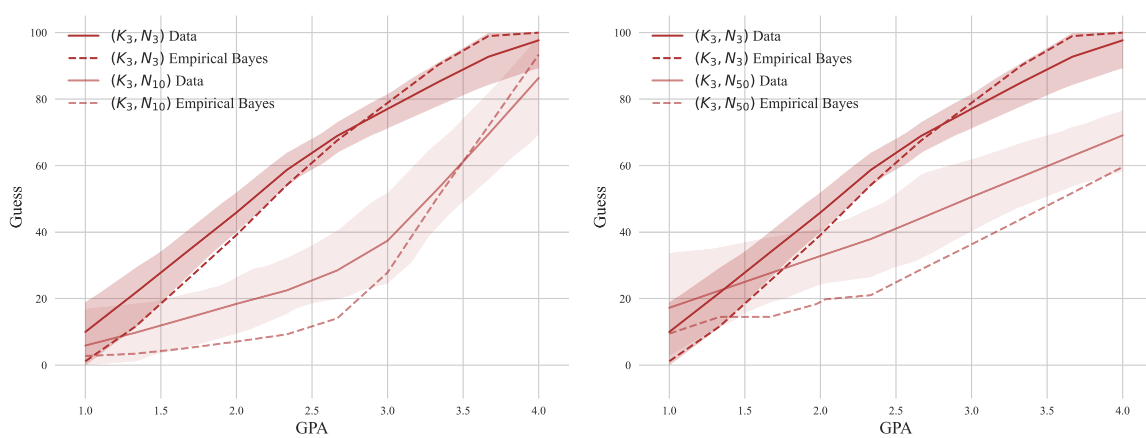


Figure C.11: Receivers' Average Guesses for $K = 3$

The first pattern we can observe is that receivers' guesses are higher when the disclosed information becomes more favorable. The second notable pattern that we can observe in the figure emerges from the comparison between $N = K$ and $N > K$: the receivers' guesses decrease in N for each message GPA and the decrease is particularly pronounced for higher values of the GPA. The only exception is the comparison between (K_3, N_{10}) and (K_3, N_3), where the guesses are similar for high values of the GPA. This suggests that receivers account for the fact that evidence is more selected when N is larger and they adjust their guesses accordingly.

Comparing the guesses of an idealized Bayesian receiver with the behavior of receivers in the data, we note that the qualitative patterns are similar. However, the receivers do not adjust their guesses enough when moving from $N = K$ to $N > K$. When N is large, subjects tend to overguess for every value of the GPA (except for the $N = K = 3$ treatment in which receivers tend to underguess). As discussed in Section 4.2.1, this behavior is in line with the bias of *selection neglect*: when making inferences given the disclosed information receivers may fail to account for the nature of the undisclosed information.

C.4 Additional Material on Receiver Optimism

C.4.1 A Behavioral Model

In this subsection, we present a simple behavioral model and discuss how it can rationalize the empirical findings presented in Section 4.2.2.

In the model, there is a sender and a receiver. The receiver is either sophisticated or naïve, with probabilities that are constant across treatments. The sophisticated receiver and the sender behave as in the theory of Section 2.1. Namely, the sender plays a maximally selective strategy and the sophisticated receiver best responds to it. The naïve receiver, by contrast, takes both concealed and selected evidence at *face value*: she neglects selection when evidence is disclosed and draws no inference when evidence is concealed.

In the discussion below, we argue how this simple model explains the facts reported in Section 4.2.2. For conciseness, we present the arguments for $K = 1$; the case $K = 3$ follows similar logic.

We begin by explaining the first finding. Note that sophisticated receivers never contribute to the response gap (by construction). In contrast, naïve receivers always contribute positively, as the sender plays a maximally selective strategy and they fail to recognize that the evidence they see (or do not see) has been strategically disclosed to inflate their responses. Therefore, in all treatments the average response gap is strictly positive.

Let us now focus on the second finding. Consider treatment (K_1, N_1). By construction, a sophisticated receiver always guesses correctly for all m, and thus never contributes to the response gap. Conversely, a naïve receiver makes no mistakes when $m \neq o$ (response gap is zero), but is overly optimistic when $m = o$ (response gap is positive). Indeed, in this latter case, the naïve receiver's guess equals the unconditional expectation of the state, $\mathbb{E}(\theta)$, whereas the sophisticated receiver would guess a strictly lower value, recognizing that the sender conceals only relatively low signal realizations. As a consequence, the response gap is strictly positive if $m = o$ and it is zero otherwise. Therefore, the response gap conditional on $m = o$ is larger than the response gap when $m \neq o$.

We now explain the third finding. In treatment (K_1, N_1), the response gap conditional on $m = o$ is a convex combination of 0 (the response gap of a sophisticated receiver) and $\mathbb{E}(\theta) - \mathbb{E}(\theta|D) > 0$ (the response gap of a naïve receiver).[2] By contrast, in treatment (K_1, N_{50}), with probability close to 1 the realized message is $m = A$. Conditional on $m = A$, the response gap is a convex combination of 0 (sophisticated receiver) and approximately $\mathbb{E}(\theta|A) - \mathbb{E}(\theta)$ (naïve receiver). Indeed, upon observing $m = A$, the naïve receiver guesses $\mathbb{E}(\theta|A)$, whereas the sophisticated receiver (recognizing selection) guesses approximately $\mathbb{E}(\theta)$, since N is large relative to K. Because in our setting the prior is uniform and f is symmetric, we obtain $\mathbb{E}(\theta|A) - \mathbb{E}(\theta) = \mathbb{E}(\theta) - \mathbb{E}(\theta|D)$. Hence, the response gap in these two cases is the same.

Finally, we explain the fourth finding. Note that $m = A$ occurs almost with probability 1 in (K_1, N_{50}). From Fact 1 we know that the response gap in treatment (K_1, N_1) is entirely explained by responses to message $m = o$, which happens with probability strictly smaller than 1. From Fact 3, we know that the response gap conditional on $m = o$ in (K_1, N_1) is approximately equal to the response gap conditional on $m = A$ in treatment (K_1, N_{50}). Therefore, the average response gap in (K_1, N_1) must be smaller than the average response gap in (K_1, N_{50}).

This model can further explain an additional empirical finding about receiver optimism that we did not report in the main text. Fix $K = 1$ and consider treatments with $K < N$. The response gap conditional on message $m \in \mathcal{M}$ is positive for all m, is non-monotonic in m, and is maximal for the second-highest message, $m = B$ (see Table C.1).[3] This fact highlights a particular pattern that emerges when we study receiver optimism within the same treatment but across different messages (in this respect, it is similar to the exercise we performed for the second finding). We now argue that our behavioral model explains this peculiar pattern. Consider again treatment (K_1, N_{50}). Under our experimental parameters, a naïve receiver would guess approximately 0.81, 0.55, 0.44, and 0.18 after observing messages A, B, C, and D, respec-

[2] Recall that $\mathbb{E}(\theta|s) = p(\theta)f(s|\theta) / \sum_{\theta'} p(\theta')f(s|\theta')$, for all $s \in S$.

[3] Recall that when $K = 1$, the message space $\mathcal{M} = \{A, B, C, D, o\}$ is totally ordered. Performing the same exercise for $K = 3$ is challenging, as there are many more possible messages and fewer data points for each message.

Table C.1: Response gaps by message for (K_1, N_{10}) and (K_1, N_{50})

	$m = A$	$m = B$	$m = C$	$m = D$
$N = 10$	6.47	32.18	18.93	21.81
	389 obs	*80 obs*	*14 obs*	*26 obs*
$N = 50$	7.31	26.53	12.35	15.19
	425 obs	*60 obs*	*17 obs*	*36 obs*

tively. By contrast, a sophisticated receiver would guess approximately $\mathbb{E}(\theta) = \frac{1}{2}$ when $m = A$ and 0 when $m \neq A$. Hence, the response gap is non-monotonic in m and is largest at $m = B$.

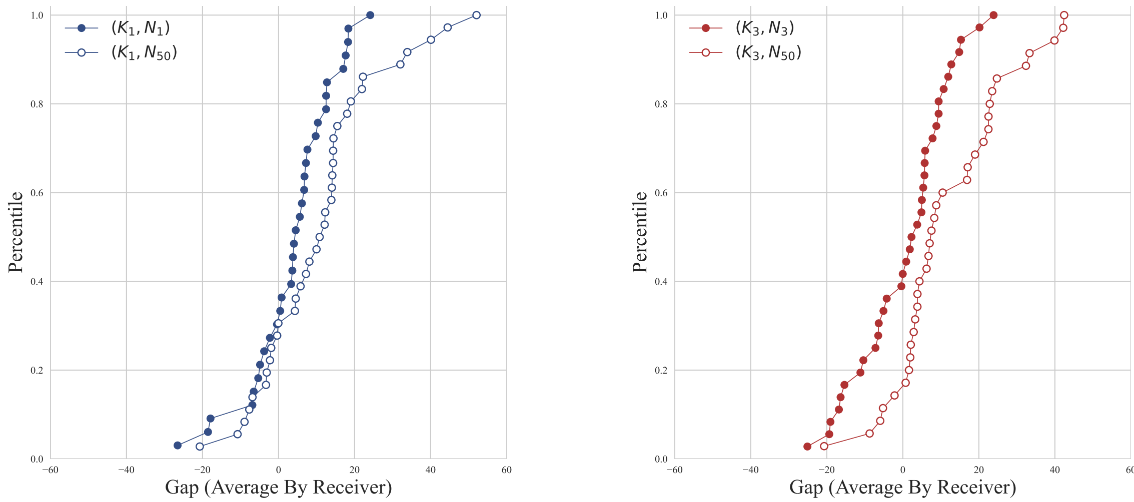
C.4.2 Receiver Optimism Increases with Selection

In Section 4.2.2, we established that the response gap increases in N. That is, receivers make larger mistakes in treatments where selection is the dominant distortion. We investigated the statistical significance of this finding in the regression model of Table 9. In this appendix, we present a robustenss exercise. We show that the same results hold when we look at receiver-level effects as opposed to average effects. Figure C.12 reports the CDF of the receiver-level response gaps, controlling for the message distribution. It shows that, for both values of K, these CDFs increase in a FOSD sense as N increases from K to 50 (p-value < 0.05 for $K = 1$ and p-value < 0.01 for $K = 3$). This indicates that, percentile by percentile, receivers make more mistakes when N is large.

C.5 Mandated Disclosure Simulation

We start by describing our baseline simulation. Assume $K = N = 1$. We use the same procedure discussed in Section 4.1.3 to cluster senders into three groups: 55% are classified as equilibrium types, 18% as deception-averse types, and 27% belong to a residual cluster, which in this case represents a hybrid between the former and the latter (see Figure C.6 in Online Appendix C.2). For each group of senders, we compute their "average" strategy in the form of a probabilistic mapping from the state θ to the messages sent. We then compute the average receiver's guess conditional on each message, as well as the variance of these guesses.

We use these numbers to simulate the play of $100,000$ interactions ($200,000$ subjects). This proceeds along the following steps: 1) Randomly generate θ and one signal s. Independently



(a) (K_1, N_1) and (K_1, N_{50}) (b) (K_3, N_3) and (K_3, N_{50})

Figure C.12: CDF of receivers' response gaps

of (θ, s), generate the sender's behavioral type (one of the three clusters identified above). 2) Randomly generate a message, $m \in \{o, s\}$, according to the average sender strategy estimated above. 3) Given m, generate a random guess from a truncated normal (between 0 and 1). The variance of the truncated normal is determined so that the average guess conditional on the message comes closest to the estimated guessing strategy from above, while the overall variance of guesses matches that observed in the data. 4) Compute the correlation between the state and the guess.

When simulating mandated disclosure, we impose that whatever signal a sender has must be sent. In the text we discuss two cases. One where the behavior of receivers is unchanged. In the other, since mandated disclosure makes it clear how messages relate to evidence, we assume that, on average, receivers correctly estimate the probability that the state is high given the message. However, they still make mistakes in their guesses, and those are simulated following the same process as in our baseline, namely: a truncated normal distribution with the same standard deviation as in step 3 above. We then compute the correlation between the state and those guesses.

D Robustness to an Alternative Design

This section covers an additional experiment with an alternative design to the one presented in the paper.[4] After a brief description (section D.1), we replicate the key tables using this new data set (section D.2). In the dimensions that are comparable to our main design, the results are qualitatively consistent. The only exception is that we find that senders do not overcommunicate, and we explain how some design features are likely to limit this possibility.

D.1 Experimental Design

The experiment featured a disclosure game very similar to the one considered in the paper. The sender observes a randomly drawn number $\theta \in \{1, 2, ...99\}$, where each number is equally likely. She also observes N draws with replacement from a (virtual) urn containing 100 balls, of which θ are white and the rest are black. Thus, θ is also the percentage probability of drawing a white ball. For each ball she observes, she must decide whether to show it to the receiver, and she can show a maximum of K balls in total. The receiver observes the shown balls, if any, displayed in a random order, and makes a guess a about θ. The payoffs of the sender and the receiver are $3 + 8a/100$ and $10 - 8(\frac{\theta - a}{100})^2$, respectively.

The design consisted of three treatments that varied K and N, as summarized in Table D.2. Theoretical predictions on equilibrium informativeness are in Figure D.13.[5] Each subject participated in only one treatment and was randomly assigned the role of sender or receiver for the entire experiment. She played fifteen rounds of the game with a randomly selected participant in the opposite role in each round. At the end of the round, both subjects observed the sender's type θ, the receiver's guess a, and the payoffs, but the receiver did not learn the color of the undisclosed balls. More details about the design can be found in Ispano (2024).

The experiment was completely computerized and implemented with O-tree. All experimental sessions were conducted at the LEEM experimental laboratory of Montpellier in 2023. Each of the fifteen sessions, five per treatment, had an average of 18.5 participants and lasted about one hour, including reading of instructions and payment. The average payment, including a 5€ show-up fee, was 15.74€, and earnings ranged from 10.29€ to 18.84€. The experiment was pre-registered on AsPredicted.org (#144222).

[4]This alternative design originally appeared in Ispano (2024). As mentioned in our acknowledgments, that paper is now merged into the current one.

[5]For the sake of precision, these predictions obtain when θ is drawn from a continuous uniform distribution, in which case the equilibrium informativeness takes a simple closed form for any K and N.

	$N = K$	$N = 6$
$K = 2$	(K_2, N_2)	(K_2, N_6)
$K = 6$	(K_6, N_6)	

Table D.2: Treatments' Denominations

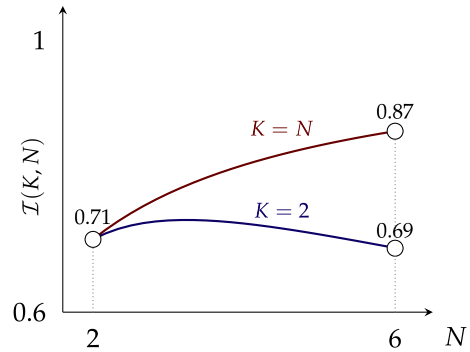


Figure D.13: Predicted Informativeness

D.2 Results

In the analysis, we use the same specifications and statistical tests as in the main body of the paper. Likewise, we focus on the second half of the rounds, i.e., from round 9 to 15.

D.2.1 Senders' Behavior

What Evidence Do Senders Disclose?

Mirroring Table 3, Table D.3 reports the average number of disclosed signals as a fraction of K in each treatment. In each treatment, the number of disclosed signals is significantly lower than K (p-value < 0.01), and increases with selection opportunities, i.e., it is significantly higher in (K_2, N_6) than in (K_2, N_2) (p-value < 0.01). Likewise, mirroring Table 4, Table D.4 reports the MGPA of the senders' messages. In computing the GPA, a disclosed white ball has a value of 4, while a disclosed black ball and a concealed ball have a value of 1. The MGPA increases significantly in N, i.e., from (K_2, N_2) to (K_2, N_6) (p-value < 0.01), and decreases significantly in K, i.e., from (K_2, N_6) to (K_6, N_6) (p-value < 0.01). Also, the MGPA is not significantly different between the two treatments in which $N = K$. Finally, in all treatments, the MGPA is significantly lower than the theoretically predicted one (p-value < 0.01).

How Much Information Do Senders Transmit?

As can be seen in Table D.5, which mirrors Table 10, for all treatment variations of K and N, the average senders' informativeness moves in the directions predicted by the theory. The increase in senders' informativeness from (K_2, N_6) to (K_6, N_6) is significant (p-value < 0.05), and so is the increase from (K_2, N_2) to (K_6, N_6) (p-value < 0.05). The decrease from (K_2, N_2) to (K_2, N_6) is not, which is however predicted to be small and insignificant even by the the-

Table D.3: The average number of signals disclosed as a fraction of K.

		$N = K$	$N = 6$
$K = 2$	Data	59.8%	80.8%
	Predictions	[48.8%, 100%]	[81.6%, 100%]
$K = 6$	Data	54.1%	
	Predictions	[53.1%, 100%]	

Table D.4: Mean grade point average (MGPA) induced by senders' messages.

		$N = K$	$N = 6$
$K = 2$	Data	2.24	3.05
	Predictions	2.46	3.45
$K = 6$	Data	2.22	
	Predictions	2.59	

Table D.5: Overall Informativeness, Senders' Informativeness, and Theoretical Predictions

		$N = K$	$N = 6$
$K = 2$	Overall Informativeness	0.42	0.21
	Senders' Informativeness	0.64	0.57
	Predictions	0.69	0.67
$K = 6$	Overall Informativeness	0.48	
	Senders' Informativeness	0.69	
	Predictions	0.85	

ory. Finally, the senders' informativeness is not significantly different from the theoretical informativeness in treatments (K_2, N_2) and (K_2, N_6), while it is significantly lower in treatment (K_6, N_6) (p-value < 0.05). That is, in this treatment, senders undercommunicate. These results contrast with those of the main experiment, where overcommunication is found for treatments with large N. Among design differences that may explain this discrepancy, e.g., equilibrium informativeness is lower to begin with for large N and a higher number of rounds may facilitate communication, we suspect that the richness of the message space relative to the type space also plays a role. Indeed, as argued in the paper, deception aversion is responsible for overcommunication. With only two types, a low type can easily separate by sending the lowest signal, especially when N is large. Such a strategy has no clear counterpart in this setting.

Evidence of Deception Aversion. Table D.6 reports the results of a basic test of deception aversion along the lines of the analysis of section 4.1.3. Namely, in each treatment, the GPA is regressed on the sender's type after controlling for the favorableness of signals as measured by

Table D.6: GPA as a Function of Type Controlling for Theoretical GPA

	(K_2, N_2)	(K_2, N_6)	(K_6, N_6)
	(1)	(2)	(3)
	GPA	GPA	GPA
Theoretical GPA	0.824***	0.732***	0.682***
	(0.0324)	(0.104)	(0.138)
Sender type	0.00258**	0.00735*	0.00405
	(0.00112)	(0.00379)	(0.00472)
Constant	0.0864	0.143	0.235**
	(0.0554)	(0.110)	(0.102)
Observations	322	343	308
Subjects	46	49	44

Notes: * $(p < 0.1)$, ** $(p < 0.05)$, *** $(p < 0.01)$. Standard errors, in parentheses, are clustered at the session level.

the maximum GPA that senders could generate, i.e., the theoretically predicted one. Contrary to theoretical predictions and consistent with deception aversion, in treatments (K_2, N_2) and (K_2, N_6) the GPA increases significantly with the type.

D.2.2 Receivers' Behavior

How Do Receivers Respond to Selected Evidence?

Mirroring Tables 6 and 7, Tables D.7 and D.8 examine how receivers' guesses and the empirical optimal guesses vary with N for fixed K, and with K for fixed N, respectively, while controlling for the GPA of the message. Consistent with the theory, guesses decrease with N and increase with K, even though, unlike in the main experiment, the effect of N for receivers' guesses is not statistically significant, possibly due to the smaller variation in N in this setting. As in the main experiment, comparing the two columns in each table shows that receivers adjust their guesses less than they should.

Receiver Optimism

Mirroring Table 8, Table D.9a reports the average response gap, i.e., the difference between the receiver's guess and the empirical optimal guess. As in the main experiment, when considering

Table D.7: Regression Results of Receivers' Responses for $K = 2$

	$K = 2$	
	(1)	(2)
	Guess	Optimal Guess
GPA	8.648***	13.46***
	(1.130)	(0.587)
D_{N_6}	-2.389	-6.236***
	(1.937)	(0.767)
Constant	30.32***	17.50***
	(2.544)	(1.336)
Observations	665	665
Subjects	95	

Notes: * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors, in parentheses, are clustered at the session level. Regression (2) does not include random effects.

Table D.8: Regression Results of Receivers' Responses for $N = 6$

	$N = 6$	
	(1)	(2)
	Guess	Optimal Guess
GPA	10.51***	14.28***
	(1.934)	(1.030)
D_{K_6}	9.404***	12.12***
	(2.294)	(1.167)
Constant	22.26***	8.754**
	(5.931)	(3.290)
Observations	651	651
Subjects	93	

Notes: * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors, in parentheses, are clustered at the session level. Regression (2) does not include random effects.

all messages the response gap is positive in each treatment, although here it statistically different from zero only in treatment (K_2, N_2) (p-value < 0.05). This difference may reflect that treatments with large N relative to K are not considered, so receivers encounter less favorable evidence overall, and that binary signals simplify equilibrium reasoning about undisclosed evidence. These features also blur the distinction between selection and concealment. Combined with receivers updating too conservatively even when messages are maximally favorable, this may explain why patterns are less clear when focusing on messages of length K or on empty messages only. Response gaps are largest for empty messages in treatments (K_2, N_2) and (K_2, N_6), and for messages of length K, the gap is not significantly negative only in treatment (K_2, N_6). Besides, mirroring Table 9, Table D.9b documents how, as in the main experiment, the response gap increases significantly with N when fixing K and controlling for the GPA.

How Much Information Do Receivers Absorb?

Finally, Table D.5 above also reports the overall informativeness resulting from the interaction between senders and receivers. As in the main experiment, by construction, overall infor-

Table D.9: Receivers' Response Gaps

(a) Average Response Gaps by Treatment

		$K = N$		$K < N$
		(K_2, N_2)	(K_6, N_6)	(K_2, N_6)
(i)	All messages	2.05^{**}	2.43	1.99
		(322 obs)	*(308 obs)*	*(343 obs)*
(ii)	Length-K	-2.87^{**}	-2.30^{**}	-1.29
		(133 obs)	*(54 obs)*	*(258 obs)*
(iii)	Empty	8.01^{**}	-4.90	8.02^{**}
		(70 obs)	*(42 obs)*	*(47 obs)*

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

(b) OLS on Response Gaps

	$K = 2$
	Gap
GPA	-4.828***
	(0.916)
D_{N_6}	3.861**
	(1.940)
Constant	12.86***
	(2.131)
Observations	665
Subjects	95

Notes: * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors, in parentheses, are clustered at the session level.

mativeness is lower than senders' informativeness. Moreover, overall informativeness moves in the directions predicted by the theory. The increase in informativeness from (K_2, N_6) to (K_6, N_6) is significant (p-value < 0.01), as is the increase from (K_2, N_2) to (K_6, N_6) (p-value < 0.05). Interestingly, even the decrease from (K_2, N_2) to (K_2, N_6), which is predicted to be insignificant by the theory, is significant (p-value < 0.05), which is consistent with selection neglect by receivers.

E Supplementary Proofs, Examples and Extensions

E.1 Proof of Lemma B.1

***Proof of Lemma** B.1.* [6] Denote by $z = (\theta, y_1, \ldots, y_n)$ a generic realization of these random variables. To show that the random variables $(\theta, y_1, \ldots, y_n)$ are affiliated, we need to show that, for any $z, z' \in \Theta \times S^N$,

$$g(z \vee z')g(z \wedge z') \geq g(z)g(z'),$$

[6] For this proof, we refer to the notation introduced at the beginning of Section B.1.

where $z \vee z'$ and $z \wedge z'$ are the component-wise max and min of the two vectors (Theorem 24 in Milgrom and Weber, 1982). $\Theta \times S^n$ is a lattice, thus the operations $\vee$ and $\wedge$ are well-defined. Additionally, letting $z = (\theta, y)$ and $z' = (\theta', y')$, if y and y' are weakly decreasing, $z \vee z'$ and $z \wedge z'$ are also weakly decreasing.

Since the product of affiliated functions is affiliated (Theorem 1(ii) in Milgrom and Weber, 1982), it is enough to show that the functions $p(\theta)f(y|\theta)$ and $B(y)\mathbb{1}_{\{y_1 \geq \ldots \geq y_n\}}$ are affiliated.

For the former, fixing $z = (\theta, y)$ and $z = (\theta', y')$, we need to show that

$$p(\theta \vee \theta')f(y \vee y'|\theta \vee \theta')p(\theta \wedge \theta')f(y \wedge y'|\theta \wedge \theta') \geq p(\theta)f(y|\theta)p(\theta')f(y'|\theta').$$

Dividing by $p(\theta)p(\theta')$ and using the fact that $f(y|\theta) = \prod_i f(y_i|\theta)$, the expression above becomes

$$\prod_i f(y_i \vee y'_i|\theta \vee \theta') \prod_i f(y_i \wedge y'_i|\theta \wedge \theta') \geq \prod_i f(y_i|\theta) \prod_i f(y'_i|\theta').$$

It suffices to show that, for all $i \in \{1 \ldots, n\}$,

$$f(y_i \vee y'_i|\theta \vee \theta')f(y_i \wedge y'_i|\theta \wedge \theta') \geq f(y_i|\theta)f(y'_i|\theta').$$

This holds thanks to a simple application of the MLRP.

For the latter, letting $h(y) = B(y)\mathbb{1}_{\{y_1 \geq \ldots \geq y_n\}}$, we need to show that for all $\bar{y}, \bar{y}' \in S^N$

$$h(\bar{y} \vee \bar{y}')h(\bar{y} \wedge \bar{y}') \geq h(\bar{y})h(\bar{y}'). \tag{4}$$

If at least one between y and y' is not weakly decreasing, $h(y)h(y') = 0$, and thus the affiliation inequality of Equation (4) holds because $h(y) \geq 0$ for any y. Therefore, assume both y and y' are weakly decreasing. We need to show that

$$B(y \vee y')B(y \wedge y') \geq B(y)B(y') = \frac{n!}{\prod_s q_s(y \vee y')!} \frac{n!}{\prod_s q_s(y \wedge y')!} \geq \frac{n!}{\prod_s q_s(y)!} \frac{n!}{\prod_s q_s(y')!} \implies$$

$$\prod_s q_s(y)! \prod_s q_s(y')! \geq \prod_s q_s(y \vee y')! \prod_s q_s(y \wedge y')! \implies$$

$$\prod_s q_s(y)!\, q_s(y')! \geq \prod_s q_s(y \vee y')!\, q_s(y \wedge y')!$$

It suffices to show that, for each $s \in S$,

$$q_s(y)!\, q_s(y')! \geq q_s(y \vee y')!\, q_s(y \wedge y')!$$

To this end, we first prove two properties of q_s.

Claim 1. *If y and y' are weakly decreasing and $s \in S$,*

$$\min\{q_s(y), q_s(y')\} \leq q_s(y \vee y') \leq \max\{q_s(y), q_s(y')\},$$

$$\min\{q_s(y), q_s(y')\} \leq q_s(y \wedge y') \leq \max\{q_s(y), q_s(y')\},$$

and

$$q_s(y) + q_s(y') = q_s(y \vee y') + q_s(y \wedge y').$$

Proof of Claim 1. Fix $s \in S$. Fix y and y', both weakly decreasing. Let $\underline{\iota} = \min\{i : y_i = s\}$, the position of the first appearance of s in y. Let $\bar{\iota} = \max\{i+1 : y_i = s\}$, the position following the last appearance of s in y. If s never appears in y, let $\underline{\iota} = \bar{\iota} = \min\{i : y_i < s\}$ if there exists some i such that $y_i < s$, and let $\underline{\iota} = \bar{\iota} = n$ if $y_i > s$ for all i. By definition, $\bar{\iota} \geq \underline{\iota}$. Note that, since y is weakly decreasing, $q_s(y) = \bar{\iota} - \underline{\iota}$. Define $\bar{\iota}'$ and $\underline{\iota}'$ for y' accordingly. It is straightforward to show that

$$q_s(y \vee y') = \max\{\bar{\iota}, \bar{\iota}'\} - \max\{\underline{\iota}, \underline{\iota}'\}, \qquad q_s(y \wedge y') = \min\{\bar{\iota}, \bar{\iota}'\} - \min\{\underline{\iota}, \underline{\iota}'\}.$$

Given these alternative definitions, it becomes easy to show that the first two inequalities in the Claim hold. In particular, they can be simplified to be

$$\min\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\} \leq \max\{\bar{\iota}, \bar{\iota}'\} - \max\{\underline{\iota}, \underline{\iota}'\} \leq \max\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\}$$

$$\min\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\} \leq \min\{\bar{\iota}, \bar{\iota}'\} - \min\{\underline{\iota}, \underline{\iota}'\} \leq \max\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\}.$$

To show that both statements are true, we need to consider four cases.

1. $\max\{\bar{\iota}, \bar{\iota}'\} = \bar{\iota}$ and $\max\{\underline{\iota}, \underline{\iota}'\} = \underline{\iota} \implies \min\{\bar{\iota}, \bar{\iota}'\} = \bar{\iota}'$ and $\min\{\underline{\iota}, \underline{\iota}'\} = \underline{\iota}'$. Our two inequalities become

$$\min\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\} \leq \bar{\iota} - \underline{\iota} \leq \max\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\}$$

$$\min\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\} \leq \bar{\iota}' - \underline{\iota}' \leq \max\{\bar{\iota} - \underline{\iota}, \bar{\iota}' - \underline{\iota}'\}$$

which are both trivially true;

2. $\max\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}'$ and $\max\{\underline{\iota},\underline{\iota}'\} = \underline{\iota}' \implies \min\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}$ and $\min\{\underline{\iota},\underline{\iota}'\} = \underline{\iota}$. The argument resembles point 1 and both inequalities are trivially true;

3. $\max\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}$ and $\max\{\underline{\iota},\underline{\iota}'\} = \underline{\iota}' \implies \min\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}'$ and $\min\{\underline{\iota},\underline{\iota}'\} = \underline{\iota}$. Our two inequalities become

$$\min\{\bar{\iota}-\underline{\iota},\bar{\iota}'-\underline{\iota}'\} \leq \bar{\iota}-\underline{\iota}' \leq \max\{\bar{\iota}-\underline{\iota},\bar{\iota}'-\underline{\iota}'\}$$

$$\min\{\bar{\iota}-\underline{\iota},\bar{\iota}'-\underline{\iota}'\} \leq \bar{\iota}'-\underline{\iota} \leq \max\{\bar{\iota}-\underline{\iota},\bar{\iota}'-\underline{\iota}'\}.$$

Notice that, under our assumptions, $\bar{\iota}-\underline{\iota} \geq \bar{\iota}'-\underline{\iota}'$ and we can further simplify the two inequalities to

$$\bar{\iota}'-\underline{\iota}' \leq \bar{\iota}-\underline{\iota}' \leq \bar{\iota}-\underline{\iota}$$

$$\bar{\iota}'-\underline{\iota}' \leq \bar{\iota}'-\underline{\iota} \leq \bar{\iota}-\underline{\iota}$$

which are both true due the fact that $\bar{\iota} \geq \bar{\iota}'$ and $\underline{\iota}' \geq \underline{\iota}$.

4. $\max\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}'$ and $\max\{\underline{\iota},\underline{\iota}'\} = \underline{\iota} \implies \min\{\bar{\iota},\bar{\iota}'\} = \bar{\iota}$ and $\min\{\underline{\iota},\underline{\iota}'\} = \underline{\iota}'$. The argument resembles part 3 and both inequalities are satisfied.

The last equation in the Claim follows trivially from our definitions. Indeed

$$\begin{aligned} q_s(y)+q_s(y') =& (\bar{\iota}-\underline{\iota})+(\bar{\iota}'-\underline{\iota}') = (\bar{\iota}+\bar{\iota}')-(\underline{\iota}+\underline{\iota}') = \\ & (\max\{\bar{\iota},\bar{\iota}'\}+\min\{\bar{\iota},\bar{\iota}'\})-(\max\{\underline{\iota},\underline{\iota}'\}+\min\{\underline{\iota},\underline{\iota}'\}) = \\ & (\max\{\bar{\iota},\bar{\iota}'\}-\max\{\underline{\iota},\underline{\iota}'\})+(\min\{\bar{\iota},\bar{\iota}'\}-\min\{\underline{\iota},\underline{\iota}'\}) = \\ & q_s(y\vee y')+q_s(y\wedge y') \end{aligned}$$

$\triangle$

We now return to our target, which is to show $q_s(y)!\ q_s(y')! \geq q_s(y\vee y')!\ q_s(y\wedge y')!$ Without loss of generality, assume $a := q_s(y) \leq q_s(y') =: b$ and let $a+b = Z$. From Claim 1, we know that $q_s(y\vee y')+q_s(y\wedge y') = Z$, $a \leq q_s(y\vee y') \leq b$ and $a \leq q_s(y\wedge y') \leq b$. Given these results, we cab write the following chain of equations

$$q_s(y\vee y')!\ q_s(y\wedge y')! \leq \max_{\substack{c,d: \\ a\leq c,d\leq b \\ c+d=Z}} c!d! = a!b! = q_s(y)!\ q_s(y')!$$

The first inequality follows from Claim 1. The second inequality follows from the fact that the maximum of the product of two factorials is achieved when they take opposite extreme values.

This completes the proof and allows us to conclude that the random variables $(\theta, y_1, \ldots, y_n)$ are affiliated. □

E.2 Proof of Remark 1

Proof of Remark 1, point 1. Denote the order statistics of the vector of signal realizations as $s_{(1:N)} \geq s_{(2:N)} \geq \ldots \geq s_{(N:N)}$, where $s_{(1:N)}$ denote the highest order statistic. Denote by $F_{(i:N)}$ the CDF of $s_{(i)}$, i.e. $F_{(i:N)}(s) = \mathbb{P}(s_{(i:N)} \leq s)$, for all $s \in S$. In any maximally selective equilibrium, the sender always reveals the K-highest signal realizations in $(s_1, \ldots, s_N)$. Let $g_{K,N}(s)$ be the probability that a signal drawn uniformly at random from the K-highest signal realizations revealed by the sender is equal to $s \in S$. That is,

$$g_{K,N}(s) \triangleq \frac{1}{K} \sum_{i=1}^{K} \mathbb{P}(s_{(i:N)} = s).$$

Thus, the CDF of the probability measure $g_{K,N}$ is

$$G_{K,N}(\hat{s}) \triangleq \sum_{s \leq \hat{s}} g_{K,N}(s) = \frac{1}{K} \sum_{i=1}^{K} \mathbb{P}(s_{(i:N)} \leq \hat{s}) = \frac{1}{K} \sum_{i=1}^{K} F_{(i:N)}(\hat{s}).$$

First, we want to show that for all N and $s \in S$, $G_{K,N+1}(s) \leq G_{K,N}(s)$ (with the rest of the argument holding by induction). This follows from the fact that, for all $i = 1, \ldots, N$ and s, $F_{(i,N)}(s) \geq F_{(i,N+1)}(s)$ (see, e.g., Shaked and Shanthikumar (2007, Corollary 1.C.38.), then use that the likelihood ratio order implies the FOSD order and that the latter is closed under mixtures). In particular, this holds for each $i \leq K$ and, therefore, $G_{K,N+1}(s) \leq G_{K,N}(s)$.

Second, we want to show that for all K and $s \in S$, $G_{K,N}(s) \leq G_{K+1,N}(s)$ (with the rest of the argument holding by induction). Using the definition above, we have that

$$\begin{aligned} G_{K+1,N}(s) =& \frac{1}{K+1} \sum_{i=1}^{K} F_{(i:N)}(s) + \frac{1}{K+1} F_{(K+1:N)}(s) \\ =& \frac{K}{K+1} \left(\frac{1}{K} \sum_{i=1}^{K} F_{(i:N)}(s) \right) + \frac{1}{K+1} F_{(K+1:N)}(s). \end{aligned}$$

Notice that $F_{(i:N)}(s) \leq F_{(j:N)}(s)$ for each $i < j$ and s. This follows directly from the fact that, by definition of the order statistics, $s_{(i:N)} \geq s_{(j:N)}$. Thus, $F_{(K+1:N)}(s) \geq \frac{1}{K}\sum_{i=1}^{K} F_{(i:N)}(s) = G_{K,N}(s)$. Thus, $G_{K+1,N}(s)$ is a convex combination of $G_{K,N}(s)$ and a greater term, and therefore $G_{K,N}(s) \leq G_{K+1,N}(s)$. □

Proof of Remark 1, point 2. We want to show that in any maximally selective equilibrium, fixing any K and any on-the-equilibrium-path message m, the conditional expectation of θ given m is decreasing in N. Any such message reveals the K-highest signal realizations in $(s_1, \ldots, s_N)$. Let $\tilde{s} = (s_1, \ldots, s_K)$ denote the vector of the K-highest signal realizations revealed by m. By Milgrom (1981), a sufficient condition for such expectation to be decreasing in N is hence that $\frac{\mathbb{P}_N(\tilde{s}|\theta)}{\mathbb{P}_{N+1}(\tilde{s}|\theta)}$ is increasing in θ, where $\mathbb{P}_n(\tilde{s}|\theta)$ denotes the probability that the K-highest signal realizations out of n draws are those in $\tilde{s}$ conditional on θ. We have that

$$\mathbb{P}_N(\tilde{s} \mid \theta) = \begin{cases} \frac{N!}{(N-L)!\prod_{j=1}^{r-1} m_j!} \prod_{j=1}^{r-1} f(y_j \mid \theta)^{m_j} f(\min S \mid \theta)^{N-L} & \text{if } y_r = \min S \\ \sum_{n_r=m_j}^{N-L} \frac{N!}{(N-L-n_r)!\prod_{j=1}^{r-1} m_j! n_r!} \prod_{j=1}^{r-1} f(y_j \mid \theta)^{m_j} f(y_r \mid \theta)^{n_r} F(y_r^- \mid \theta)^{N-L-n_r} & \text{if } y_r > \min S \end{cases}$$

where:

- $\tilde{s} = (s_1, \ldots, s_K)$ is the vector of the K-highest signal realizations revealed by m;
- $y(\tilde{s}) = (y_1, \ldots, y_r)$ is the vector of the *distinct* values in $\tilde{s}$, with $y_r \triangleq \min\{\tilde{s}_i \in \tilde{s}\}$. Namely, $y(\tilde{s})$ differs from $\tilde{s}$ in that identical realizations are collapsed into single elements, and, moreover, the lowest of these elements, denoted by y_r, is arranged at the end;
- $y_r^- \triangleq \max\{s \in S \mid s < y_r\}$ denotes the largest signal in S that is strictly less than y_r;
- each element $y_j \in y(\tilde{s})$ has multiplicity m_j in $\tilde{s}$ and n_j in the *entire sample*;
- for each y_j with $j < r$, $n_j = m_j$, since realizations strictly higher than y_r cannot appear outside the top K. Thus, $L \triangleq \sum_{j=1}^{r-1} m_j$ denotes the total multiplicity of such realizations in the entire sample;
- as for y_r, if $y_r = \min S$, all realizations outside the top K must be equal to $\min S$ as well, so $n_r = N - L$, while if $y_r > \min S$, n_r varies from m_r to $N - L$.

First, we consider the case $y_r = \min S$. Then,

$$\frac{\mathbb{P}_N(\tilde{s}|\theta)}{\mathbb{P}_{N+1}(\tilde{s}|\theta)} = \frac{\frac{N!}{(N-L)!\prod_{j=1}^{r-1} m_j!} \prod_{j=1}^{r-1} f(y_j \mid \theta)^{m_j} f(\min S \mid \theta)^{N-L}}{\frac{(N+1)!}{(N+1-L)!\prod_{j=1}^{r-1} m_j!} \prod_{j=1}^{r-1} f(y_j \mid \theta)^{m_j} f(\min S \mid \theta)^{N+1-L}}$$

$$= \frac{N+1-L}{N+1} \frac{1}{f(\min S \mid \theta)}.$$

Thus, the result follows since, by the MLRP, $f(\min S \mid \theta)$ is decreasing in θ (the MLRP implies FOSD, i.e., $F(s \mid \theta') \leq F(s \mid \theta)$ for $\theta' > \theta$ and $s \in S$, which in particular holds at $s = \min S$).

Next, we consider the case $y_r > \min S$. Then,

$$\frac{\mathbb{P}_N(\tilde{s}|\theta)}{\mathbb{P}_{N+1}(\tilde{s}|\theta)} = \frac{\sum_{n_r=m_r}^{N-L} \frac{N!}{(N-L-n_r)!\,n_r!} f(y_r \mid \theta)^{n_r} F(y_r^- \mid \theta)^{N-L-n_r}}{\sum_{n_r=m_r}^{N+1-L} \frac{(N+1)!}{(N+1-L-n_r)!\,n_r!} f(y_r \mid \theta)^{n_r} F(y_r^- \mid \theta)^{N+1-L-n_r}}$$

$$= \frac{N+1-L}{N+1} \frac{\sum_{n_r=m_r}^{N-L} \binom{N-L}{n_r} f(y_r \mid \theta)^{n_r} F(y_r^- \mid \theta)^{N-L-n_r}}{\sum_{n_r=m_r}^{N+1-L} \binom{N+1-L}{n_r} f(y_r \mid \theta)^{n_r} F(y_r^- \mid \theta)^{N+1-L-n_r}}$$

$$= \frac{N+1-L}{N+1} \frac{(f(y_r \mid \theta) + F(y_r^- \mid \theta))^{N-L} \sum_{n_r=m_r}^{N-L} \binom{N-L}{n_r} p_\theta^{n_r} (1-p_\theta)^{N-L-n_r}}{(f(y_r \mid \theta) + F(y_r^- \mid \theta))^{N+1-L} \sum_{n_r=m_r}^{N+1-L} \binom{N+1-L}{n_r} p_\theta^{n_r} (1-p_\theta)^{N+1-L-n_r}}$$

$$= \frac{N+1-L}{N+1} \frac{1}{F(y_r \mid \theta)} \frac{\mathbb{P}_{\mathrm{Bin}(N-L,p_\theta)}(n_r \geq m_r)}{\mathbb{P}_{\mathrm{Bin}(N+1-L,p_\theta)}(n_r \geq m_r)},$$

where the second equation obtains using that $\frac{N!}{(N-L-n_r)!\,n_r!} = \binom{N-L}{n_r} \frac{N!}{(N-L)!}$, in the third equation $p_\theta \triangleq \frac{f(y_r|\theta)}{f(y_r|\theta)+F(y_r^-|\theta)} = \frac{f(y_r|\theta)}{F(y_r|\theta)}$ and, in the forth equation, $\mathbb{P}_{\mathrm{Bin}(N-L,p_\theta)}(n_r \geq m_r)$ is the right tail (i.e., the probability of at least m_r successes) of a binomial distribution with success probability p_θ and $N-L$ trials. By the MLRP, $F(y_r \mid \theta)$ is decreasing in θ while p_θ is increasing in θ (since the MLRP implies reverse hazard rate dominance). Hence, a sufficient condition for $\frac{\mathbb{P}_N(\tilde{s}|\theta)}{\mathbb{P}_{N+1}(\tilde{s}|\theta)}$ to be increasing in θ is that the ratio of right tails in the last equation is increasing in p_θ.

To see this, note that by a known Beta-Binomial identity, $\mathbb{P}_{\mathrm{Bin}(N,p)}(n \geq x) = \frac{\mathrm{B}(p;x,1+N-x)}{\mathrm{B}(x,1+N-x)}$, where $\mathrm{B}(.,.,.)$ and $\mathrm{B}(.,.)$ are the incomplete and complete Beta functions, respectively. Thus, ignoring multiplicative constants,

$$R(p) \triangleq \frac{\mathbb{P}_{\mathrm{Bin}(N,p)}(n \geq x)}{\mathbb{P}_{\mathrm{Bin}(N+1,p_\theta)}(n \geq x)} \propto \frac{\int_0^p t^{x-1}(1-t)^{N-x}\,dt}{\int_0^p t^{x-1}(1-t)^{N+1-x}\,dt}$$

Defining $f(p) \triangleq \int_0^p t^{x-1}(1-t)^{N-x}\,dt$ and $h(p) \triangleq \int_0^p t^{x-1}(1-t)^{N+1-x}\,dt$, so that $R(p) \propto$

$\frac{f(p)}{h(p)}$, and using that $f'(p) = p^{x-1}(1-p)^{N-x}$ and $h'(p) = p^{x-1}(1-p)^{N+1-x}$,

$$R'(p) \propto \frac{p^{x-1}(1-p)^{N-x}\left[h(p) - (1-p)f(p)\right]}{h(p)^2}.$$

The sign of $R'(p)$ is hence determined by the sign of

$$\begin{aligned} h(p) - (1-p)f(p) &= \int_0^p t^{x-1}\left[(1-t)^{N+1-x} - (1-p)(1-t)^{N-x}\right]dt \\ &= \int_0^p t^{x-1}(1-t)^{N-x}(p-t)dt > 0. \end{aligned}$$

□

E.3 Informativeness

In this Section, we establish the link between the ex ante receiver's expected payoff and the expected variance of the state θ given the disclosed message.

Remark 4. *Consider any PBE of our game. Fix the message $m \in \mathcal{M}$, the receiver's strategy $\xi : \mathcal{M} \to A$, the sender's strategy $\sigma : \Theta \times S^N \to \Delta(\mathcal{M})$ and the receiver's posterior belief $\mu(\cdot|m) \in \Delta(\Theta)$. The correlation between the state θ and the receiver's action induced by $\xi(\cdot)$ is a monotonic transformation of the ex ante expected payoff of the receiver.*

Proof. Consider any PBE of our game. Fix the message $m \in \mathcal{M}$, the receiver's strategy $\xi : \mathcal{M} \to A$, the sender's strategy $\sigma : \Theta \times S^N \to \Delta(\mathcal{M})$ and the receiver's posterior belief $\mu(\cdot|m) \in \Delta(\Theta)$. We have that

$$\mathbb{E}_\theta\left[u\left(\theta, \xi(m)\right)\right] = -\sum_{\theta' \in \Theta} \mu(\theta'|m)\left(\mathbb{E}\left[\theta|m\right] - \theta'\right)^2$$

Given this, we can derive the ex-ante expected payoff of the receiver as:

$$\begin{aligned} \mathbb{E}_{\theta,m}\left[u_R\left(\theta, \xi(m)\right)\right] &= -\sum_{m \in \mathcal{M}} Prob(m) \sum_{\theta' \in \Theta} \mu(\theta'|m)\left(\mathbb{E}\left[\theta|m\right] - \theta'\right)^2 \\ &\quad - \sum_{m \in \mathcal{M}} \sum_{\theta' \in \Theta} Prob(m, \theta')\left(\mathbb{E}\left[\theta|m\right] - \theta'\right)^2. \end{aligned}$$

At this point notice that

$$Prob(m, \theta) = Prob(m) \cdot \mu(\theta|m).$$

Rearranging the expression we get

$$\mathbb{E}_{\theta,m}\left[u_R\left(\theta,\xi(m)\right)\right] = -\sum_{m\in\mathcal{M}} Prob(m) \sum_{\theta'\in\Theta} \mu(\theta|m)\left(\mathbb{E}\left[\theta|m\right]-\theta'\right)^2$$

which implies

$$\mathbb{E}_{\theta,m}\left[u_R\left(\theta,\xi(m)\right)\right] = -\sum_{m\in\mathcal{M}} Prob(m) Var\left[\theta|m\right] = -\mathbb{E}_m\left[Var\left[\theta|m\right]\right]$$

This argument directly links the ex-ante expected payoff of the receiver with the variance of θ given the disclosed message. We can now show that $-\mathbb{E}_m\left[Var[\theta|m]\right]$ is monotonic transformation of $Corr(\theta,a)$, where a is the random variable generated by $\xi(\cdot)$ and

$$Corr(\theta,a) = \frac{\mathbb{E}\left[\theta\cdot a\right]-\mathbb{E}[\theta]\mathbb{E}[a]}{\sqrt{Var[\theta]Var[a]}} = \frac{\mathbb{E}_m\left[\mathbb{E}_\theta\left[\theta\cdot\xi(m)|m\right]\right]-\mathbb{E}[\theta]\mathbb{E}_m[\xi(m)]}{\sqrt{Var[\theta]Var_m[\xi(m)]}}$$

Notice that:

- $\mathbb{E}_m[\xi(m)] = \mathbb{E}[\theta]$;
- $\mathbb{E}_m\left[\mathbb{E}_\theta\left[\theta\cdot\xi(m)|m\right]\right] = \mathbb{E}_m\left[\xi(m)^2\right]$ since $\mathbb{E}_\theta[\theta|m] = \xi(m)$;
- $Var_m[\xi(m)] = \mathbb{E}_m\left[\xi(m)^2\right]-\mathbb{E}[\theta]^2$.

This implies that

$$Corr(\theta,a) = \frac{\sqrt{Var_m[\xi(m)]}}{\sqrt{Var[\theta]}} = \frac{\sqrt{\mathbb{E}[a^2]-\mathbb{E}[\theta]^2}}{\sqrt{Var[\theta]}}.$$

Given that $-\mathbb{E}_m[Var[\theta|m]] = \mathbb{E}_m\left[\mathbb{E}\left[\theta|m\right]^2\right] - \mathbb{E}_m\left[\mathbb{E}\left[\theta^2|m\right]\right] = \mathbb{E}[a^2]-\mathbb{E}[\theta^2]$, we can derive the following relation:

$$-\mathbb{E}_m[Var[\theta|m]] = Var[\theta]\cdot Corr(\theta,a)^2 - Var[\theta].$$

□

This argument allows us to conclude that both $-\mathbb{E}_m[Var[\theta|m]]$ and $Corr(\theta,a)$ can be used to study the level of information transmitted in equilibrium. Indeed, both measures provide the same comparative statics with respect to the main parameters of our model.

E.4 Examples of the Non-Monotonicity of $\mathcal{I}(K,N)$

As stated in Proposition 2, $\mathcal{I}(K,N)$ can be non-monotonic in N. In Section 2.2 we provided an intuition for this result by introducing two effects, the imitation and the revelation effect. The following example illustrates more in detail both effects and how their interaction is the key determinant of the shape of $\mathcal{I}(K,N)$.

Throughout, we suppose that $\Theta = \{0,1\}$, with each state equally likely, and we fix $K = 1$. Also, we assume $S = \{B, A\}$, with $f(A|1) = \gamma > \eta = f(A|0)$. Besides, we let $\mu(m) = \mu(1|m)$ denote the receiver's belief that $\theta = 1$, which also coincides with the optimal guess given such belief. Finally, among all possible maximally selective strategies of the sender, we focus on the one that prescribes to always disclose a signal, namely, send $m = A$ if an A is available and $m = B$ otherwise.

In equilibrium, the receiver's beliefs, hence actions, upon each disclosed signal are pinned down by Bayes' rule, i.e.,

$$a(B) = \mu(B) = \frac{(1-\gamma)^N}{(1-\gamma)^N + (1-\eta)^N}$$

$$a(A) = \mu(A) = \frac{1-(1-\gamma)^N}{2-(1-\gamma)^N-(1-\eta)^N}.$$

Thus, the receiver's expected utility in equilibrium is

$$\begin{aligned}
\mathbb{E}\left[u\left(\theta,a\right)\right] = &-\frac{1}{2}\left(\left(1-(1-\gamma)^N\right)(1-a(A))^2 + (1-\gamma)^N\left(1-a(B)\right)^2\right) \\
&-\frac{1}{2}\left(\left(1-(1-\eta)^N\right)(a(A))^2 + (1-\eta)^N\left(a(B)\right)^2\right) \\
= &-\frac{(1-\gamma)^N\left(1-(1-\gamma)^N\right) + (1-\eta)^N\left(1-(1-\eta)^N\right)}{2\left((1-\gamma)^N+(1-\eta)^N\right)\left(2-(1-\gamma)^N-(1-\eta)^N\right)}.
\end{aligned}$$

According to our definition of equilibrium informativeness, $(\mathcal{I}(K,N))^2 = 1 + \mathbb{E}\left[u\left(\theta,a\right)\right]/Var[\theta]$ (see Section E.3). Hence, defining $\mathcal{I}^2(K,N) \triangleq (\mathcal{I}(K,N))^2$ and using that $Var[\theta] = 1/4$,

$$\mathcal{I}^2(K,N) = \frac{\left((1-\gamma)^N-(1-\eta)^N\right)^2}{\left(2-(1-\gamma)^N-(1-\eta)^N\right)\left((1-\gamma)^N+(1-\eta)^N\right)}.$$

To build intuition, we first relax the assumption that f has full support and study two extreme cases. Consider first the extreme case of *perfect good news*, i.e., the case in which $\eta = 0$ and

$\gamma < 1$. Under this parametrization, $m = A$ perfectly reveals that $\theta = 1$. Then,

$$\mathcal{I}^2(K, N) = \frac{1 - (1-\gamma)^N}{1 + (1-\gamma)^N},$$

which is strictly increasing in N. Hence, an increase in N always leads to more information being transmitted in equilibrium. The intuition behind this result is straightforward. Only $\theta = 1$ can draw a signal equal to A and, when this happens, the value of the state is fully revealed. The larger the number of available signals, the more likely it is that a high-type sender can send $m = A$. In addition, as a consequence of the previous fact, when N grows, observing $m = B$ makes the receiver more confident that $\theta = 0$. These two channels together are responsible for the fact that more available signals lead to more information being transmitted in equilibrium: We refer to this as the *revelation effect*.

Now consider the other extreme case of *perfect bad news*, i.e., the case in which $\eta > 0$ and $\gamma = 1$. Under this parametrization, $m = B$ perfectly reveals that $\theta = 0$. Then,

$$\mathcal{I}^2(K, N) = \frac{(1-\eta)^N}{2 - (1-\eta)^N},$$

which is strictly decreasing in N. That is, the amount of information transmitted in equilibrium decreases with the number of available signals. Again, the intuition is simple. As N grows, the low-type sender is increasingly more likely to draw at least one A. Thus, sending the fully-revealing message $m = B$ becomes less likely and, at the same time, $m = A$ becomes weaker evidence of $\theta = 1$. This leads to less information being transmitted in equilibrium: We refer to this as the *imitation effect*.

Finally, consider the case in which f has full support, i.e., in which both $\eta > 0$ and $\gamma < 1$. Intuitively, in this case, both the revelation and imitation effects play a role in how the information transmitted in equilibrium changes with the number of available signals. Which effect prevails determines the direction of the change in $\mathcal{I}(K, N)$ after an increase in N. We now show that when the signal has symmetric precision, i.e., $\gamma > 1/2$ and $\eta = 1 - \gamma$, informativeness always increases when moving from $N = 1$ to $N = 2$. Indeed,

$$\mathcal{I}^2(1, 2) - \mathcal{I}^2(1, 1) = \frac{(2\gamma - 1)^2}{1 - 4(1-\gamma)^2\gamma^2} - (2\gamma - 1)^2 > 0,$$

which is positive because the denominator of the first term is positive and less than one. Since $\mathcal{I}(K, N)$ always converges to zero in the limit as N goes to infinity, it is nonmonotonic in N.

Hence, the example discussed above allows us to illustrate that the effect of a change in N on informativeness is the result of positive and negative forces. The exact parametrization of the model pins down the overall effect. However, as N becomes extremely large, the negative force always prevails, making communication fully ineffective.

E.5 A Model with Random N

This section considers an extension of the theoretical predictions of Section 2.2 when N is not common knowledge. For simplicity we assume that the state and signals are binary, i.e., $\Theta = \{0, 1\}$ and $S = \{A, B\}$, and at most one signal can be disclosed, i.e., $K = 1$. This setting is similar to that analyzed in Section E.4. For ease of exposition, we also assume that the state is uniformly distributed and signals have symmetric precision, i.e., $f(A|1) = f(B|0) = \gamma \in (1/2, 1)$. All insights generalize to an arbitrary prior and likelihood function.

In contrast to the main model, we assume that N is a random variable. While its distribution is common knowledge, the realized value of N is observed only by the sender. Specifically, for a given $n \in \mathbb{N}$, we assume that $N = n$ with probability $\delta \in (0, 1)$ and $N = 1$ with complementary probability $1 - \delta$. This specification is not intended to be fully general, but we believe it captures the core features of a setting in which N is no longer common knowledge. We show that the comparative-static results with respect to N in Section 2.2 carry through in this variation with respect to δ and n.[7]

We begin by computing the unconditional probability that each message is sent, under the assumption that the sender plays the maximally selective strategy that consists of disclosing the highest signal available. Note that:

$$\begin{aligned}
\Pr(m = A) &= \frac{1}{2}\left[(1-\delta)\gamma + \delta(1-(1-\gamma)^n)\right] + \frac{1}{2}\left[(1-\delta)(1-\gamma) + \delta(1-\gamma^n)\right] \\
&= \frac{1}{2}\left[(1-\delta)(\gamma + 1 - \gamma) + \delta\left(2 - (1-\gamma)^n - \gamma^n\right)\right] \\
&= \frac{1}{2}\left[1 + \delta\left(1 - (1-\gamma)^n - \gamma^n\right)\right].
\end{aligned}$$

[7] As in our baseline model, we continue to assume that $\mathbb{P}(N \geq K) = 1$—that is, the receiver knows that N is at least as large as K. This is an important assumption. When it does not hold, the sender may have fewer than K signals to disclose, which introduces incentives to hide behind a veil of ignorance—a feature reminiscent of Dye (1985). In such cases, it is known that a maximally selective equilibrium may fail to exist.

Similarly,

$$\begin{aligned}\Pr(m=B) &= \frac{1}{2}\left[(1-\delta)(1-\gamma)+\delta(1-\gamma)^n\right]+\frac{1}{2}\left[(1-\delta)\gamma+\delta\gamma^n\right] \\ &= \frac{1}{2}\left[(1-\delta)(1-\gamma+\gamma)+\delta\left((1-\gamma)^n+\gamma^n\right)\right] \\ &= \frac{1}{2}\left[1-\delta\left(1-(1-\gamma)^n-\gamma^n\right)\right].\end{aligned}$$

Finally, since the sender always discloses either A or B, we have $\Pr(m=o)=0$.

We now discuss the predictions of Remark 1. First, we note that the distribution of signals disclosed by the sender increases in a first-order stochastic sense with both δ and n. Since $K=1$, this simply means that the probability of receiving message $m=A$ increases in both δ and n, while the probability of receiving message $m=B$ decreases in both δ and n. In other words, as δ or n increases, the message received appears more favorable, as predicted by Remark 1.1.

Next, we compute the receiver's optimal response to each message. We have:

$$\begin{aligned}a(A)=\Pr(\theta=1\mid m=A) &= \frac{\Pr(\theta=1)\Pr(m=A\mid\theta=1)}{\Pr(m=A)} \\ &= \frac{(1-\delta)\gamma+\delta\left(1-(1-\gamma)^n\right)}{1+\delta\left(1-(1-\gamma)^n-\gamma^n\right)}.\end{aligned}$$

Similarly,

$$\begin{aligned}a(B)=\Pr(\theta=1\mid m=B) &= \frac{\Pr(\theta=1)\Pr(m=B\mid\theta=1)}{\Pr(m=B)} \\ &= \frac{(1-\delta)(1-\gamma)+\delta(1-\gamma)^n}{1-\delta\left(1-(1-\gamma)^n-\gamma^n\right)}.\end{aligned}$$

Since message $m=o$ is off-path, we assign to it the most pessimistic belief, which induces the lowest possible action, $a(o)=0$. It can be verified that $a(A)>a(B)>a(o)$. Therefore, it is optimal for the sender to employ a maximally selective strategy, and a maximally selective equilibrium exists. This extends the result in Proposition 1 to the current setting.

Next, a direct computation of the derivative $\frac{d\,a(A)}{d\delta}$ shows that its sign is determined by the numerator:

$$\gamma(1-\gamma)^n-2\gamma+\gamma^{n+1}-(1-\gamma)^n+1=1-2\gamma-(1-\gamma)^{n+1},$$

which is negative for all $\gamma>\frac{1}{2}$. Therefore, $a(A)$ is decreasing in δ. Similarly, the derivative

$\frac{d\,a(B)}{d\delta}$ has numerator

$$\gamma(1-\gamma)^n - \gamma^n + \gamma^{n+1} = \gamma(1-\gamma)\left[(1-\gamma)^{n-1} - \gamma^{n-1}\right],$$

which is also negative for all $\gamma > \frac{1}{2}$. Thus, $a(B)$ is decreasing in δ. Together, these results imply that, for any n, as δ increases, the receiver becomes more skeptical of the messages received—responding with lower actions—regardless of whether A or B is observed. This is in line with the comparative statics of Remark 1.2.

Having derived the probabilities of each message and the receiver's best response, we can now compute the receiver's expected utility, which, as shown in Online Appenix E.3, is a monotone transformation of informativeness. The derivation is somewhat algebraically intensive, so we omit the intermediate steps and report the final expression directly. Letting $X \triangleq \gamma(1-\gamma)$, $W \triangleq \gamma^{n-1} + (1-\gamma)^{n-1}$, and $Z \triangleq \gamma^n + (1-\gamma)^n$, the receiver's expected utility in the maximally selective equilibrium simplifies to

$$\mathbb{E}[u_R] = -\frac{1}{2}\left(\frac{2X(1-\delta)\left[1-\delta(1-W)\right] + 2\delta^2 X^n + \delta(1-Z)\left[1-\delta(1-Z)\right]}{1-\delta^2(1-Z)^2}\right)$$

Figure E.14 plots the receiver's expected utility as a function of n for various values of δ. We highlight that the pattern mirrors the one observed in our baseline model (see Proposition 2 and Figure 1). Notably, as $n \to \infty$, the receiver's utility does not converge to the value it would attain under complete ignorance, namely -0.25 (the negative of the prior variance). This is because, in the current model, the sender's strategy remains partially informative, as $N = 1$ with probability $1-\delta$.

F Design

F.1 Graphical Interface

The figures in this section show the software interface of our main experiment, namely the one described in the main body of the paper. Figures F.15 and F.16 show the sender's screen at the time of selecting her message. Figure F.17 shows the receiver's screen at the time of making her guess. Figure F.18 shows the feedback screen.

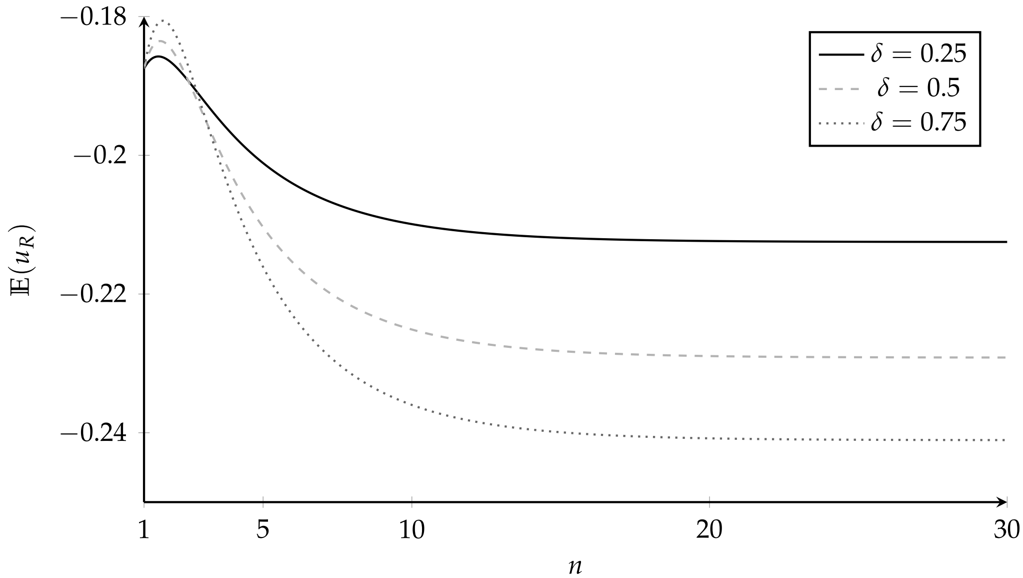


Figure E.14: Receiver's expected utility with random N. We set $\gamma = 0.75$ and let $\delta \in \{\frac{1}{4}, \frac{2}{4}, \frac{3}{4}\}$.

Figure F.15: Sender's Interface Before the Message Choice

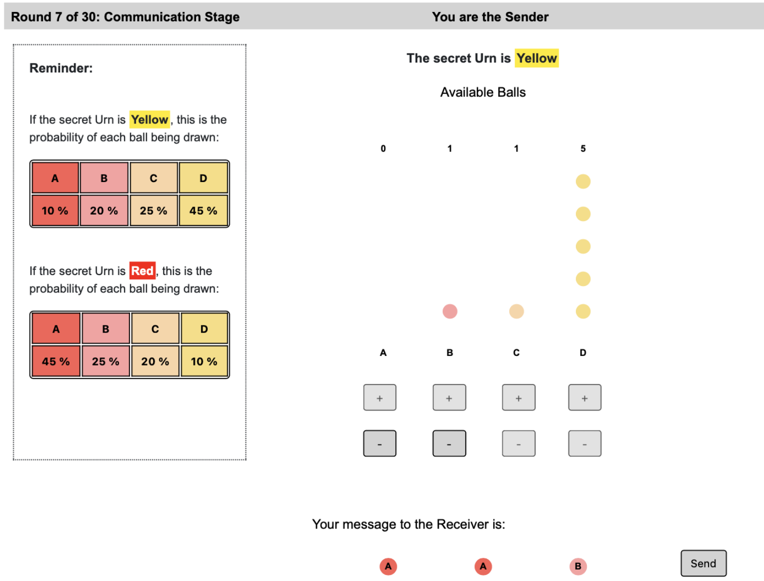


Figure F.16: Sender's Interface After the Message Choice

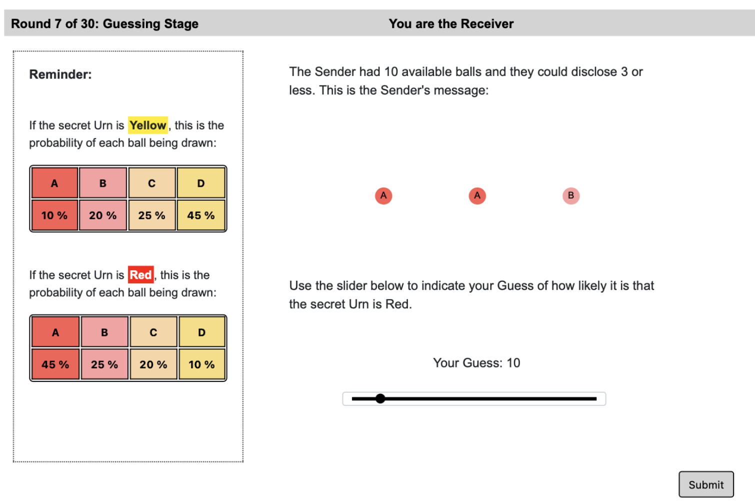


Figure F.17: Receiver's Interface

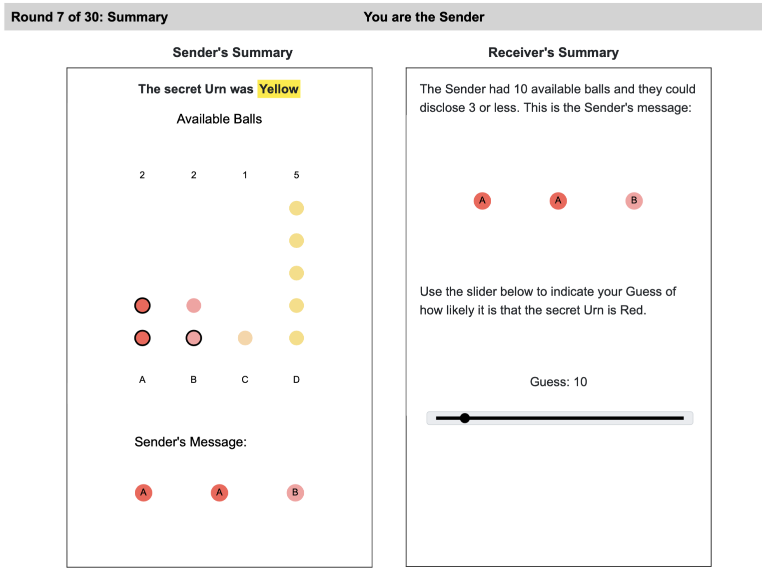


Figure F.18: Feedback Interface

F.2 Sample Instructions

We reproduce instructions for one of our treatments, (K_3, N_{10}). These instructions were read out aloud at the beginning of each session. Additionally, a copy of the instructions was handed out to the subject and it was available to them at any point during the experiment.

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation with cash vouchers (privately) at the end of the session. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

Please turn off phones and tablets now. Please close any program you may have open on the computer. The entire session will take place through computer terminals, and all interaction among you will take place through computers. Please do not talk or in any way try to communicate with other participants during the session.

We will start with a brief instruction period. During the instruction period you will be given a description of the main features of the session and will be shown how to use the computer interface. If you have any questions during this period, raise your hand and your question will be answered privately.

Instructions

You will play for 30 matches in either of two roles: **Sender** or **Receiver**. At the end of each round, you will be randomly paired with a new player.

There are two urns: one **Red** and one **Yellow**. Each urn contains four types of balls, labeled A, B, C, and D.

The Round

At the beginning of each round, the Computer randomly selects one of the two urns (we will refer to the selected urn as secret Urn). The secret Urn has a 50% chance of being Red and a 50% chance of being Yellow.

The Computer randomly draws 10 balls from the secret Urn. Depending on the color of the secret Urn, each ball has a chance of being drawn that is reported in the table below:

Urn	A	B	C	D
Red Urn	45%	25%	20%	10%
Yellow Urn	10%	20%	25%	45%

The 10 balls are drawn independently, meaning that the chance of drawing a ball is not affected by previous draws.

1. Communication Stage—Sender is Active

The sender observes the color of the secret Urn and sees the 10 balls that were drawn from it by the Computer.

The Sender can disclose to the Receiver up to 3 of these 10 balls. We call this the Sender's **Message**.

2. Guessing Stage—Receiver is Active

The Receiver observes the Sender's Message but does not observe the color of the secret Urn.

The Receiver must guess how likely it is that the secret Urn is Red. Specifically, the Receiver chooses a number from 0 to 100. We call this the Receiver's **Guess**.

For example, a Guess of 20 indicates that the Receiver believes there is a 20% chance that the secret Urn is Red. A Guess of 80, instead, indicates that the Receiver believes there is an 80% chance that the secret Urn is Red. More generally, a higher Guess indicates a greater chance that the secret Urn is Red.

3. Feedback

At the end of each round, both Sender and Receiver will see screens that summarize information from the Round. You will learn the color of the secret Urn; the balls that were available to the Sender; the Message sent by the Sender; the Receiver's Guess; and your payoff. You will also see a history of what happened in previous rounds.

How Payoffs Are Determined

In each round, you earn points that will be converted into cash at the end of the experiment.

Sender

The number of points the Sender earns in a round depends only on the Receiver’s Guess and not on the color of the secret Urn. Specifically, the number of earned points is equal to the Receiver’s Guess. Therefore, the higher the Receiver’s Guess, the greater the number of points earned by the Sender.

Receiver

The number of points the Receiver earns depends on the Guess, on the color of the secret Urn, and on chance. Specifically, the number of earned points is determined as follows:

> The Computer randomly generates two numbers between 0 and 100, where each integer number is equally likely. Let’s call them the *Computer’s Random Numbers*.
>
> The Receiver earns 100 points if one of the following two things happens:
>
> - The secret Urn is Red and the Receiver’s Guess is greater than or equal to the smallest of the two *Computer’s Random Numbers*.
> - The secret Urn is Yellow and the Receiver’s Guess is smaller than or equal to the largest of the two *Computer’s Random Numbers*.
>
> The Receiver earns 0 points otherwise.

This compensation rule was designed so that the Receiver has the greatest chance of earning 100 points when they choose a Guess that equals their true belief that the secret Urn is Red.

Final Payments

At the end of the experiment, the total number of points you earned will be converted to dollars at the rate of:

- $0.012 per point ($1.20 per 100 points) if you are the Sender.
- $0.009 per point ($0.90 per 100 points) if you are the Receiver.

In addition, you will receive a flat participation fee of $10.

Practice Rounds:

The experiment will begin with 2 practice rounds, to make you familiar with the interface and the tasks of both Sender and Receiver. All the choices you make in the Practice Rounds are unpaid and do not affect in any way the rest of the experiment.

Summary

Before we start, let me remind you that:

- You will play 30 Rounds in the same role: Sender or Receiver. You will be assigned your role at the end of the Practice Rounds.
- The secret Urn has an equal chance of being Red or Yellow.
- The Computer randomly draws 10 balls from the secret Urn.
- The Sender can disclose to the Receiver up to 3 of these 10 balls.
- The Receiver has to guess how likely it is that the secret Urn is Red.
- The Receiver has the greatest chance of earning points when they choose a Guess that equals their true belief that the secret Urn is Red.
- The higher the Receiver's Guess, the greater the number of points earned by the Sender.
- At the end of each Round, you are randomly paired with a new participant.