Online Appendix for

Disentangling Probabilistic and Strategic Errors

Sen Geng Menglong Guan

CONTENTS:

A. Experiment Parameters and Equilibrium Belief Specification

A.1. *Experiment Parameters*

To ensure meaningful comparison between endogenous and exogenous settings, we impose two key conditions. The first condition concerns payoff and information comparability. Specifically, the games share identical payoff parameters (u_0, u_1, u_2) and information diagnosticity (q_H, q_L). Since the initial belief about the state ω_1 differs structurally, i.e., fixed at p^* (exogenously given) in exogenous settings and strategically chosen by the persuader ($p \in [p_L, p_H]$) in endogenous settings, we enforce comparability by requiring $p^* \in [p_L, p_H]$, ensuring beliefs span the same range.[58] The second condition concerns equilibrium uniqueness and consistency. To cleanly define and decompose errors, we require a unique equilibrium outcome in all settings and identical receiver behavior ($ARRR$ in $I_1 - I_4$) across equilibria. Propositions 1 and 2 show these requirements are met under:

$$q_H > \frac{u_2-u_0}{u_1+u_2-2u_0} > max\{q_L, \frac{p_H q_H - p_L q_L}{p_H - p_L + q_H - q_L}\} \text{ and } \mu_1(I_1) < p_L < \mu_1(I_2) < p^* \leq p_H < \mu_1(I_3).$$

Recall that the threshold belief is given by $\hat{u} = \frac{v_0-v_2}{v_1-v_2}$ and $(\mu_1(I_1), \mu_1(I_2), \mu_1(I_3))$ are fully determined by $(\hat{u}, q_H, q_L)$. The sufficient condition presented earlier defines a parameter space for $(u_0, u_1, u_2, v_0, v_1, v_2, p_H, p_L, q_H, q_L, p^*)$ that meets our experimental requirements.[59] We chose the following values from the admissible parameter space: (1) payoffs: $u_0 = v_0 = 10$, $u_1 = 12$, $u_2 = 21$, $v_1 = 15$, $v_2 = 6$; (2) belief parameters: $p_H = 0.6$, $p_L = 0.2$; (3) information diagnosticity: $(q_H, q_L) = (1, 0.75)$ in Revealing settings (with perfect state revelation being possible), and $(q_H, q_L) = (0.95, 0.7)$ in Not Revealing settings (imperfect revelation).

Finally, we set the exogenous prior belief to $p^* = 0.4$, matching the empirical average of the persuader's chosen p in endogenous settings. This choice satisfies both the sufficient equilibrium conditions outlined above and our additional requirement that p^* reflects the actual mean p chosen by persuaders. Specifically, the average chosen p values were 0.39 in the Endogenous-Revealing setting (95 percent confidence interval $[0.36, 0.41]$) and 0.42 in the Endogenous-Not Revealing setting (95 percent confidence interval $[0.40, 0.44]$). Thus, $p^* =$

[58] An alternative approach is to set p^* equal to the endogenous equilibrium choice of p. However, to account for potential off-equilibrium behavior, we adopt the weaker condition $p^* \in [p_L, p_H]$.

[59] Ideally, p^* should equal the mean value of the persuader's actual p choices in endogenous settings to maximize comparability. This requirement implies that the empirical mean of p falls between $\mu_1(I_2)$ and p_H. To increase the likelihood of satisfying this, we further constrained the parameter space by requiring $\frac{\mu_1(I_2)-p_L}{p_H-\mu_1(I_2)} \leq \frac{1}{5}$. Ex post, this additional requirement was satisfied in our implementation.

0.4 provides comparable belief levels between exogenous settings and endogenous settings.[60]

A.2. *Equilibrium Belief Specification under Experiment parameters*

In Revealing settings, for a given level of information diagnosticity, the likelihood ratio of the receiver observing a signal conditional on the underlying state, is as follows: $l(s=r|q_H) = \frac{q_H}{1-q_H} = \infty$, $l(s=b|q_H) = \frac{1-q_H}{q_H} = 0$, $l(s=r|q_L) = \frac{q_L}{1-q_L} = 3$, and $l(s=b|q_L) = \frac{1-q_L}{q_L} = \frac{1}{3}$. In Not Revealing settings, the corresponding likelihood ratios are: $l(s=r|q_H) = \frac{q_H}{1-q_H} = 19$, $l(s=b|q_H) = \frac{1-q_H}{q_H} = \frac{1}{19}$, $l(s=r|q_L) = \frac{q_L}{1-q_L} = \frac{7}{3}$, and $l(s=b|q_L) = \frac{1-q_L}{q_L} = \frac{3}{7}$.

In the Endogenous-Revealing (EN-R) setting, the persuader's equilibrium strategy is to choose $(p_H, q_H) = (0.6, 1)$ and select $p_L = 0.2$ after choosing $q_L = 0.75$. The likelihood ratio of the persuader choosing a given level of information diagnosticity is $m(q_H) = \frac{\pi(q_H|p_H)}{\pi(q_H|p_L)} = \frac{1}{0} = \infty$ and $m(q_L) = \frac{\pi(q_L|p_H)}{\pi(q_L|p_L)} = \frac{0}{1} = 0$. The receiver's equilibrium initial belief that the underlying state is ω_1 is $\mu_0 = 0.6$. Her interim beliefs, conditional on the observed diagnosticity, are: $\mu_1(q_H) = p_L + (p_H - p_L) * \frac{1}{1+\frac{p_H-\mu_0}{\mu_0-p_L} * \frac{1}{m(q_H)}} = 0.6$ and $\mu_1(q_L) = 0.2$. Her posterior belief after observing a signal s and information diagnosticity q, is given by $\mu_2 = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1}{l(s|q)}}$. In information set $I_1 - I_4$, the resulting posterior beliefs are: 1, $\frac{3}{7}$, $\frac{1}{13}$ and 0.

In the Endogenous-Not Revealing (EN-NR) setting, the persuader's equilibrium strategy is to choose $(p_H, q_H) = (0.6, 0.95)$ and select $p_L = 0.2$ after choosing $q_L = 0.70$. The likelihood ratio of the persuader choosing information diagnosticity is: $m(q_H) = \frac{\pi(q_H|p_H)}{\pi(q_H|p_L)} = \frac{1}{0} = \infty$ and $m(q_L) = \frac{\pi(q_L|p_H)}{\pi(q_L|p_L)} = \frac{0}{1} = 0$. The receiver's equilibrium initial belief that the underlying state is ω_1 is $\mu_0 = 0.6$. Her interim beliefs are $\mu_1(q_H) = 0.6$ and $\mu_1(q_L) = 0.2$. Her posterior belief is computed as: $\mu_2 = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1}{l(s|q)}}$, with the corresponding values being $\frac{57}{59}$, $\frac{7}{19}$, $\frac{3}{31}$ and $\frac{3}{41}$ in information set $I_1 - I_4$, respectively.

In the Exogenous-Revealing (EX-R) setting, the persuader's pooling equilibrium strategy is given by $\pi(q_H|\omega_1) = \pi(q_H|\omega_2) = 1$. The receiver's initial belief that the underlying state is ω_1 is $\mu_0 = p^* = 0.4$. The likelihood ratio associated with the persuader's choice of information diagnosticity q_H is $m(q_H) = \frac{\pi(q_H|\omega_1)}{\pi(q_H|\omega_2)} = \frac{1}{1} = 1$. Given this, her interim belief after observing q_H is $\mu_1(q_H) = \frac{1}{1+\frac{1-\mu_0}{\mu_0}\frac{1}{m(q_H)}} = 0.4$. Based on the divinity refinement of Banks and Sobel (1987), her interim belief after observing q_L must satisfy $\mu_1(q_L) < \mu_1(I_2) = \frac{4}{19}$, where $\mu_1(I_2)$ is derived from the equation: $\frac{1}{1+\frac{1-\mu_1(I_2)}{\mu_1(I_2)}\frac{1-q_L}{q_L}} = \hat{u} = \frac{4}{9}$. Therefore, the likelihood

[60]Notably, with our parameter choices ($p_H = 0.6$, $p_L = 0.2$ in endogenous settings and $p^* = 0.4$ in exogenous settings), a naive receiver who ignores both information diagnosticity and realized signals would choose to reject (action R) since the threshold belief of choosing to accept is $\hat{\mu} = \frac{4}{9} \approx 0.44$.

ratio for q_L must satisfy: $m(q_L) < \frac{1-\mu_0}{\mu_0} * \frac{\mu_1(I_2)}{1-\mu_1(I_2)} = \frac{2}{5}$, while $m(q_L) = \frac{\pi(q_L|\omega_1)}{\pi(q_L|\omega_2)} = \frac{0}{0}$ is not well defined. Her posterior belief, denoted: $\mu_2 = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1}{l(s|q)}}$, takes on the following values in each information set: $\mu_2 = 1$ in I_1, $\mu_2(I_2) \in [0, \frac{4}{9})$ in I_2, $\mu_2 = \frac{\mu_2(I_2)}{9-8\mu_2(I_2)} \in [0, \frac{4}{49})$ in I_3, and $\mu_2 = 0$ in I_4. In the separating equilibrium, the persuader's strategy is $\pi(q_H|\omega_1) = \pi(q_L|\omega_2) = 1$. This results in likelihood ratios: $m(q_H) = \frac{\pi(q_H|\omega_1)}{\pi(q_H|\omega_2)} = \infty$ and $m(q_L) = \frac{\pi(q_L|\omega_1}{\pi(q_L|\omega_2)} = 0$. The receiver's initial belief remains $\mu_0 = p^* = 0.4$, while her her interim beliefs are: $\mu_1(q_H) = 1$ and $\mu_1(q_L) = 0$. Her posterior belief μ_2, again computed as: $\mu_2 = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1}{l(s|q)}}$, takes the values 1, 0, 0 and 0 in $I_1 - I_4$, respectively.

In the Exogenous-Not Revealing (EX-NR) setting, the persuader's pooling equilibrium strategy is $\pi(q_H|\omega_1) = \pi(q_H|\omega_2) = 1$. The receiver's initial belief that the underlying state is ω_1 is $\mu_0 = p^* = 0.4$. The likelihood ratio associated with the persuader's choice of information diagnosticity q_H is $m(q_H) = \frac{\pi(q_H|\omega_1)}{\pi(q_H|\omega_2)} = \frac{1}{1} = 1$. Her interim belief after observing q_H is $\mu_1(q_H) = \frac{1}{1+\frac{1-\mu_0}{\mu_0}\frac{1}{m(q_H)}} = 0.4$. According to the divinity refinement, her interim belief after observing q_L must satisfy: $\mu_1(q_L) < \mu_1(I_2) = \frac{12}{47}$, where $\mu_1(I_2)$ is the solution to the equation: $\frac{1}{1+\frac{1-\mu_1(I_2)}{\mu_1(I_2)}\frac{1-q_L}{q_L}} = \hat{u} = \frac{4}{9}$. This implies that the likelihood ratio of choosing q_L must satisfy: $m(q_L) < \frac{1-\mu_0}{\mu_0} * \frac{\mu_1(I_2)}{1-\mu_1(I_2)} = \frac{18}{35}$, while $m(q_L) = \frac{\pi(q_L|\omega_1)}{\pi(q_L|\omega_2)} = \frac{0}{0}$ is not well defined. Her posterior belief is given by: $\mu_2 = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1}{l(s|q)}}$. The corresponding posterior beliefs across information sets $I_1 - I_4$ are: $\frac{38}{41}$, $[0, \frac{4}{9})$, $\frac{9\mu_2(I_2)}{49-40\mu_2(I_2)} \in [0, \frac{36}{281})$, and $\frac{2}{59}$.

B. Additional Figures and Tables

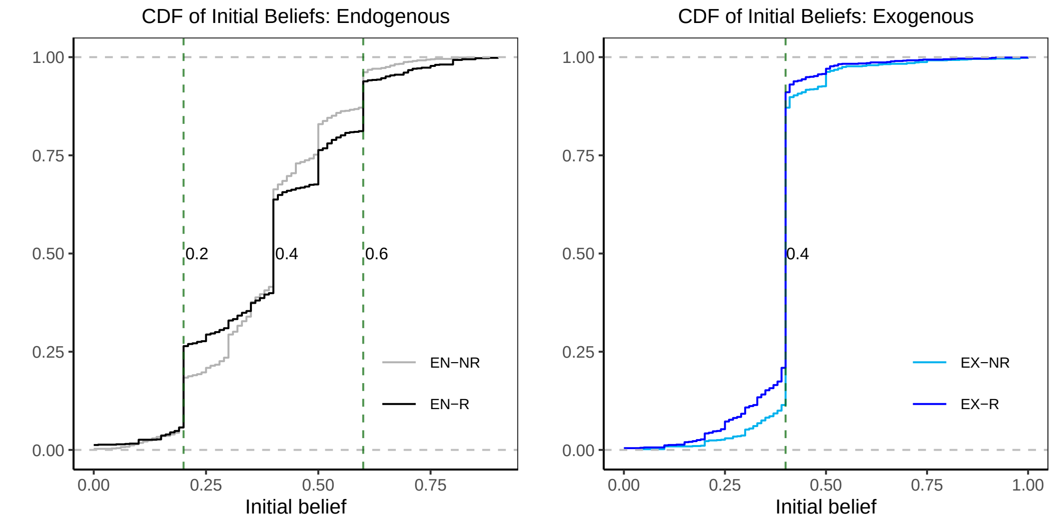


FIGURE 7. DISTRIBUTION OF INITIAL BELIEFS

Notes: The figure displays the distribution of receivers' initial beliefs about the underlying state being ω_1 (i.e., the red box), prior to learning persuaders' choices of information diagnosticity q, across all four settings. EN and EX refer to the endogenous and exogenous settings, while R and NR refer to the Revealing and Not Revealing settings, respectively.

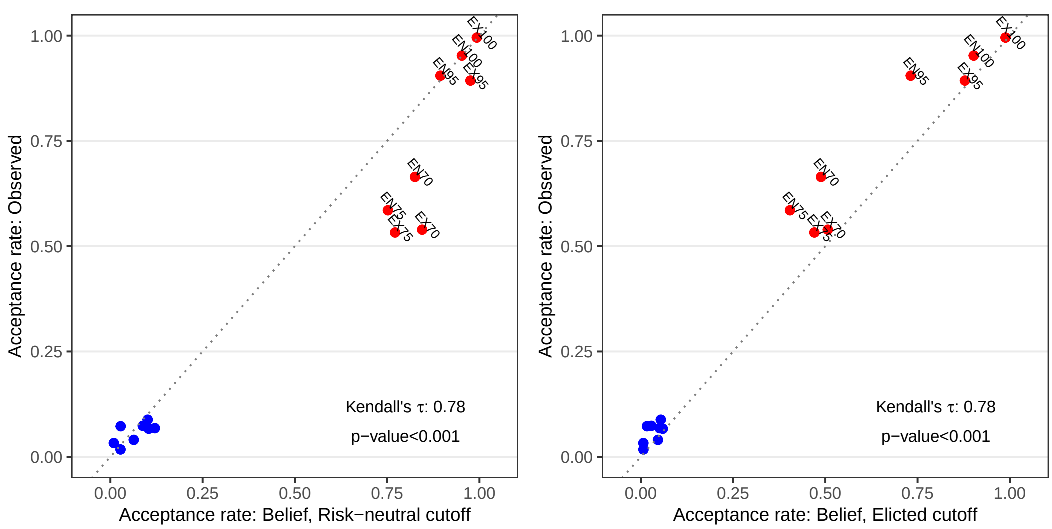


FIGURE 8. ACTION CHOICES OF RECEIVERS: OBSERVED VERSUS INDICATED BY POSTERIOR BELIEFS

Notes: Each data point represents a possible information set (q, s) observed by receivers. Red (blue) dots indicate information sets where $s = r$ $(s = b)$. EN and EX refer to the endogenous and exogenous settings, respectively, with the numbers next to EN or EX indicating the persuaders' chosen q in percentage points. The y-axis in both panels shows the observed likelihood of receivers choosing to accept under each information set. The x-axis in the left panel shows the likelihood of acceptance implied by receivers' posterior beliefs, assuming a risk-neutral threshold probability of $\frac{4}{9}$ for the underlying state being ω_1, which makes "accept" a weakly better action choice. The x-axis in the right panel uses each receiver's self-reported threshold probability for choosing "accept".

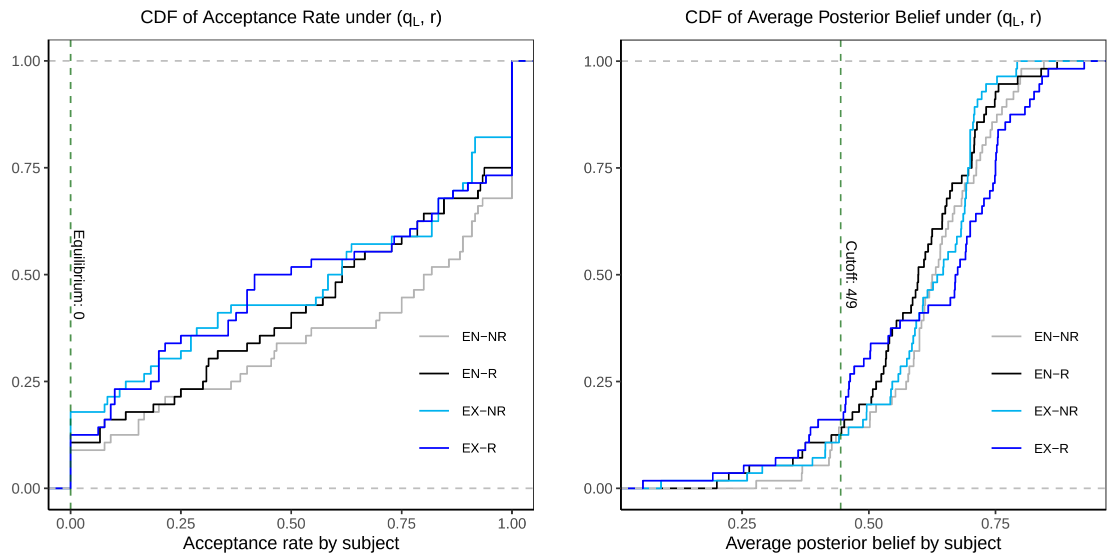


FIGURE 9. DISTRIBUTION OF ACTION CHOICES AND AVERAGE POSTERIOR BELIEFS UNDER (q_L, r)

Notes: EN and EX refer to endogenous and exogenous settings. Equilibrium analysis predicts that a rational receiver should choose not to accept when observing (q_L, r), i.e., the acceptance rate being 0. $\frac{4}{9}$ is the threshold probability of the underlying state being ω_1 that makes "accept" a (weakly) better choice.

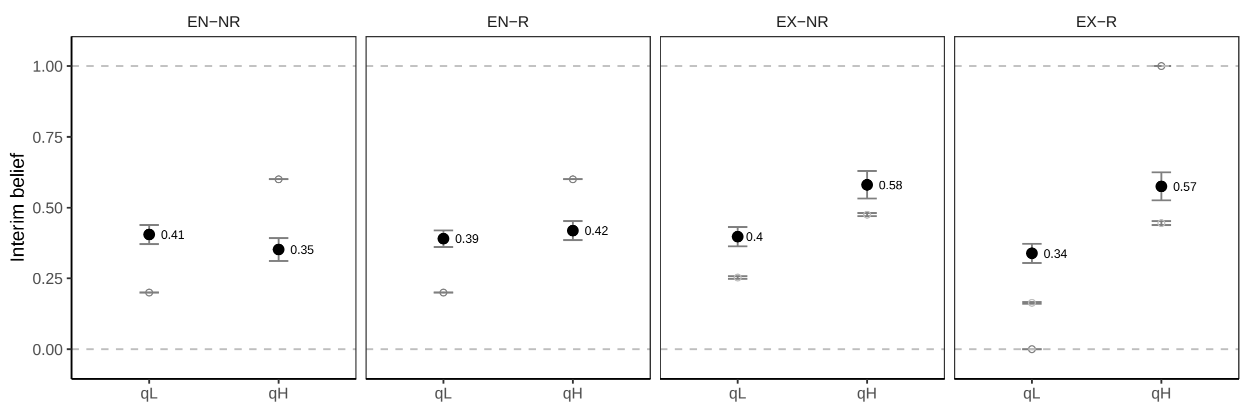


FIGURE 10. AVERAGE INTERIM BELIEFS OF RECEIVERS: OBSERVED VERSUS COUNTERFACTUAL

Notes: Black dots represent the average interim beliefs about the underlying state being ω_1. Gray hollow dots represent the average interim beliefs computed under the assumption that receivers made only probabilistic error (based on the estimated α and β in Table 3) and no strategic error. Short vertical lines are 95% confidence intervals. EN-R, EN-NR, EX-R, and EX-NR refer to Endogenous-Revealing, Endogenous-Not Revealing, Exogenous-Revealing, and Exogenous-Not Revealing, respectively. In endogenous settings and under the separating equilibrium of EX-R setting, the computed interim beliefs align exactly with the equilibrium predictions, as the probabilistic error becomes inconsequential when the strategic error is absent.

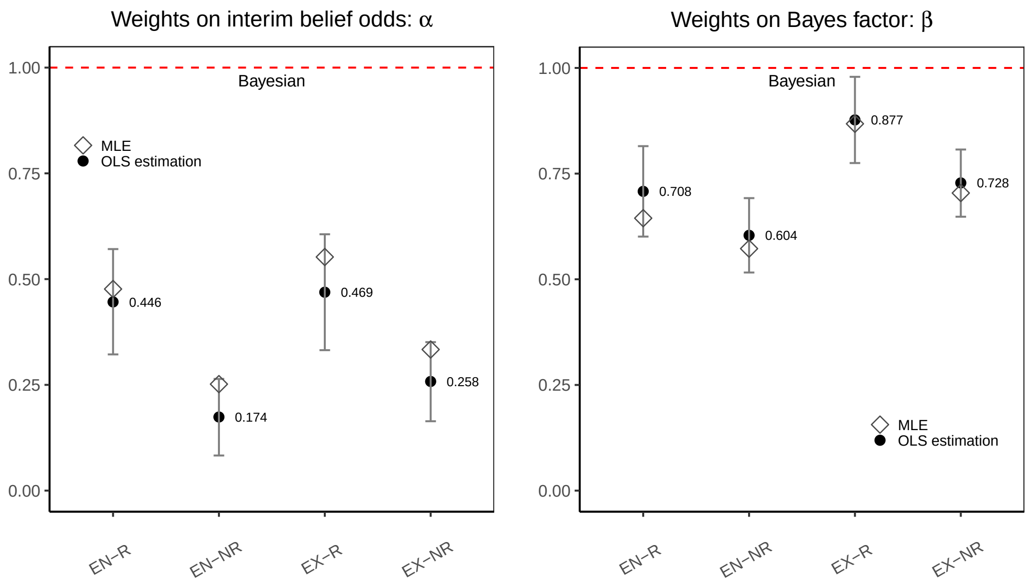


Figure 11. Estimation of Non-Bayesian Updating: OLS versus MLE

Notes: The estimations focus on the belief update from interim to posterior beliefs. The OLS estimation results also appear in Table 3. Short vertical lines represent the 95% confidence intervals of the OLS estimates. MLE refers to the estimation results of the belief-updating specification described in Section IV.A, conducted at the setting level.

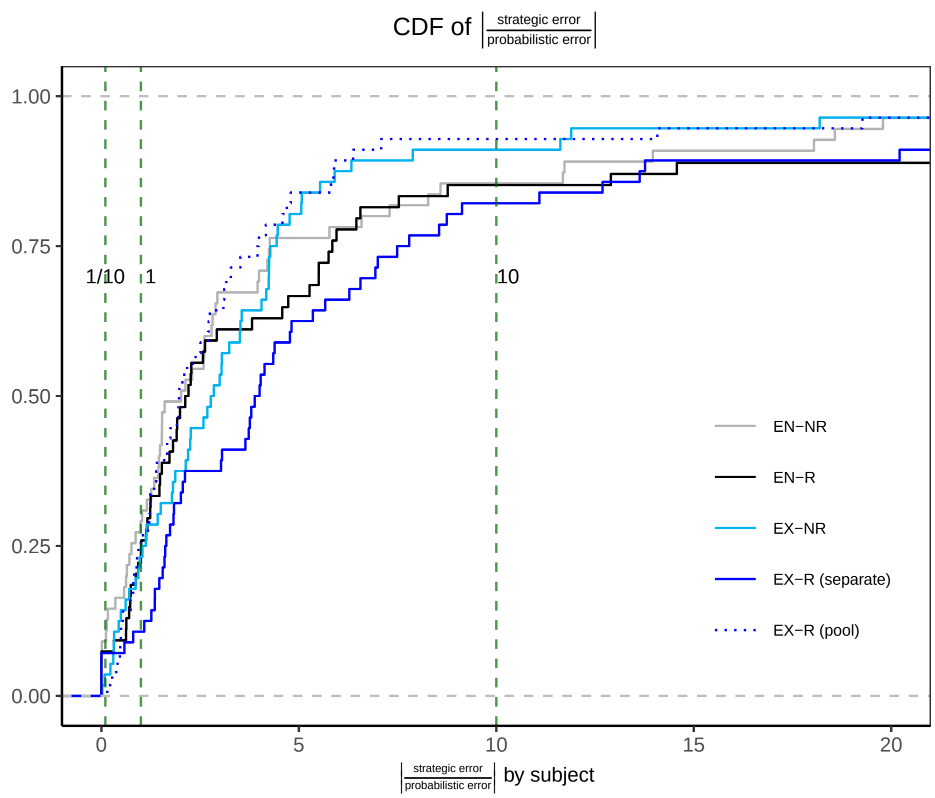


Figure 12. Distribution of Impact Size Ratios: Strategic versus Probabilistic Errors

Notes: The measures of impact sizes of two reasoning errors are the same as in Figure 5. CDF lines do not reach exactly 1 because truncated plots are used for better readability.

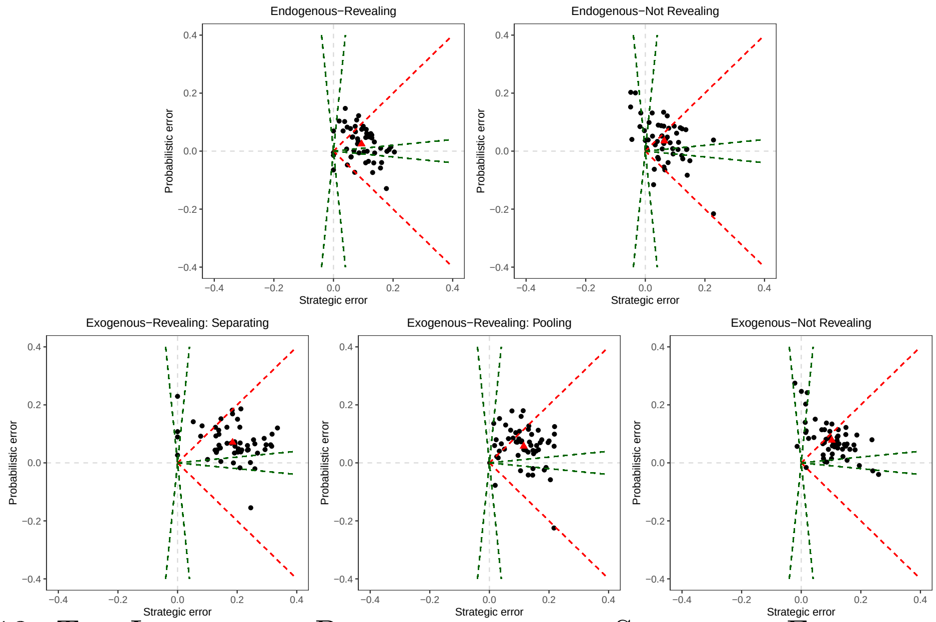


FIGURE 13. THE IMPACTS OF PROBABILISTIC AND STRATEGIC ERRORS: ALL (q, s)

Notes: The analysis considers all possible information sets (q, s). The plots illustrate the impacts of probabilistic and strategic errors on posterior beliefs (similar to those shown in Figure 5) across receivers in the four experimental settings. The impacts of probabilistic and strategic reasoning errors are comparable for 74%, 89%, 84%/96%, and 84% of receivers, and strategic error has a relatively greater impact than the probabilistic error for 72%, 64%, 86%/70% and 71% of receivers in the EN-R, EN-NR, EX-R (under separating/pooling equilibria), and EX-NR settings, respectively.

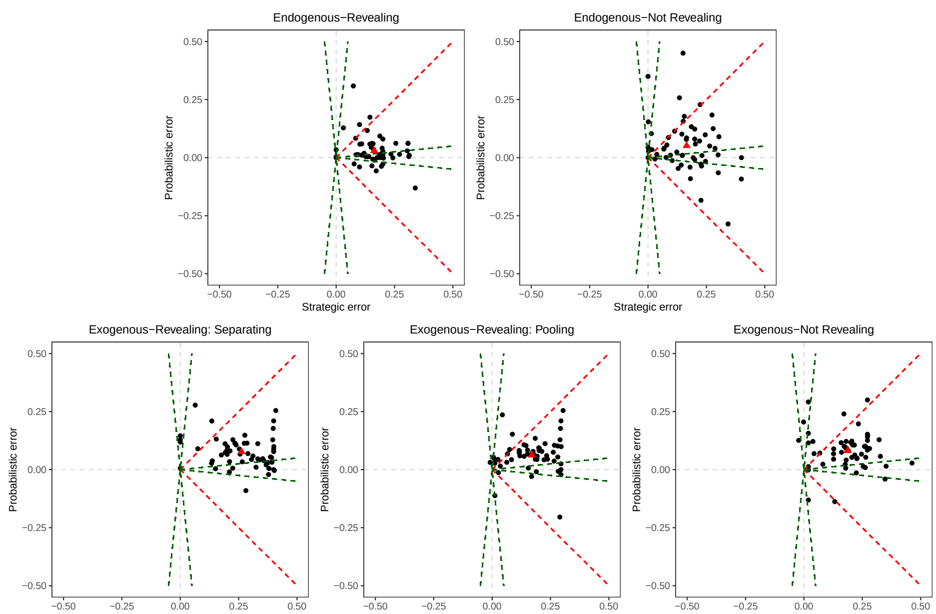


FIGURE 14. THE IMPACTS OF PROBABILISTIC AND STRATEGIC ERRORS ON INTERIM BELIEFS

Notes: The analysis focuses on the information set associated with q_L. The plots illustrate the impacts of strategic error, defined as $\mu_{i1}(\chi_i^{est}, \alpha_i = \beta_i = 1) - \mu_1^e$, and probabilistic errors, $\mu_{i1}^{ob} - \mu_{i1}(\chi_i^{est}, \alpha_i = \beta_i = 1)$, on interim beliefs given the information diagnosticity q_L, across receivers in the four experimental settings. The impacts of probabilistic and strategic reasoning errors are comparable for 59%, 76%, 73%/89%, and 84% of receivers, and strategic error has a relatively greater impact than the probabilistic error for 89%, 80%, 88%/82% and 79% of receivers in the EN-R, EN-NR, EX-R (under separating/pooling equilibria), and EX-NR settings, respectively.

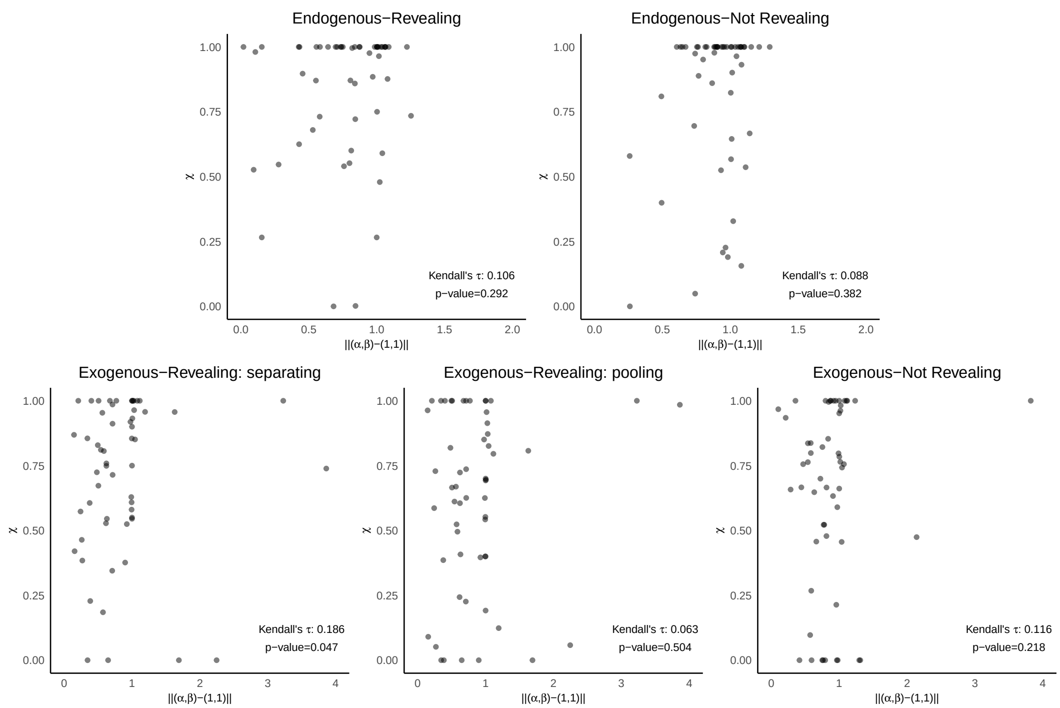


Figure 15. Measures of Probabilistic and Strategic Errors

Notes: α, β, and χ are behavioral model parameters as defined in Section IV.A. Each data point corresponds to a receiver in the experiment.

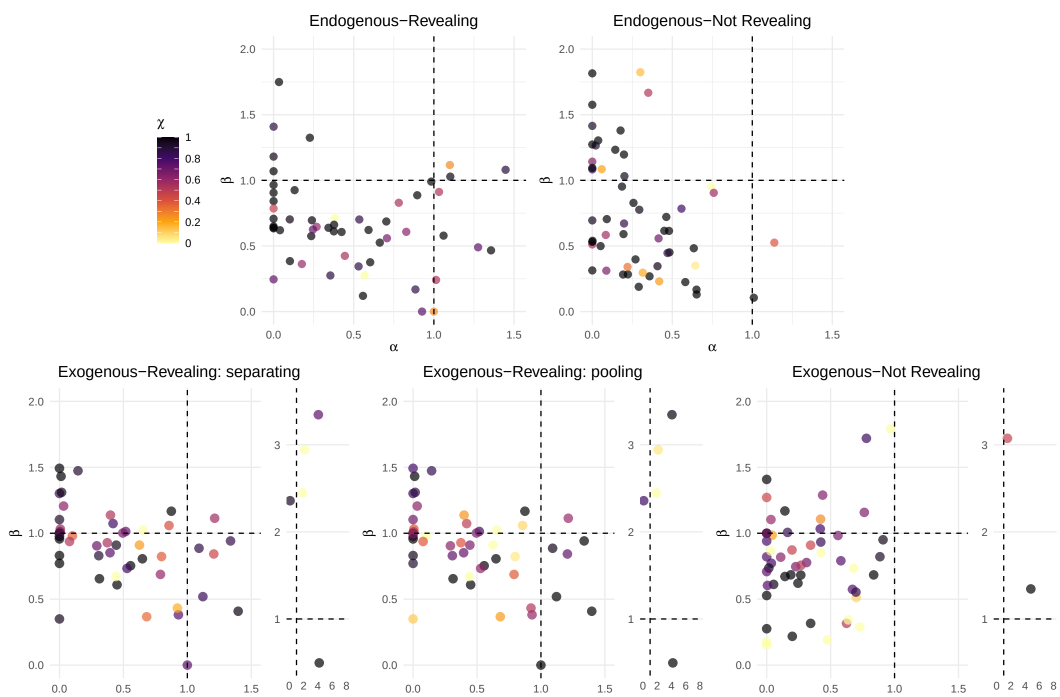


Figure 16. Estimated Behavioral Model Parameters (α, β, χ) Across Receivers

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. Each data point corresponds to a receiver in the experiment. Black dashed lines represent the reference values $\alpha = 1$ and $\beta = 1$. In each of the three plots in the second row, the main cluster of data and the more extreme data are shown separately to improve readability.

Table 6—Impacts of Individual Characteristics on Strategic and Probabilistic Reasoning Errors

	χ		$\|\|(\alpha, \beta) - (1,1)\|\|$
	(1)	(2)	(3)
Raven	-0.033	-0.036	-0.029
	(0.014)	(0.015)	(0.021)
pstat	0.026	0.003	-0.135
	(0.046)	(0.047)	(0.067)
game theory	-0.057	-0.050	-0.026
	(0.057)	(0.059)	(0.084)
risk preference	0.004	0.014	-0.029
	(0.012)	(0.013)	(0.018)
gender	0.058	0.069	-0.053
	(0.044)	(0.046)	(0.065)
Endogenous	0.158	0.205	-0.033
	(0.041)	(0.042)	(0.060)
Revealing	0.039	-0.008	-0.049
	(0.041)	(0.043)	(0.061)
Constant	0.844	0.788	1.411
	(0.139)	(0.143)	(0.205)
N	221	221	224

Notes: α, β, and χ are behavioral model parameters defined in Section IV.A. The OLS regressions use two dependent variables: the measure of strategic naivete, χ (columns 1-2), and the deviation from Bayesian updating, $||(\alpha, \beta) - (1,1)||$ (column 3). In the EX-R setting, the estimate of χ differs depending on whether the analysis conditions on the separating equilibrium or the pooling one. Columns (1) and (2) report regression results using each of these estimates, respectively. Raven denotes scores from Raven's test completed by participants at the end of the experiment. Endogenous and Revealing are indicator variables for the four experimental settings. pstat and game theory are indicator variables for whether a participant has taken any probability and statistics courses or game theory courses, respectively. Risk preference is the switching point in the multiple-price-list task, where a higher value indicates greater risk aversion. Gender equals 1 for male participants and 0 for female participants. The χ parameter could not be estimated for three participants (two from EN-R and one from EN-NR). Standard errors are reported in parentheses.

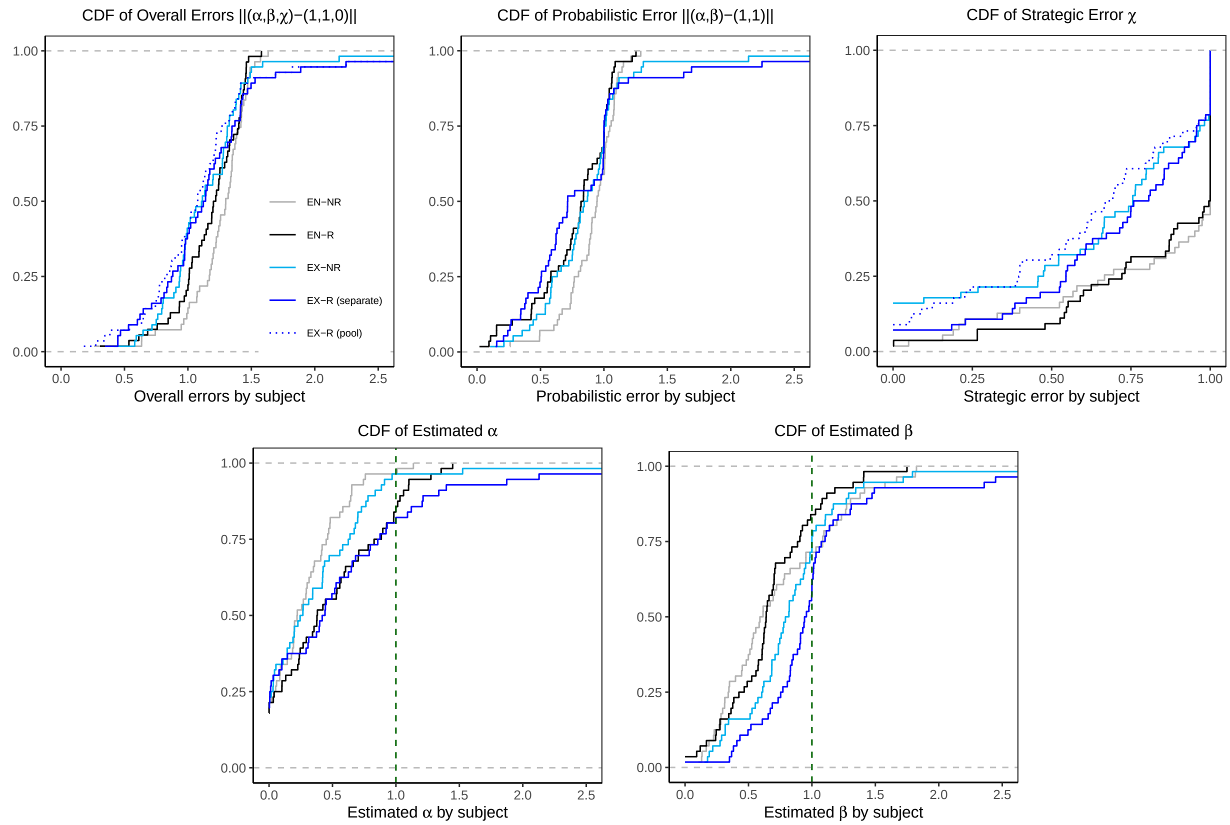


FIGURE 17. ERROR COMPARISON ACROSS SETTINGS

Notes: Each plot shows the distribution of individual-level error measures across settings. Some CDF lines do not reach exactly 1 because truncated plots are used for better readability, except in the case of χ.

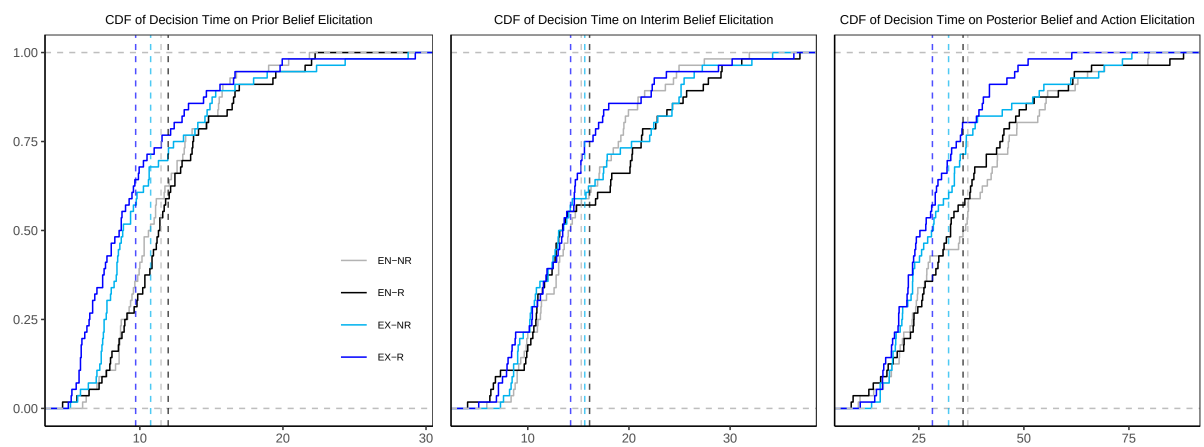


FIGURE 18. DISTRIBUTION OF DECISION TIME OF RECEIVERS

Notes: The vertical dashed lines indicate the aggregate average decision time (in seconds) of receivers in the same setting.

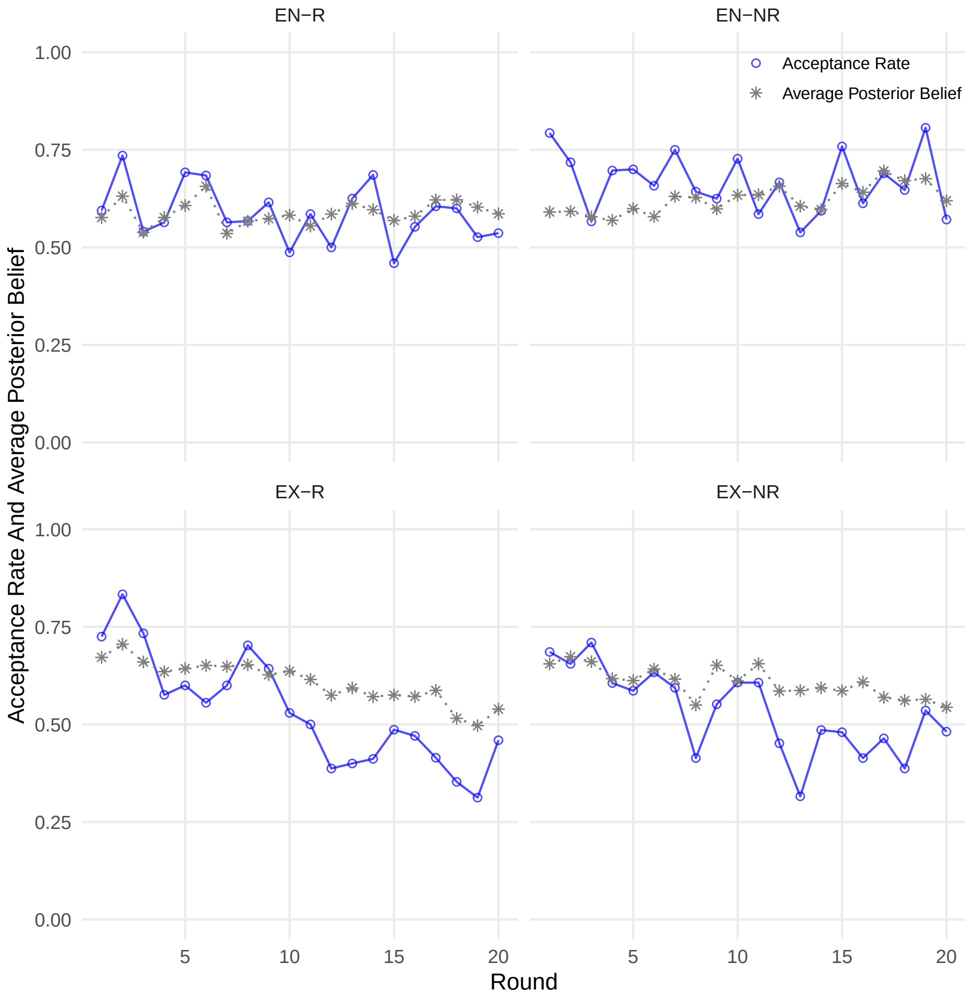


FIGURE 19. ACTION CHOICES AND POSTERIOR BELIEFS OF RECEIVERS UNDER (q_L, r) BY ROUND

Notes: EN-R, EN-NR, EX-R, and EX-NR refer to Endogenous-Revealing, Endogenous-Not Revealing, Exogenous-Revealing, and Exogenous-Not Revealing, respectively. The four plots correspond to the (q_L, r) case under the EN-R, EN-NR, EX-R and EX-NR settings, respectively.

C. Robustness

C.1. *Comparable in Magnitude*

C.1.1. *Reduced-form Analysis to Quantify Impacts of Two Errors*

Our experimental design allows us to isolate the impact of each type of reasoning error. We begin by estimating the receivers' non-Bayesian updating parameters, (α, β), using both interim and posterior beliefs. This estimation follows the same regression approach described in Section III.C.1. Next, we infer the receivers' strategic reasoning, captured by the diagnosticity measure $m(q)(\equiv \frac{\pi(q|p_H)}{\pi(q|p_L)}$ or $\frac{\pi(q|\omega_1)}{\pi(q|\omega_2)})$, by examining how initial beliefs are updated to interim beliefs, taking into account the estimated (α, β). Using the identified parameters $(m(q), \alpha, \beta)$, we then quantify the separate contributions of each type of reasoning error to the deviation from the equilibrium prediction.

Specifically, the belief update in the first stage can be expressed, in endogenous settings, as: $\frac{\mu_{p_H|q}}{1-\mu_{p_H|q}} = (\frac{\mu_{p_H}}{1-\mu_{p_H}})^{\alpha}(m(q))^{\beta}$, where $\mu_{p_H|q} \equiv P(p_H|q) = \frac{\mu_1 - p_L}{p_H - p_L}$ and $\mu_{p_H} \equiv P(p_H) = \frac{\mu_0 - p_L}{p_H - p_L}$. In exogenous settings, it is expressed as: $\frac{\mu_1}{1-\mu_1} = \left(\frac{\mu_0}{1-\mu_0}\right)^{\alpha}(m(q))^{\beta}$. Taking logarithms, we can solve for $\ln m(q) = \frac{\ln \frac{\mu_{p_H|q}}{1-\mu_{p_H|q}} - \alpha \ln \frac{\mu_{p_H}}{1-\mu_{p_H}}}{\beta}$ and $\ln m(q) = \frac{\ln \frac{\mu_1}{1-\mu_1} - \alpha \ln \frac{\mu_0}{1-\mu_0}}{\beta}$ in endogenous and exogenous settings, respectively. Let μ_2^{ob} and μ_2^{e} denote the observed and equilibrium posterior beliefs, respectively. Let $m^{est}(q_H)$, $m^{est}(q_L)$, α^{est}, and β^{est} represent the estimated subjective likelihood and belief update parameters. The deviation of posterior beliefs from the equilibrium prediction can be decomposed into two parts:

$$\mu_2^{ob} - \mu_2^{e} = \mu_2(\mu_0^{ob}, m^{est}(q), \alpha = \beta = 1) - \mu_2^{e} \tag{1}$$

$$+ \mu_2^{ob} - \mu_2(\mu_0^{ob}, m^{est}(q), \alpha = \beta = 1), \tag{2}$$

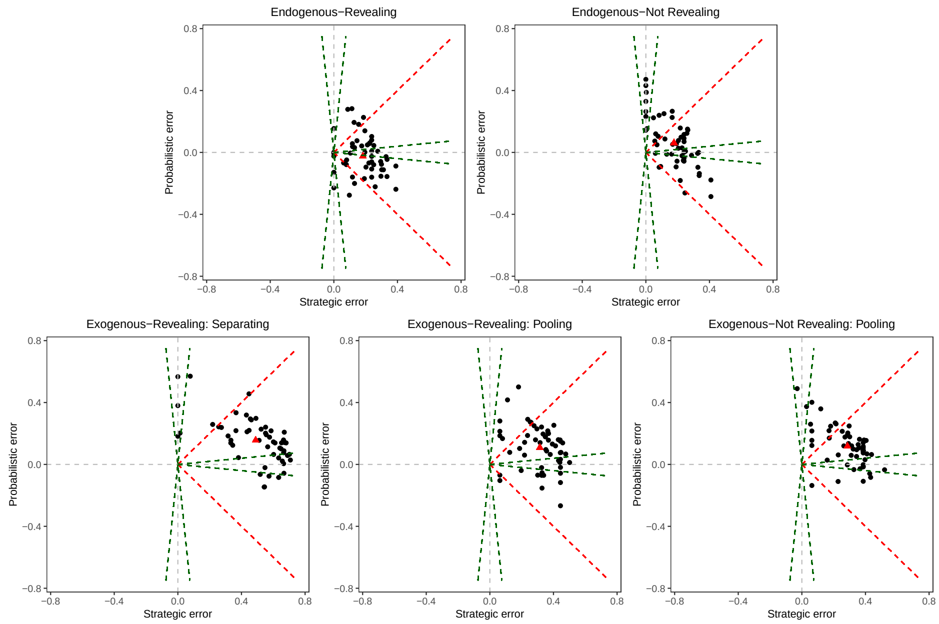
where the two parts capture the impacts of strategic and probabilistic reasoning errors on the deviation of posterior beliefs, respectively.

We find that for a substantial majority of receivers, the impact of one error is not more than ten times that of the other: 89%, 79%, 86%/82%, and 81% in EN-R, EN-NR, EX-R (separating/pooling equilibrium), and EX-NR settings, respectively.[61,62]

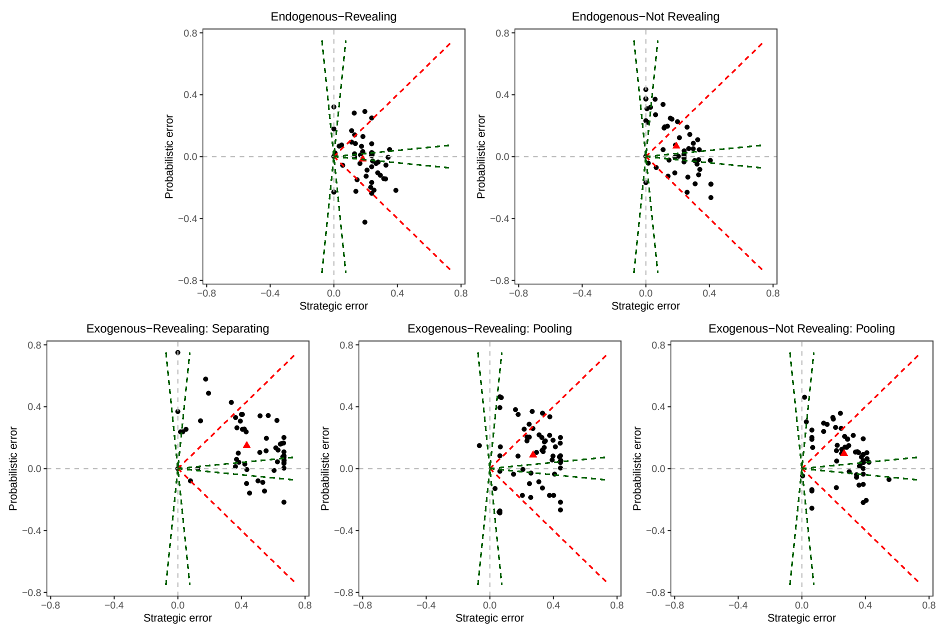
[61] Considering all possible information sets (q, s), the percentages are 92%, 83%, 86%/92% and 87%.

[62] Since the logarithms of belief odds in the expression for $\ln m(q)$ become undefined when beliefs (i.e., $\mu_{p_H|q}$ and μ_{p_H} in endogenous settings, and μ_0 and μ_1 in exogenous settings) are 0 or 1, in this analysis, we exclude observations with degenerate beliefs. In endogenous settings, we also exclude observations where μ_0, $\mu_1 \notin [p_L, p_H]$, because in these cases $\mu_{p_H|q}$ and μ_{p_H} are not well defined.

C.1.2. *Robust to Other Ways of Parameter Estimation*



(a) First 10 Rounds



(b) Last 10 Rounds

Figure 20. The Impacts of Probabilistic and Strategic Errors: First 10 and Last 10 Rounds *Notes:*

Notes: The plots replicate Figure 5, but the (individual-level) structural estimation and the computation of the impacts of the two errors on posterior beliefs (in information set (q_L, r)) use only the first or last 10 rounds of data. In panel (a), the impacts of probabilistic and strategic reasoning errors are comparable in magnitude for 77%, 69%, 79%/89%, and 84% of receivers in the EN-R, EN-NR, EX-R (conditional on the separating/pooling equilibrium), and EX-NR settings. In panel (b), the corresponding percentages are 76%, 67%, 77%/93%, and 86%. In EX-R (pooling) of panel (b), one receiver exhibits a negative measure of strategic error. This is driven by the influence of deviations in initial beliefs.

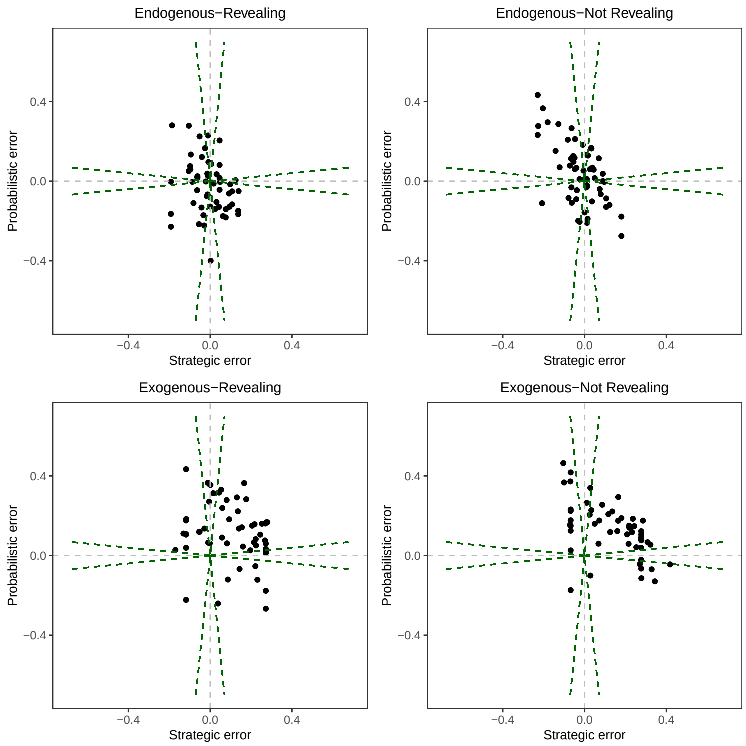


Figure 21. The Impacts of Probabilistic and Strategic Errors: Empirically Optimal Benchmarks

Notes: The plots replicate Figure 5, but the (individual-level) structural estimation and the computation of the impacts of the two errors on posterior beliefs (in information set (q_L, r)) use empirically optimal benchmarks, which are derived from the distribution of persuaders' actual choices, rather than equilibrium benchmarks. The impacts of probabilistic and strategic reasoning errors are comparable in magnitude for 80%, 84%, 86%, and 93% of receivers in the EN-R, EN-NR, EX-R (conditional on the separating/pooling equilibrium), and EX-NR settings, respectively.

C.2. *Orthogonality*

Robust to other ways of parameter estimation

TABLE 7—CORRELATION BETWEEN STRATEGIC NAIVETE AND DEVIATION FROM BAYESIAN UPDATING: ROBUSTNESS CHECKS

		Correlation($\|(\alpha, \beta) - (1, 1)\|$, χ)		
		(1)	(2)	(3)
		First 10 rounds	Last 10 rounds	Empirical (all 20 rounds)
EN-R		0.054 (0.610) [-0.150, 0.258]	0.125 (0.217) [-0.062, 0.312]	0.117 (0.250) [-0.079, 0.312]
EN-NR		0.019 (0.853) [-0.191, 0.228]	0.208 (0.046) [-0.002, 0.417]	-0.002 (0.987) [-0.208, 0.204]
EX-R	separating	0.107 (0.253) [-0.101, 0.314]	0.164 (0.079) [-0.020, 0.348]	0.165 (0.081) [-0.032, 0.363]
	pooling	0.144 (0.125) [-0.042, 0.330]	0.117 (0.216) [-0.068, 0.302]	
EX-NR		0.133 (0.170) [-0.067, 0.333]	0.072 (0.450) [-0.131, 0.275]	0.163 (0.095) [-0.044, 0.369]

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. The table presents Kendall's τ as correlation coefficients, with the numbers in parentheses indicating p-values and the numbers in brackets representing the 95% confidence intervals of the correlation coefficients. We conduct individual-level structural estimations of the behavioral model parameters using three additional approaches: (1) using only the first 10 rounds of data with equilibrium benchmarks, (2) using only the last 10 rounds of data with equilibrium benchmarks, and (3) using all 20 rounds with empirical optimal benchmarks, which are derived from the persuaders' actual choices.

Robust to adjustment for potential measurement error

To further validate our finding that probabilistic and strategic errors are orthogonal, we apply the Obviously Related Instrumental Variables (ORIV) method proposed by Gillen, Snowberg and Yariv (2019), which accounts for potential measurement errors that may otherwise lead to an underestimation of the correlation between characteristics in experimental data.

In our setting, the parameters capturing both types of reasoning errors are structurally estimated rather than directly elicited. Accordingly, measurement error refers to the discrepancy between the latent true parameter values and their estimates. The ORIV method addresses such measurement error by using duplicate estimates of the variables of interest as instruments. Under the assumption that measurement errors in the duplicates are orthogonal, ORIV can consistently correct for these errors. Even when this assumption does not

Table 8—Correlation between Strategic Naivete and Deviation from Bayesian Updating: Adjusted for Measurement Error Using ORIV

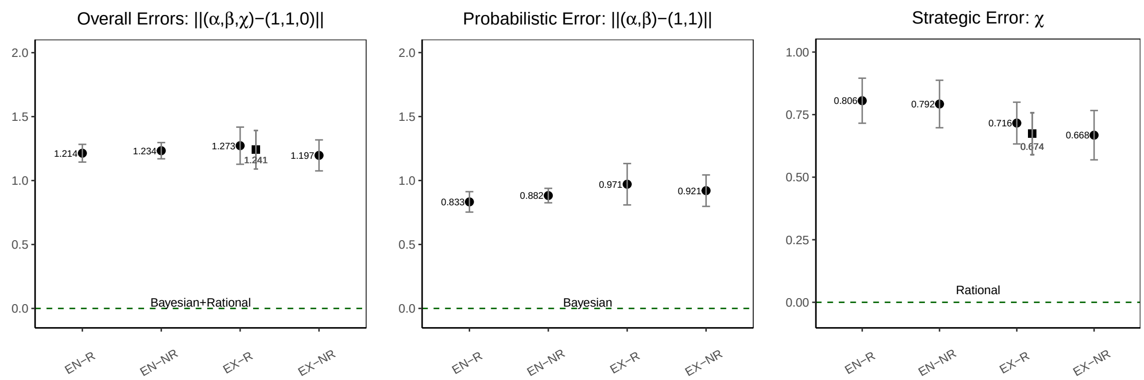
		Correlation($\|(\alpha, \beta) - (1,1)\|$, χ) (1)	(2)
EN-R		0.107 (0.341)	0.200 (0.262)
EN-NR		0.197 (0.166)	0.118 (0.142)
EX-R	Separating	0.097 (0.114)	0.275 (0.136)
	Pooling	0.231 (0.097)	0.221 (0.128)
EX-NR		0.076 (0.136)	0.132 (0.151)

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. The table presents adjusted correlation coefficients using the ORIV method from Gillen, Snowberg and Yariv (2019). Column (1) reports the results based on the original values of the estimated behavioral parameters, corresponding to the Pearson linear correlation coefficient. Column (2) presents the results derived from the ranks of behavioral parameters, approximating rank-based correlation coefficients. The numbers in parentheses represent the standard errors obtained through Bootstrapping.

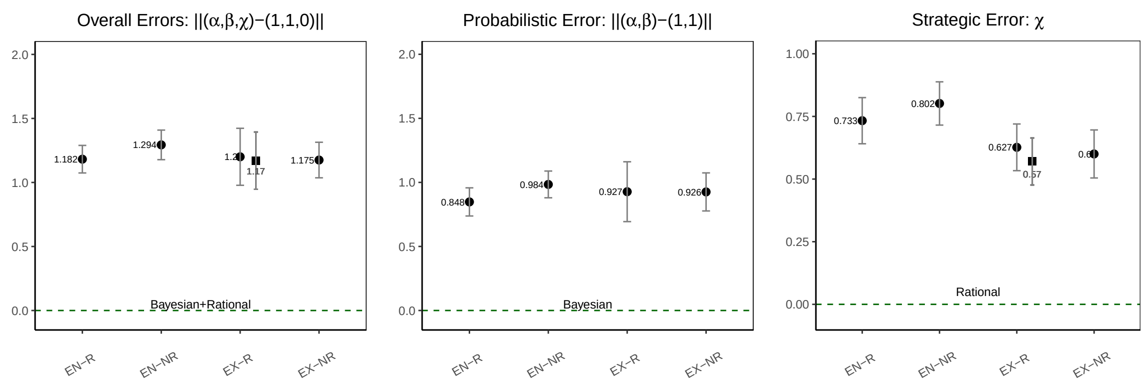
fully hold, Gillen, Snowberg and Yariv (2019) recommend using ORIV, as it can still correct for at least some component of measurement error. We implement ORIV by instrumenting the error parameters estimated from the data of odd-numbered rounds with those estimated from even-numbered rounds, and vice versa.

Table 8 presents the estimated correlation coefficients using the ORIV method. Column (1) reports results based on the original estimated values of the behavioral parameters, corresponding to the Pearson correlation coefficient. Column (2) shows results based on the ranks of the behavioral parameters, which approximate rank-based correlation measures. After accounting for potential measurement error, the correlation between the two types of reasoning errors remains small across all four settings.

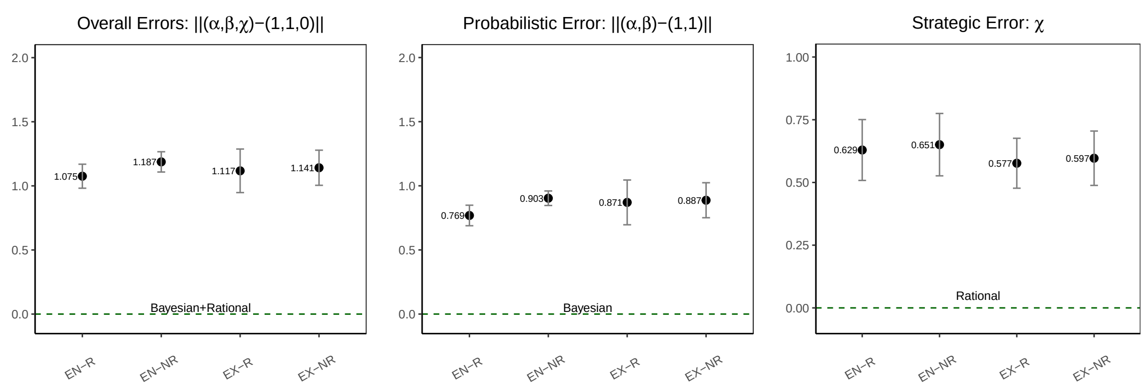
C.3. *Stability*



(a) First 10 Rounds



(b) Last 10 Rounds



(c) Empirical

FIGURE 22. ERROR COMPARISON ACROSS SETTINGS: ROBUSTNESS CHECK

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. The plots replicate Figure 6, and the displayed results correspond to our three additional approaches of estimating individual-level behavioral model parameters: (a) using only the first 10 rounds of data with equilibrium benchmarks, (b) using only the last 10 rounds with equilibrium benchmarks, and (c) using all 20 rounds with empirical optimal benchmarks, which are derived from the persuaders' actual choices. Under any approach, on average, the degree of overall errors (shown on the left of each panel) is not significantly different across settings ($p > 0.05$ for all pairwise comparisons under the two-sample t-test).

C.4. *Estimating* (α, β, χ) *Jointly*

We also jointly estimate the three behavioral model parameters (α, β, χ) using the elicited initial, interim, and posterior beliefs altogether. Based on these estimates, we conduct the same analyses on the comparability, orthogonality, and stability of probabilistic and strategic errors as shown in the main text and present the results below. Our key findings on the comparability and orthogonality of the two reasoning errors remain robust even when using joint estimates.

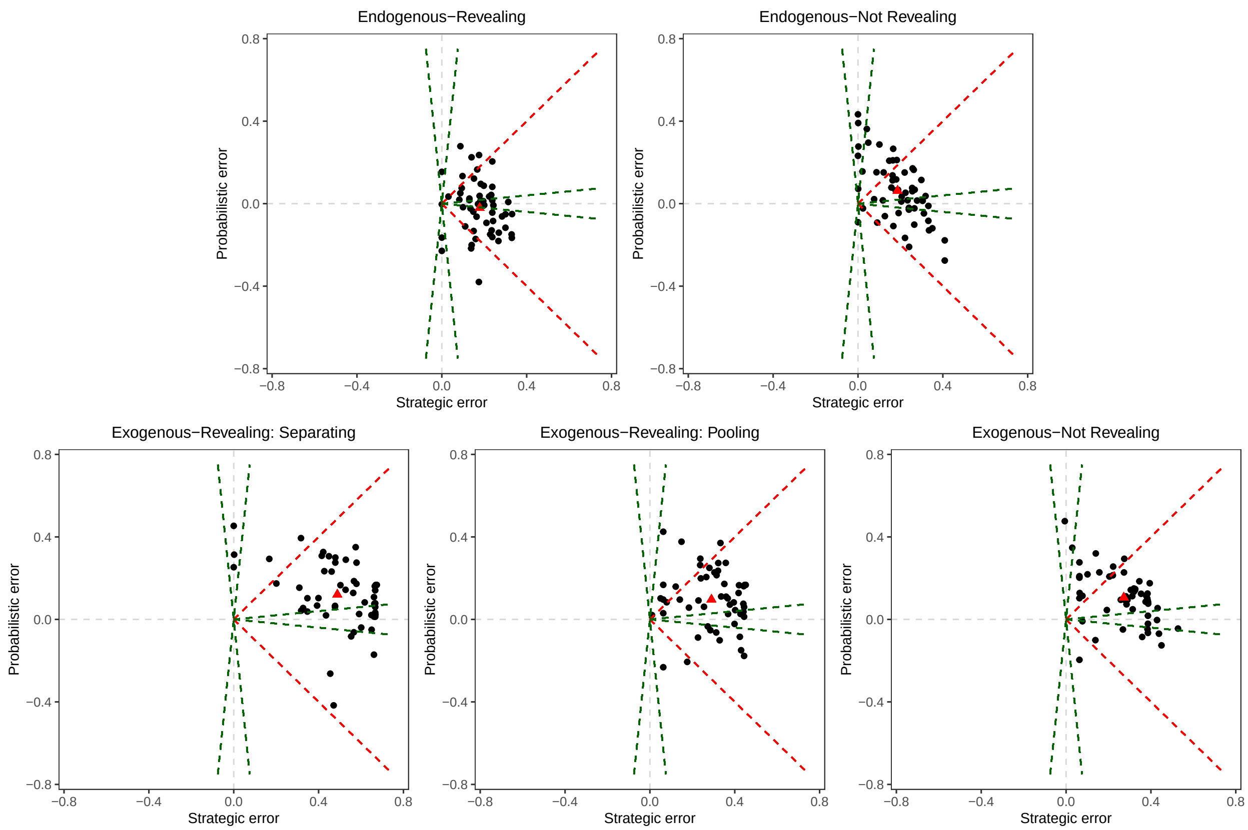


FIGURE 23. THE IMPACTS OF PROBABILISTIC AND STRATEGIC ERRORS – BASED ON JOINT ESTIMATION OF (α, β, χ)

Notes: The analysis focuses on the information set (q_L, r). The plots illustrate the impacts of probabilistic and strategic errors on posterior beliefs (as defined above) for participants across the four settings. The third and fourth plots show results conditional on the separating and pooling equilibria in the Exogenous-Revealing setting, respectively. Each data point represents the average impacts of probabilistic and strategic errors for a receiver. The red triangles indicate the setting-level averages. The red dashed lines represent $y = \pm x$. The green dashed lines represent $y = \pm \frac{x}{10}$ and $y = \pm 10x$. The impacts of probabilistic and strategic reasoning errors are comparable in magnitude for 78%, 77%, 75%/91%, and 89% of receivers in the EN-R, EN-NR, EX-R (conditional on the separating/pooling equilibrium), and EX-NR settings, respectively.

Table 9—Correlation between Strategic Naivete and Deviation form Bayesian Updating – Based on Joint Estimation of (α, β, χ)

		Correlation($\|(\alpha,\beta)-(1,1)\|$, χ)	Correlation($\lvert\alpha-1\rvert$, χ)	Correlation($\lvert\beta-1\rvert$, χ)
EN-R		0.028 (0.780) [-0.171,0.226]	0.173 (0.084) [-0.029,0.375]	-0.169 (0.090) [-0.382,0.045]
EN-NR		-0.119 (0.229) [-0.322,0.084]	0.060 (0.549) [-0.152,0.271]	-0.146 (0.140) [-0.343,0.052]
EX-R	separating	-0.004 (0.963) [-0.204,0.195]	-0.045 (0.644) [-0.230,0.140]	0.099 (0.302) [-0.095,0.293]
	pooling	-0.013 (0.890) [-0.194,0.167]	-0.071 (0.464) [-0.224,0.082]	0.158 (0.102) [-0.021,0.336]
EX-NR		0.014 (0.880) [-0.197,0.226]	0.079 (0.415) [-0.117,0.275]	-0.059 (0.537) [-0.277,0.159]

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. The table presents Kendall's τ as correlation coefficients, with the numbers in parentheses representing p-values and the numbers in brackets indicating the 95% confidence intervals of the correlation coefficients. The first column lists the four settings of our experiment. In the EX-R setting, the estimate of χ differs depending on whether it is conditional on the separating equilibrium or the pooling equilibrium. Therefore, we compute the correlation coefficients for each case separately.

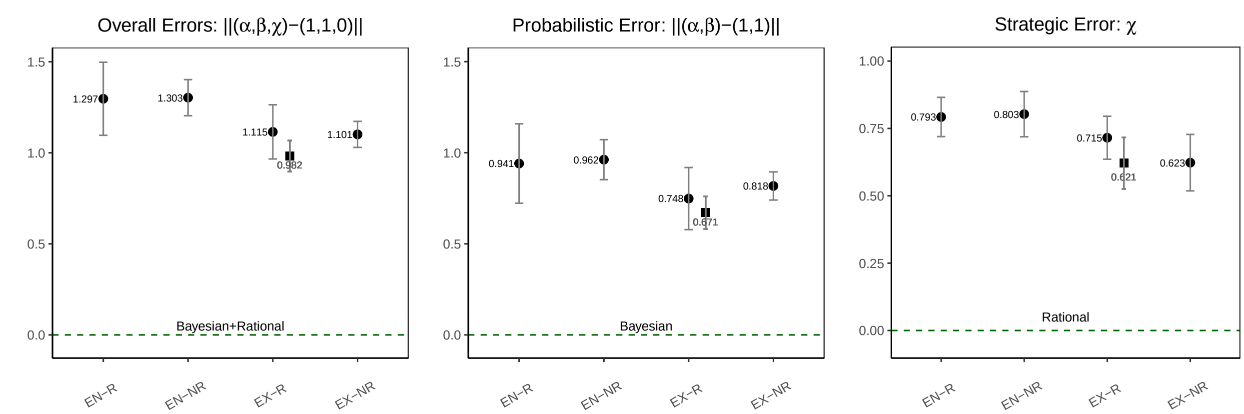


Figure 24. Error Comparison Across Settings – Based on Joint Estimation of (α, β, χ)

Notes: α, β, and χ are the behavioral model parameters defined in Section IV.A. Black dots (and rectangles) represent setting-level averages. The EX-R setting includes both separating and pooling equilibria. In the first and third plots, the black dot and rectangle corresponding to EX-R represent results based on χ estimated from the separating and pooling equilibrium benchmarks, respectively. Short vertical lines indicate 95% confidence intervals.

D. Choices of Persuaders

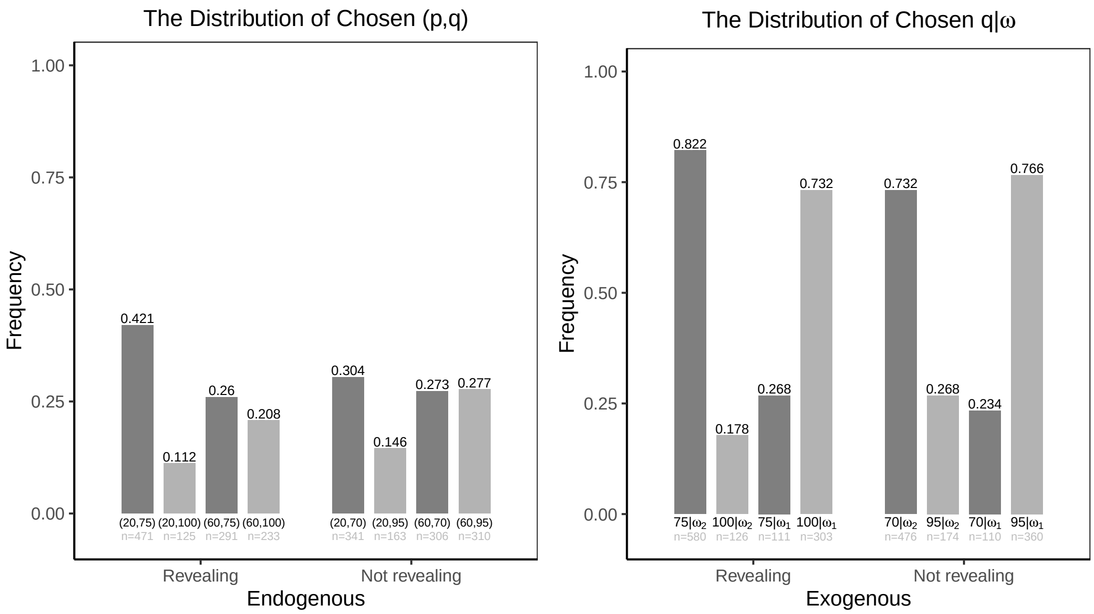


Figure 25. Choices of Persuaders

Notes: In the left panel, the labels at the bottom of each bar represent the chosen (p, q) values in percentage points, with the corresponding number of observations shown below. In the right panel, the labels represent the choice of q (in percentage points) conditional on the observed state realization ω, along with the corresponding number of observations.

The key observation from Figure 25 is that persuaders are more likely to choose q_H (q_L) together with p_H (p_L) in endogenous settings, or when ω_1 (ω_2) is observed in exogenous settings. These patterns are qualitatively consistent with the equilibrium-predicted association between the persuader's choice of information diagnosticity and their private choice or information about the underlying state.

E. The Choice-Only Experiment

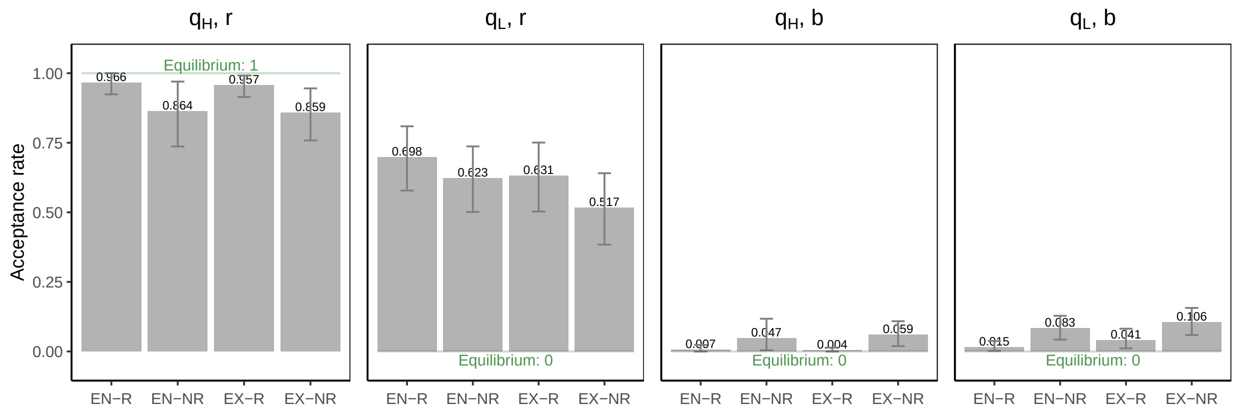


Figure 26. Action Choices of Receivers: Choice-Only Experiment

Notes: EN-R, EN-NR, EX-R, and EX-NR refer to Endogenous-Revealing, Endogenous-Not Revealing, Exogenous-Revealing, and Exogenous-Not Revealing, respectively. Short vertical lines represent 95% confidence intervals obtained via Bootstrapping. In all scenarios (i.e., information sets (q, s) that receivers can observe) except 'EN-R, (q_L, b)', there is no significant difference in the acceptance rate between the main experiment and the choice-only experiment.

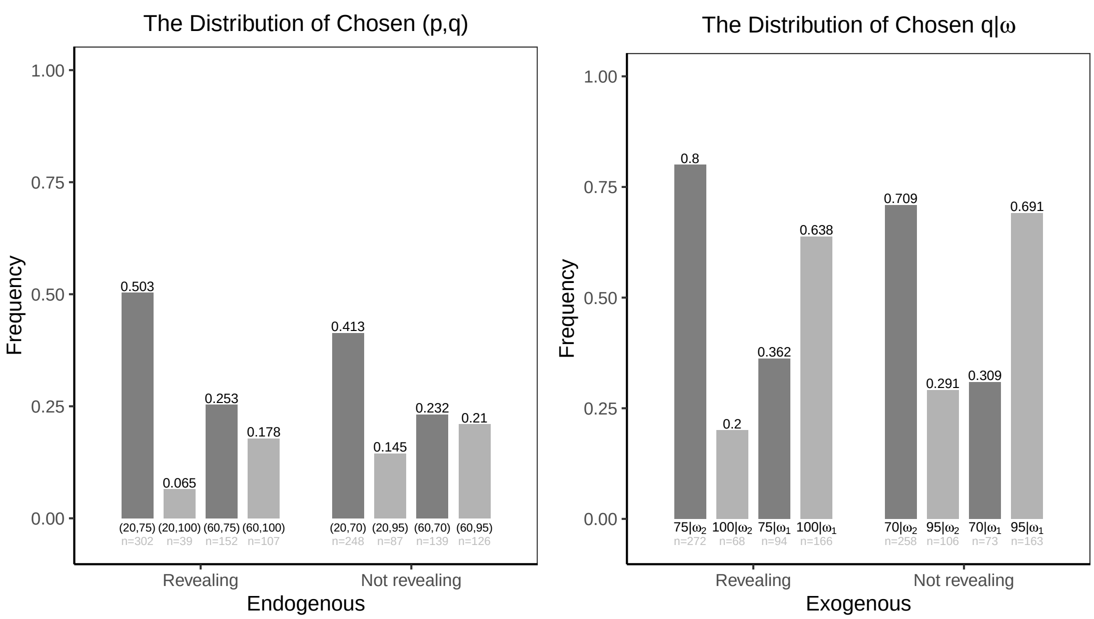


Figure 27. Choices of Persuaders: Choice-Only Experiment

Notes: In the left panel, the labels at the bottom of each bar represent the chosen (p, q) values in percentage points, with the corresponding number of observations shown below. In the right panel, the labels represent the choice of q (in percentage points) conditional on the observed state realization ω, with the corresponding number of observations. Consistent with the main experiment, the patterns of persuaders' choices qualitatively align with the predicted association between the persuader's choice of information diagnosticity and their private choice or information about the underlying state in equilibrium.

F. Simulation

In the simulation exercises, we sample 200 receivers (with replacement) from each of the four settings and retain the sequence of q values they experienced during their 20 rounds of game play. We then simulate the belief data for each sampled receiver based on pre-specified behavioral model parameters. Using the simulated belief data, we perform individual-level structural estimation (the same procedure as in Section IV). Table 10 presents the median and average parameter estimates (over 800 sampled receivers) for each pre-specified parameter combination. The estimates recover the pre-specified true parameters with substantial accuracy.

TABLE 10—SIMULATION EXERCISES

Pre-specified parameters					Estimations	
α	β	χ	σ_1	σ_2	median: $(\tilde{\alpha},\ \tilde{\beta},\ \tilde{\chi},\ \tilde{\sigma}_1,\ \tilde{\sigma}_2)$	mean: $(\bar{\alpha},\ \bar{\beta},\ \bar{\chi},\ \bar{\sigma}_1,\ \bar{\sigma}_2)$
0.31	0.77	0.87			(0.31, 0.77, 0.86, 0.15, 0.11)	(0.33, 0.78, 0.84, 0.15, 0.11)
1	1	0.87			(1.01, 1.02, 0.86, 0.14, 0.11)	(1.03, 1.03, 0.84, 0.15, 0.11)
1.5	1.5	0.87			(1.52, 1.53, 0.87, 0.15, 0.11)	(1.54, 1.54, 0.85, 0.15, 0.12)
0.31	0.77	1			(0.32, 0.78, 1, 0.15, 0.11)	(0.34, 0.79, 0.93, 0.15, 0.11)
1	1	1	0.15	0.11	(1.01, 1, 1, 0.15, 0.11)	(1.02, 1.02, 0.93, 0.15, 0.11)
1.5	1.5	1			(1.52, 1.52, 0.99, 0.15, 0.11)	(1.53, 1.53, 0.94, 0.15, 0.11)
0.31	0.77	0			(0.32, 0.79, 0, 0.15, 0.11)	(0.45, 0.88, 0.06, 0.15, 0.14)
1	1	0			(1.04, 1.03, 0, 0.14, 0.11)	(1.13, 1.14, 0.06, 0.15, 0.15)
1.5	1.5	0			(1.57, 1.57, 0, 0.15, 0.11)	(1.64, 1.65, 0.08, 0.15, 0.15)

Notes: The parameter combinations we consider in the simulation analysis are quite diverse. (0.33,0.77,0.87,0.15,0.11) are the medians of $(\alpha, \beta, \chi, \sigma_1, \sigma_2)$ from our individual-level structural estimation results. The value $\chi = 0(1)$ denotes fully rational (naive) strategic inference. The combination $(\alpha, \beta) = (1, 1)$ represents perfect Bayesian. The combination $(\alpha, \beta) = (1.5, 1.5)$ corresponds to a case of overweighting both prior beliefs and information diagnosticity.

G. Proofs

G.1. *Proof of Proposition 1*

Proof. Assume the specified parameter restrictions are imposed. We first show that in any sequential equilibrium, the receiver's best response at information set $I_1 = (q_H, r)$ and $I_4 = (q_H, b)$ must be A and R, respectively. In I_1, the receiver's interim belief lies between p_L and p_H. Accordingly, her Bayesian posterior belief falls within the range $[\mu_2(\mu_1 = p_L, I_1), \mu_2(\mu_1 =$

$p_H, I_1)]$, where $\mu_2(\mu_1 = p_L, I_1) = \frac{1}{1+\frac{1-p_L}{p_L}\frac{1-q_H}{q_H}}$ and $\mu_2(\mu_1 = p_H, I_1) = \frac{1}{1+\frac{1-p_H}{p_H}\frac{1-q_H}{q_H}}$. Given the parameter condition $\mu_1(I_1) < p_L$, it follows that $\mu_2(\mu_1 = p_L, I_1) > \mu_2(\mu_1(I_1), I_1) = \mu_2^*$. Therefore, her optimal choice in I_1 must be A. A similar logic applies to I_4. The receiver's posterior belief lies in the range $[\mu_2(\mu_1 = p_L, I_4), \mu_2(\mu_1 = p_H, I_4)]$, where $\mu_2(\mu_1 = p_L, I_4) = \frac{1}{1+\frac{1-p_L}{p_L}\frac{q_H}{1-q_H}}$ and $\mu_2(\mu_1 = p_H, I_4) = \frac{1}{1+\frac{1-p_H}{p_H}\frac{q_H}{1-q_H}}$. Given the parameter condition $\mu_1(I_4) > p_H$, we have $\mu_2(\mu_1 = p_H, I_4) < \mu_2(\mu_1(I_4), I_4) = \mu_2^*$. Hence, her optimal choice in I_4 must be R.

We next show that in any sequential equilibrium, the receiver's best response at information set $I_2 = (q_L, r)$ and $I_3 = (q_L, b)$ cannot be A and R, respectively. Suppose by contradiction that this is the case. Then her best response across all information sets, from I_1 to I_4, would be $AARR$. That is, the receiver chooses to accept when the signal realization is $s = r$, regardless of q. Given the receiver's such strategy, the persuader's expected payoff from choosing (p, q) is $p(qu_1+(1-q)u_0)+(1-p)((1-q)u_2+qu_0)$. Now consider the following parameter conditions: (1) $q_H > \frac{u_2-u_0}{u_1+u_2-2u_0}$ implies that the persuader prefers (p_H, q_H) over (p_L, q_H), (2) $q_L < \frac{u_2-u_0}{u_1+u_2-2u_0}$ implies that the persuader prefers (p_L, q_L) over (p_H, q_L), (3) $\frac{p_Hq_H-p_Lq_L}{p_H-p_L+q_H-q_L} < \frac{u_2-u_0}{u_1+u_2-2u_0}$ implies that the persuader prefers (p_L, q_L) over (p_H, q_H). Therefore, the persuader's best response must be to choose (p_L, q_L). Given this, the receiver's Bayesian posterior at I_2 is: $\mu_2(\mu_1 = p_L, I_2) = \frac{1}{1+\frac{1-p_L}{p_L}\frac{1-q_L}{q_L}}$. Since $p_L < \mu_1(I_2)$, it follows that $\mu_2(\mu_1 = p_L, I_2) < \mu_2(\mu_1(I_2)), I_2) = \mu_2^*$. This implies that the receiver would prefer to choose R, not A, in I_2, contradicting our assumption. Thus, the receiver's best responses in I_2 and I_3 cannot be A and R , respectively.

It is clear that in any sequential equilibrium, the receiver's best response at information sets I_2 and I_3 cannot both be A. The reason is as follows: if the receiver were to choose A in both I_2 and I_3, the persuader's optimal choice would be (p_L, q_L), yielding an expected payoff of $p_Lu_1 + (1 - p_L)u_2$. This payoff exceeds that from choosing (p_H, q_L), which is $p_Hu_1 + (1 - p_H)u_2$, and also exceeds the payoff from either (p_H, q_H) or (p_L, q_H), which is $p(q_Hu_1 + (1 - q_H)u_0) + (1 - p)((1 - q_H)u_2 + q_Hu_0)$. Given that $p_L < \mu_1(I_2) < \mu_1(I_3)$, the receiver would update her beliefs and switch from choosing A to R in both I_2 and I_3. Hence, the profile AA at I_2 and I_3 cannot be sustained in equilibrium.

Furthermore, the receiver's best responses at I_2 and I_3 cannot be R and A, respectively. This is because $\mu_2(\mu_1, I_3) = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{q_L}{1-q_L}} < \mu_2(\mu_1, I_2) = \frac{1}{1+\frac{1-\mu_1}{\mu_1}\frac{1-q_L}{q_L}}$. That is, the posterior belief at I_3 is strictly lower than at I_2, so it cannot be rational for the receiver to accept in I_3 while rejecting in I_2.

Taken together, we conclude that the receiver's best responses at information sets I_1

through I_4 must be A, R, R, R, respectively, in any sequential equilibrium.

We now show that the strategy profile in which the persuader chooses (p_H, q_H) and the receiver chooses A, R, R, R at information sets I_1 through I_4, respectively, constitutes a sequential equilibrium. As established earlier, A, R, R, R is the receiver's best response across $I_1 - I_4$ in any sequential equilibrium. Given this, it is also optimal for the persuader to choose (p_H, q_H) since the condition $q_H > \frac{u_2-u_0}{u_1+u_2-2u_0}$ holds, making (p_H, q_H) his unique best response. To verify that the receiver's off-equilibrium actions—choosing R, R in information sets $I_2 = (q_L, r)$ and $I_3 = (q_L, b)$—are sequentially rational, consider the following. Suppose the persuader adopts a fully mixed strategy over the four possible pairs: (p_H, q_H) , (p_L, q_H), (p_H, q_L) and (p_L, q_L), assigning them probabilities $1 - \epsilon_1 - \epsilon_2 - \epsilon_3$, ϵ_1, ϵ_2 and ϵ_3, respectively. Upon observing the off-equilibrium signal q_L, the receiver forms an interim belief: $Pr(\omega_1|q_L) = \frac{\epsilon_2}{\epsilon_2+\epsilon_3}p_H + \frac{\epsilon_3}{\epsilon_2+\epsilon_3}p_L$. As $\frac{\epsilon_3}{\epsilon_2+\epsilon_3} \to 1$, for example, by letting $\epsilon_3 = \epsilon, \epsilon_2 = \epsilon^2$, the belief $Pr(\omega_1|q_L)$ converges to $p_L(< \mu_1(I_2))$. Consequently, the receiver's posterior belief in both I_2 and I_3 falls below $\hat{\mu}$, justifying the choice of action R in both information sets. Therefore, the strategy profile in which the persuader chooses $(p = p_H, q = q_H)$ and the receiver chooses A, R, R, R at I_1 through I_4 constitutes a sequential equilibrium and indeed, the unique one.

A reordered game á la In and Wright (2018) essentially refines the sequential equilibrium. Since the sequential equilibrium is already unique, it is straightforward to verify that the refined equilibrium prescribes the persuader choosing q_H at the initial node, choosing p_H following q_H, and choosing p_L following q_L. The receiver's strategy in information sets I_1 through I_4 remains A, R, R, R, respectively. The purpose of this refinement is to discipline the receiver's interim belief upon observing the off-equilibrium action q_L: she will then hold a point belief that the probability of the state being ω_1 is p_L. Moreover, we can show that the parameter condition required for supporting this refined equilibrium is slightly weaker. Specifically, the condition $\frac{p_H q_H - p_L q_L}{p_H - p_L + q_H - q_L} < \frac{u_2-u_0}{u_1+u_2-2u_0}$ can be dropped. This is because the equilibrium in the subgame where $q = q_H$ involves the persuader choosing p_H and the receiver choosing A and R in I_1 and I_4, respectively, while the subgame where $q = q_L$ involves the persuader choosing p_L and the receiver choosing R in both I_2 and I_3. These subgame equilibria are sufficient to ensure that the persuader prefers q_H at the initial node, without requiring the more stringent parameter condition. □

G.2. *Proof of Proposition 2*

Proof. We analyze the Not Revealing setting ($q_H < 1$) and the Revealing setting ($q_H = 1$) separately.

Case 1: $q_H < 1$.

We first show that no separating equilibrium exists, that is, there is no equilibrium in which the persuader's strategy is $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$ or (q_L, q_H).[63] In both configurations, the receiver's best response would involve accepting in the information sets where the persuader chooses the corresponding q and rejecting where another q is chosen. However, in such a case, the persuader in state ω_2 would find it profitable to mimic the strategy of the persuader in state ω_1, thereby undermining the separation. As a result, only pooling equilibria may exist.

Next, we show that in any sequential equilibrium, the receiver's best response: in I_1 and I_4 cannot be (a) RA or (b) AA, and in I_2 and I_3 cannot be (c) RA or (d) AA. Cases (a) and (c) are ruled out because, for any interim belief after observing q, the posterior belief after additionally observing $s = r$ is at least as high as after observing $s = b$. Cases (b) and (d) are ruled out because we are in a pooling equilibrium, and the condition $\mu_1(I_1) < \mu_1(I_2) < p^* < \mu_1(I_3) < \mu_1(I_4)$ implies that accepting in both I_1 and I_4 (or I_2 and I_3) is not sequentially rational.

Thus, the receiver's best response in I_1 through I_4 in any sequential equilibrium must be one of the following: (a) $AARR$, (b) $RRRR$, (c) $RARR$, and (d) $ARRR$. Option (a) cannot be an equilibrium because it induces the persuader to choose $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$, contradicting the optimality of the receiver's choice in I_2. Option (b) also fails, because if the persuader is pooling on either q_H or q_L, then the receiver should accept after observing $s = r$ in the pooling q due to the condition $\mu_1(I_1) < \mu_1(I_2) < p^* < \mu_1(I_3) < \mu_1(I_4)$.

(c) can be part of a sequential equilibrium strategy. To see this, consider the case where $(\pi(\omega_1), \pi(\omega_2)) = (q_L, q_L)$. In this case, the receivers choice of AR in I_2 and I_3 is sequentially rational, and it is optimal for the persuader to choose (q_L, q_L). To establish that the receiver's choice of RR in I_1 and I_4 is also sequentially rational, suppose that the persuader plays a slightly mixed strategy, with $Pr(\pi(\omega_1) = q_H) = \epsilon_1$ and $Pr(\pi(\omega_2) = q_H) = \epsilon_2$. Then, after observing q_H, the receiver's interim belief about the state being ω_1 is given by $\frac{p^*\epsilon_1}{p^*\epsilon_1+(1-p^*)\epsilon_2}$. As $\frac{\epsilon_1}{\epsilon_2} \to 0$, for instance by letting $\epsilon_1 = \epsilon^2$ and $\epsilon_2 = \epsilon$, the receiver's interim belief after observing

[63] Since we focus on pure strategy, $\pi(\omega)$ denotes the persuader's choice between q_H and q_L in state ω with slight abuse of notation.

q_H converges to zero. Consequently, her Bayesian posterior belief in I_1 also converges to zero, which is less than $\hat{\mu}$. Hence, choosing RR in I_1 and I_4 is sequentially rational.

However, the divinity refinement proposed by Banks and Sobel (1987) rules out the proposed equilibrium. Let D_1 denote the set of the receiver's best responses off the equilibrium path that induce the persuader in state ω_1 to strictly prefer deviating from the proposed equilibrium. Define D_2 analogously for the persuader in state ω_2. Under the proposed equilibrium, it is clear that $D_1 = \{AA, AR\}$ and $D_2 = \{AA\}$, so $D_2 \subset D_1$. In other words, whenever the persuader in state ω_2 is tempted to deviate, the persuader in state ω_1 is even more so. The divinity criterion requires that the receiver's interim belief about being in state ω_1 after the off-equilibrium signal $q = q_H$, denoted α_{q_H}, cannot increase the likelihood ratio of type ω_2 relative to type ω_1. Formally, this means: $\frac{1-\alpha_{q_H}}{\alpha_{q_H}} \leq \frac{1-p^*}{p^*}$, which implies $\alpha_{q_H} > p^*$. Given this belief, the receiver's optimal action in information set I_1 is to accept, since $\alpha_{q_H} > p^* > \mu_1(I_1)$, where the second inequality holds by assumption. Therefore, at least the persuader in state ω_1 has an incentive to deviate, and the proposed equilibrium cannot be sustained.

(d) can be part of a sequential equilibrium strategy. To see this, consider the persuader's strategy $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_H)$. The receiver's actions AR in I_1 and I_4 are sequentially rational, and it is optimal for the persuader to choose (q_H, q_H). To see why the receiver's actions RR in I_2 and I_3 are also sequentially rational, suppose $Pr(\pi(\omega_1) = q_L) = \epsilon_1$ and $Pr(\pi(\omega_2) = q_L) = \epsilon_2$. Then the receiver's interim belief after observing q_L in I_2 is $\frac{p^*\epsilon_1}{p^*\epsilon_1+(1-p^*)\epsilon_2}$. As $\frac{\epsilon_1}{\epsilon_2} \to 0$, for example with $\epsilon_1 = \epsilon^2$ and $\epsilon_2 = \epsilon$, this belief converges to zero. Hence, the receiver's posterior in I_2 is less than $\hat{\mu}$, making rejection in I_2 and I_3 sequentially rational.

This proposed equilibrium also satisfies the divinity refinement. Specifically, define $D_1 = \{AA\}$ as the set of receiver responses that would make the persuader in state ω_1 strictly prefer deviating, and $D_2 = \{AA, AR\}$ as the corresponding set for state ω_2. Since $D_1 \subset D_2$, whenever the persuader in state ω_1 is tempted to deviate, the persuader in state ω_2 is even more so. According to the divinity criterion, the receiver's belief α_{q_L} after observing the off-equilibrium message q_L must not increase the likelihood ratio in favor of type ω_1. Formally, $\frac{\alpha_{q_L}}{1-\alpha_{q_L}} \leq \frac{p^*}{1-p^*}$, which implies that $\alpha_{q_L} \leq p^*$. As long as $\alpha_{q_L} < \mu_1(I_2) < p^*$, where the assumed parameter condition ensures the second inequality, the equilibrium remains intact.

Taken together, there exists a unique sequential equilibrium consistent with the divinity refinement in this case: the persuader chooses $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_H)$, and the receiver accepts only in information set I_1.

Case 2: $q_H = 1$.

In this case, perfectly revealing the underlying state is feasible, and sequential rationality requires the receiver to choose AR in information sets $I_1 = (q_H = 1, r)$ and $I_4 = (q_H = 1, b)$, respectively. We now show that, in any sequential equilibrium, the receiver's best response in information set $I_2 = (q_L, r)$ and $I_3 = (q_L, b)$ cannot be: (a) RA, (b) AA, or (c) AR. Case (a): This cannot occur because for any interim belief μ_1 after observing q_L, the posterior belief in I_2 satisfies $\mu_2(\mu_1, I_2) \geq \mu_2(\mu_1, I_3)$, which rules out RA. Case (b): Suppose the receiver chooses AA. Then the persuader's best response would be $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$ or (q_L, q_L). In the former, the receiver's strategy AA is not optimal because the state is revealed and rejection is warranted in both information sets. In the latter, consider the receiver's Bayesian posterior belief in $I_3 = (q_L, b)$: $\mu_2(\mu_1 = p^*, I_3) = \frac{1}{1+\frac{1-p^*}{p^*}\frac{q_L}{1-q_L}} < \hat{\mu}$ by the assumed parameter condition $p^* < \mu_1(I_3)$. Thus, accepting in I_3 is not optimal. Case (c): If the receiver plays AR in I_2 and I_3, while also choosing AR in I_1 and I_4, then the persuader's best response is $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$. But in this case, the posterior in I_2 does not justify acceptance. Therefore, in any sequential equilibrium, the receiver's best response across I_1 to I_4 must be $ARRR$.

Given that the receiver chooses $ARRR$ in information sets I_1 through I_4, the persuader optimally chooses $\pi(\omega_1) = q_H$. Note that the persuader in state ω_2 receives an expected payoff of u_0 regardless of whether they choose $\pi(\omega_2) = q_H$ or q_L. In the separating case where $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$, the receiver's choosing $ARRR$ in $I_1 - I_4$ is clearly sequential rational. In the pooling case where $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_H) = (1, 1)$, the receiver's choice of AR in I_1 and I_4 is also clearly sequential rational. To verify that the receiver's choice of RR in I_2 and I_3 is sequential rational in the pooling case, suppose the persuader in state ω_1 chooses q_H with probability $1 - \epsilon_1$ and q_L with probability ϵ_1, while the persuader in state ω_2 chooses q_H with probability $1 - \epsilon_2$ and q_L with probability ϵ_2. Then, upon observing q_L, the receiver's interim belief is $Pr(\omega_1|q_L) = \frac{p^*\epsilon_1}{p^*\epsilon_1+(1-p^*)\epsilon_2}$. As $\frac{\epsilon_1}{\epsilon_2} \to 0$, for example with $\epsilon_1 = \epsilon^2$ and $\epsilon_2 = \epsilon$, the receiver's interim belief converges to zero. Since this belief is less than $\hat{\mu}$, choosing RR in I_2 and I_3 is optimal.

Therefor, when $q_H = 1$, there are two possible equilibria: A pooling equilibrium, where the persuader chooses $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_H)$ and the receiver chooses $ARRR$ in I_1 through I_4, and a separating equilibrium, where the persuader chooses $(\pi(\omega_1), \pi(\omega_2)) = (q_H, q_L)$ and the receiver chooses $ARRR$ in I_1 through I_4.

Finally, we show that the pooling equilibrium survives the refinement of divinity. To see this, note that $D_1 = \emptyset$ and $D_2 = \{AA, AR\}$. Thus, $D_1 \subset D_2$. In other words, whenever the persuader in state ω_1 finds deviation strictly beneficial, the persuader in state ω_2 does as well. The divinity refinement requires that, following an off-equilibrium message $q = q_L$, the receiver's updated belief about the state being ω_1, denoted α_{q_L}, must not increase the likelihood ratio of type ω_1 relative to type ω_2. That is, $\frac{\alpha_{q_L}}{1-\alpha_{q_L}} \leq \frac{p^*}{1-p^*}$. This condition implies that $\alpha_{q_L} \leq p^*$. As long as $\alpha_{q_L} < \mu_1(I_2) < p^*$, where the assumed parameter condition ensures the second inequality, the equilibrium remains intact. $\square$

H. Instructions and Screenshots

H.1. *Experimental Instructions*

The experiment was conducted in Chinese. Below, we present an English translation of the instructions used in the Endogenous-Revealing setting, followed by the variations for the other three settings.

Participants received printed copies of the experimental instructions and a summary of the game play, listed in the order of decision timing. We played pre-recorded audio of the experimental instructions to ensure consistency across sessions within the same setting. Then, we walked through the decision timing twice and addressed participants' questions about the instructions publicly.

Experimental Instructions for the Endogenous-Revealing Setting

Welcome to today's decision-making experiment! By carefully reading these instructions and making thoughtful decisions, you have the opportunity to earn a substantial reward.

Half of today's participants will be randomly assigned to the role of Player A, while the other half will be assigned to the role of Player B. Your role will remain fixed throughout the experiment. You will complete 20 rounds of a two-player game with monetary rewards. At the start of each round, one Player A and one Player B will be randomly matched to play that round's game.

Player A's Decisions: Two Choices

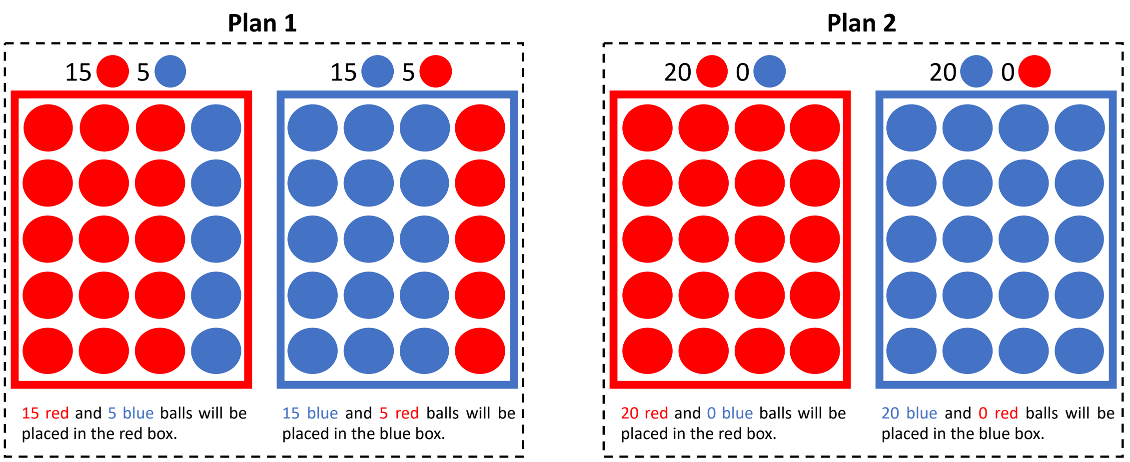
Player A will make the first decision by choosing either the first or second pack of boxes shown below.

- **First Pack**: Contains 3 red and 2 blue boxes, with the chances of selecting a red or blue box being 60% and 40%, respectively.

- **Second Pack**: Contains 1 red and 4 blue boxes, with the chances of selecting a red or blue box being 20% and 80%, respectively.

Player A will also choose one of two plans for placing balls into the red and blue boxes.

- **Plan 1**: 15 red and 5 blue balls are placed in the red box, and 15 blue and 5 red balls are placed in the blue box. Under this plan, the chance of randomly drawing a red or blue ball from the red box is 75% and 25%, respectively, while from the blue box, the chances are 25% and 75%.
- **Plan 2**: 20 red and 0 blue balls are placed in the red box, and 20 blue and 0 red balls are placed in the blue box. Under this plan, the chance of drawing a red or blue ball from the red box is 100% and 0%, respectively, while from the blue box, the chances are 0% and 100%.



The Selected Box

The computer will randomly select one box from the pack chosen by Player A for this round. It will then place the corresponding number of red and blue balls into the selected box based on Player A's chosen plan. The color of the selected box will only be revealed during the feedback phase at the end of the round.

Player B's Guess 1 and Guess 2

After Player A has made their choices, Player B will make decisions. First, Player B will estimate the chances of the selected box being red or blue under two scenarios. These guesses can range from 0% to 100%.

- **Guess 1**: Without knowing which pack of boxes or ball placement plan Player A selected, Player B estimates the chances of the selected box being red or blue.

- **Guess 2**: Without knowing which pack of boxes was selected, but knowing the ball placement plan, Player B estimates the chances of the selected box being red or blue.

The Drawn Ball

The computer will randomly draw one ball from the selected box. The color of the ball will only be revealed during the feedback phase at the end of the round.

Player B's Guess 3, Guess 4, and Choice to Trade

- **Guess 3 and Choice 1**: If the computer draws a red ball and Player B knows the ball placement plan (but not the pack of boxes selected), Player B will estimate the chances of the selected box being red or blue and decide whether to trade with Player A. This applies only if the ball drawn is red.

- **Guess 4 and Choice 2**: If the computer draws a blue ball and Player B knows the ball placement plan (but not the pack of boxes selected), Player B will estimate the chances of the picked box being red or blue and decide whether to trade with Player A. This applies only if the ball drawn is blue.

Payoff Rules for Choices

Player payoffs depend on the selected box and Player B's trade decision, as shown in Table 1:

Table 1: Payoff for Choices (unit: ECU)

Player B's Trade Decision	The Picked Box	Player A's Payoff	Player B's Payoff
No	Red or Blue Box	10	10
Yes	Red Box	12	15
Yes	Blue Box	21	6

Payoff Rules for Guesses

Player B's payoffs for guesses are calculated as follows: Suppose Player B reports the chance of the selected box being red as X%. The computer randomly selects a number Y% between 0% and 100%. The payoff is determined by the relationship between X% and Y%, as shown in Table 2.

Table 2: Player B's Payoff for Guess (unit: ECU)

Computer's Random Number	**Relation**	**Player B's Reported Probability**	**Payoff=15**	**Payoff=6**
Y%	≥	X%	Y%	1- Y%
Y%	<	X%	When the box is red	When the box is blue

According to the payoff rules for guesses outlined above, the most advantageous choice for Player B in each guess is to truthfully report the chance they estimate for the selected box being red or blue. Below are some examples to illustrate this.

Example 1: Suppose Player B estimates the chance of the selected box being red as 50%. The optimal choice for Player B is to report X% = 50%. Here's why:

If Player B reports a lower chance than their true estimate, for example, X% = 30%, the situation would unfold as follows: If the computer selects a random number Y% between 30% and 50%, Player B will have a Y% chance of earning a payoff of 15. However, if Player B truthfully reports X% = 50%, the chance of earning 15 increases to 50%, which is greater than Y%.

Similarly, if Player B reports a higher chance than their true estimate, for example, X% = 70%, the situation would unfold as follows: If the computer selects a random number Y% between 50% and 70%, Player B would have a 50% chance of earning a payoff of 15. However, if Player B truthfully reports X% = 50%, the chance of earning 15 would be Y%, which is greater than 50%.

Example 2: Suppose Player B estimates the chance of the selected box being red as 100%. In this case, the optimal choice for Player B is to report X% = 100%. Here's why:

If Player B reports a lower chance than their true estimate, for example, X% = 80%, and the computer selects a random number Y% between 80% and 100%, Player B will earn a payoff of 15 with a chance of Y%. On the other hand, if Player B truthfully reports X% = 100%, the chance of earning 15 would be 100%, which is higher than Y%.

Example 3: Suppose Player B estimates the chance of the selected box being red as 0%. In this case, the optimal choice for Player B is to report X% = 0%. Here's why:

If Player B reports a higher chance than their true estimate, for example, X% = 20%, and the computer selects a random number Y% between 0% and 20%, Player B will earn a payoff of 15 with a chance of 0%. However, if Player B truthfully reports X% = 0%, the chance of earning 15 would be Y%, which is higher than 0%.

In summary, even if you are not fully clear about the payoff calculation rules for guesses, then you only need to remember this: According to these rules, **it is in the best interest of player B to truthfully report their guess about the chance of the selected box being red or blue.**

Feedback and Payoff Calculation for Each Round

At the end of each round, the following information will be displayed to both Player A and Player B:

- The pack of boxes and the ball placement pan that Player A chose.
- Player B's Choices 1 and 2.
- The selected box for this round.
- The color of the ball randomly drawn from the selected box.

Player A's payoff for the round is based on the outcome of their choices. For Player B, the computer will randomly select one of the following four decision-based payoffs: the payoff from the active Choice 1 or Choice 2, the payoff from Guess 1, the payoff from Guess 2, or the payoff from the active Guess 3 or Guess 4.

Final Payoff Rules

One of the 20 rounds will be randomly selected to determine each participant's final payoff. The experimental currency will be converted to RMB at the rate of 1 ECU = 3.5 RMB.

After completing the 20 rounds of the game, each participant will answer two sets of test questions and complete one questionnaire. The final reward for each participant will include: a 10 RMB show-up fee, earnings from the two-player game, and earnings from the two sets of test questions.

Practice Rounds and Experiment Guidelines

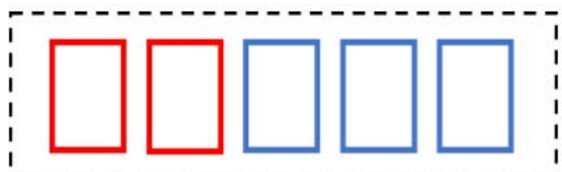
We will begin by reviewing the decision-making sequence and key points of the game. You will then have the opportunity to ask any questions you may have about the experiment instructions. Once all questions are answered, the experiment will proceed to two practice rounds of the two-player game. During these practice rounds, you will take on both the Player A and Player B roles. The purpose of the practice rounds is to help you become familiar with the two-player game. Your decisions during the practice rounds will NOT affect your earnings in the subsequent 20 rounds.

During the experiment, you must keep your cell phone turned off and refrain from speaking with other participants. If you have any questions during the experiment, please raise your hand, and a staff member will assist you.

For the Exogenous-Revealing Setting*, the instruction regarding the selected box and Player A's decisions is replaced with the following. Additionally, any references to the choice of the pack of boxes are removed from the entire instruction where applicable.*

The selected Box

At the start of each round, the computer will randomly and equally likely select one box from a pack consisting of 2 red and 3 blue boxes to use for that round.



Player A's Decision

After observing the color of the selected box, Player A chooses one of two plans for placing balls into the red and blue boxes: ......[64]

For the Endogenous-Not Revealing Setting and the Exogenous-Not Revealing Setting*, the experimental instructions resemble those for the Endogenous-Revealing Setting and the Exogenous-Revealing Setting, respectively, with the 15-5 and 20-0 plan replaced by the 14-6 and 19-1 plans.*

[64]The omitted part is the same as the corresponding part for the Endogenous-Revealing setting.

H.2. *Screenshots of Experimental Interface*

The following figures illustrate experimental interfaces of the persuader's and receiver's tasks in the order of decision timing. These images are an English translation of the actual experimental interface that was used in the Endogenous-Revealing and Exogenous-Not Revealing settings. The original interface, which was in Chinese, is available upon request.

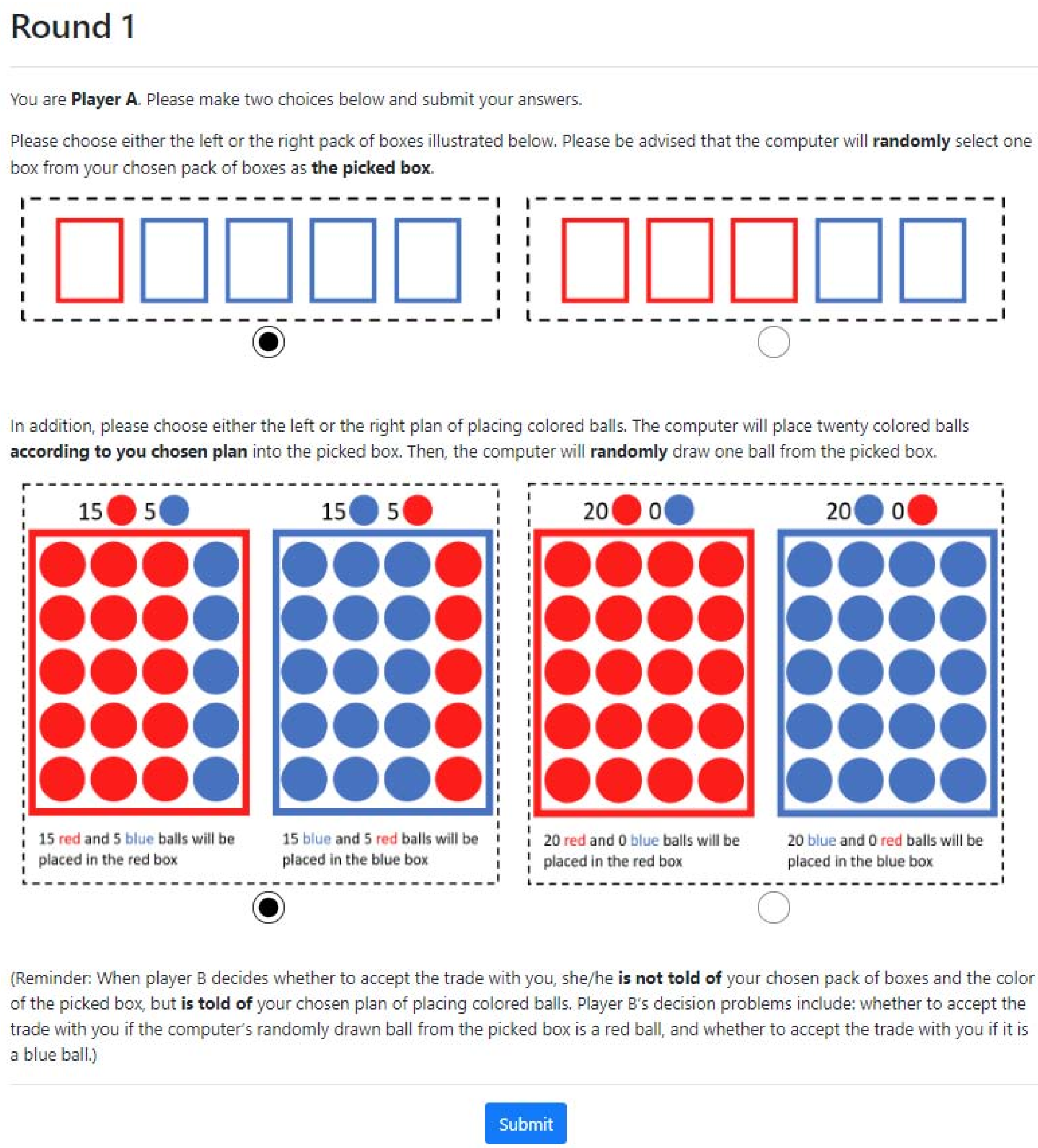


Figure 28. The Persuader's Tasks in the Endogenous-Revealing Setting

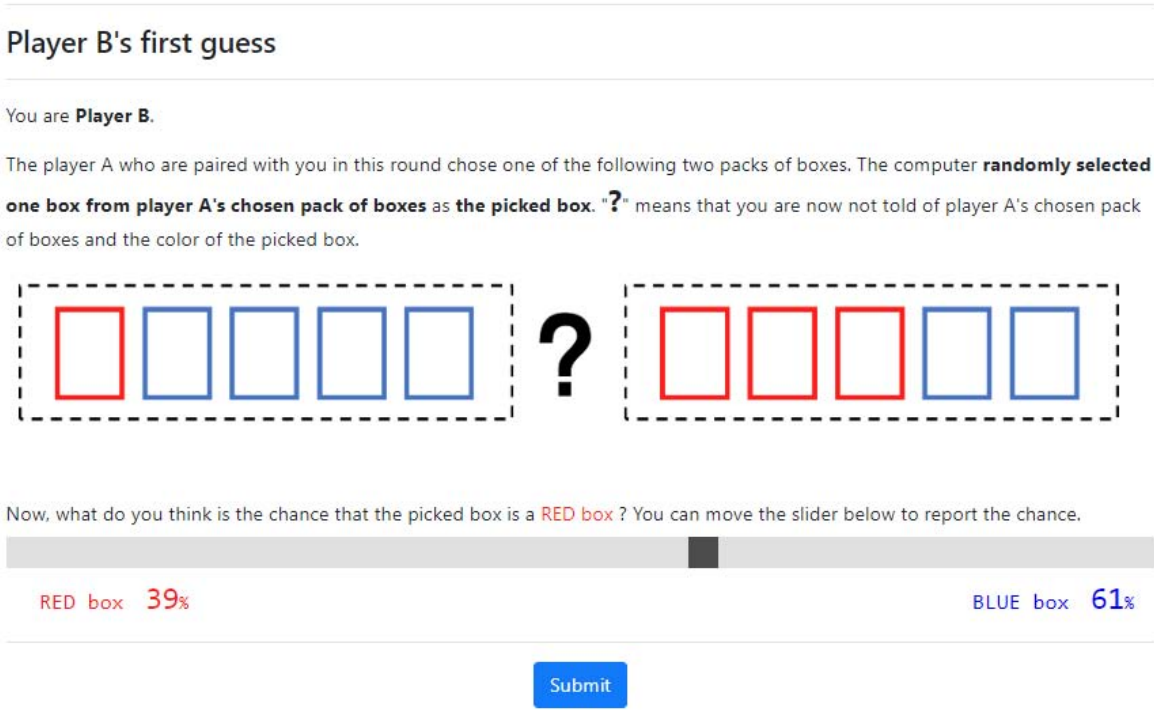


Figure 29. The Receiver's Initial Belief Elicitation in the Endogenous-Revealing Setting

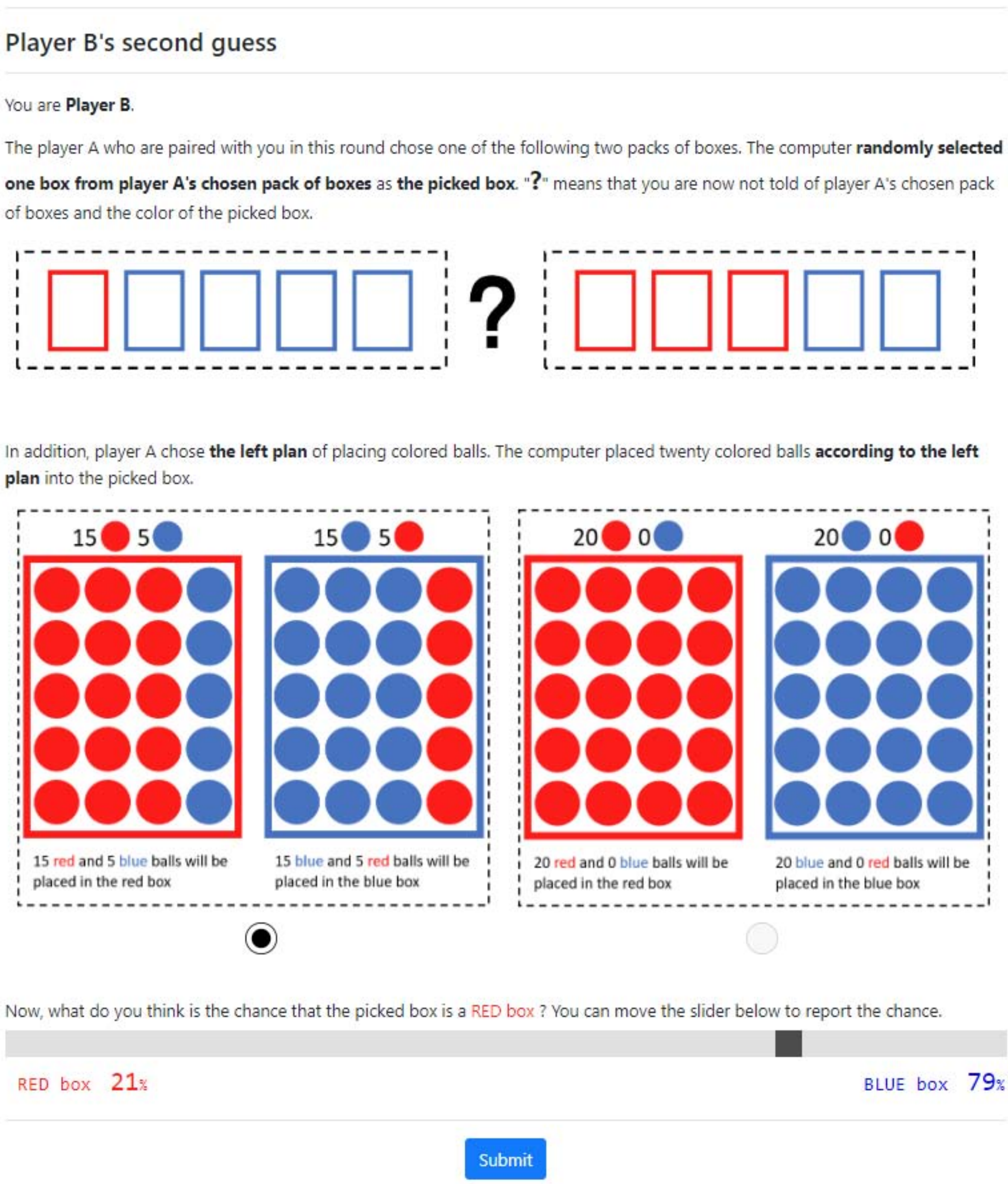


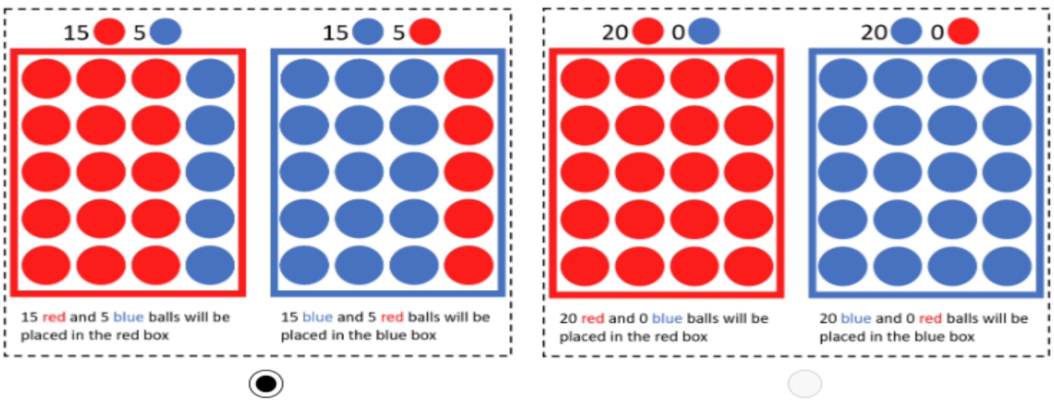
Figure 30. The Receiver's Interim Belief Elicitation in the Endogenous-Revealing Setting

Player B's third and fourth guesses, and decisions of whether to accept the trade with player A

You are **Player B**.

The player A who are paired with you in this round chose one of the following two packs of boxes. The computer **randomly selected one box from player A's chosen pack of boxes** as **the picked box**. "**?**" means that you are now not told of player A's chosen pack of boxes and the color of the picked box.

In addition, player A chose **the left plan** of placing colored balls. The computer placed twenty colored balls **according to the left plan** into the picked box.



According to player A's chosen plan of placing colored balls that is illustrated above, if the picked box is a RED box, it contains **15 red balls and 5 blue balls** ; if the picked box is a BLUE box, it contains **15 blue balls and 5 red balls** . The computer randomly selected one ball from the picked box.

--Suppose the randomly drawn ball from the picked box is a **red ball**,

- What do you think is the chance that the picked box is a RED box ? You can move the slider below to report the chance.

  RED box 84% BLUE box 16%

- Do you accept the trade with player A? ◉ Yes ○ No

--Suppose the randomly drawn ball from the picked box is a **blue ball**,

- What do you think is the chance that the picked box is a RED box ? You can move the slider below to report the chance.

  RED box 25% BLUE box 75%

- Do you accept the trade with player A? ○ Yes ◉ No

You need to answer all four questions before proceeding to the next page.

Submit

(Please click on the button of "Submit" to go to the next page)

Figure 31. The Receiver's Posterior Belief Elicitation and Choice Task in the Endogenous-Revealing Setting

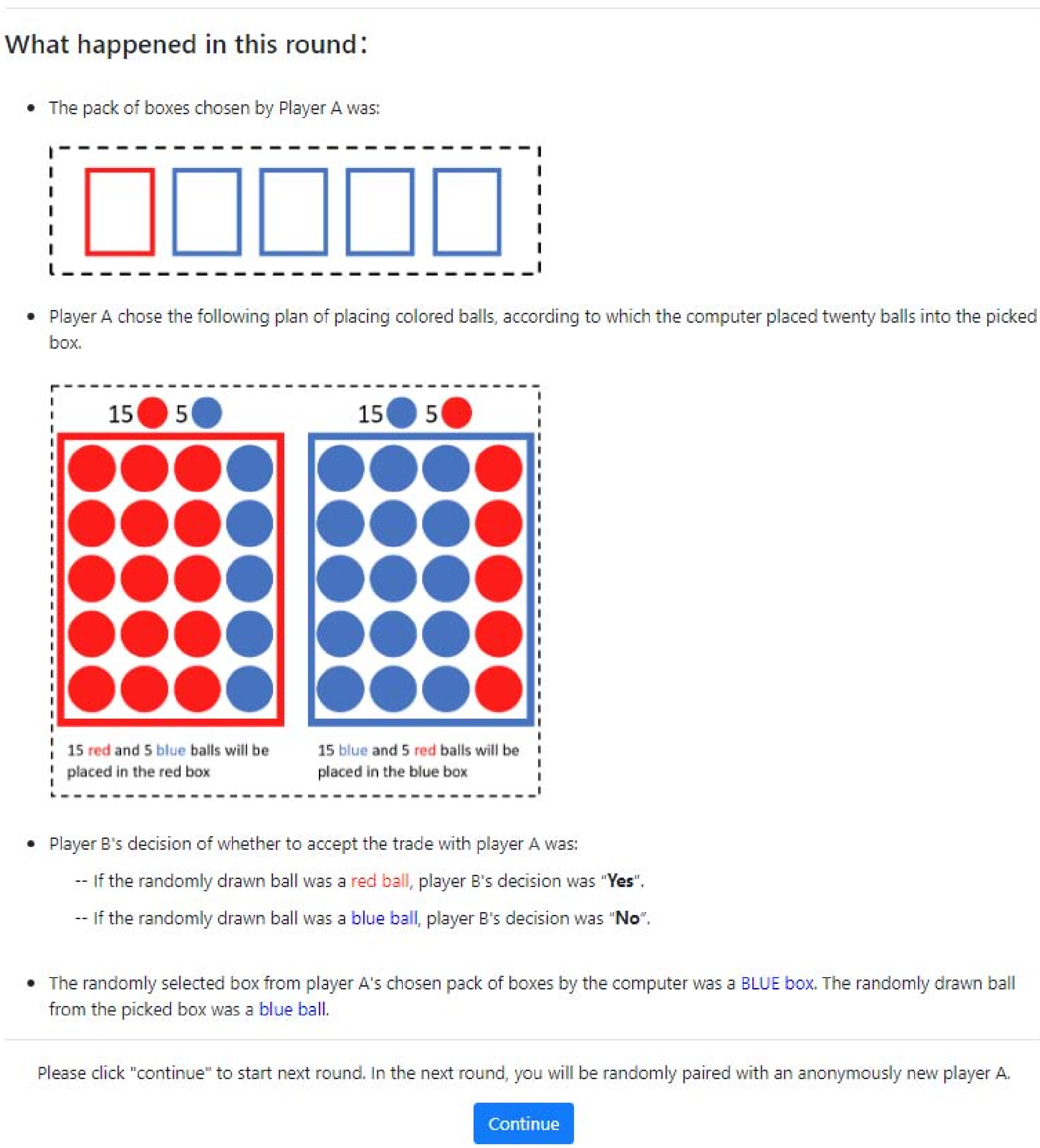
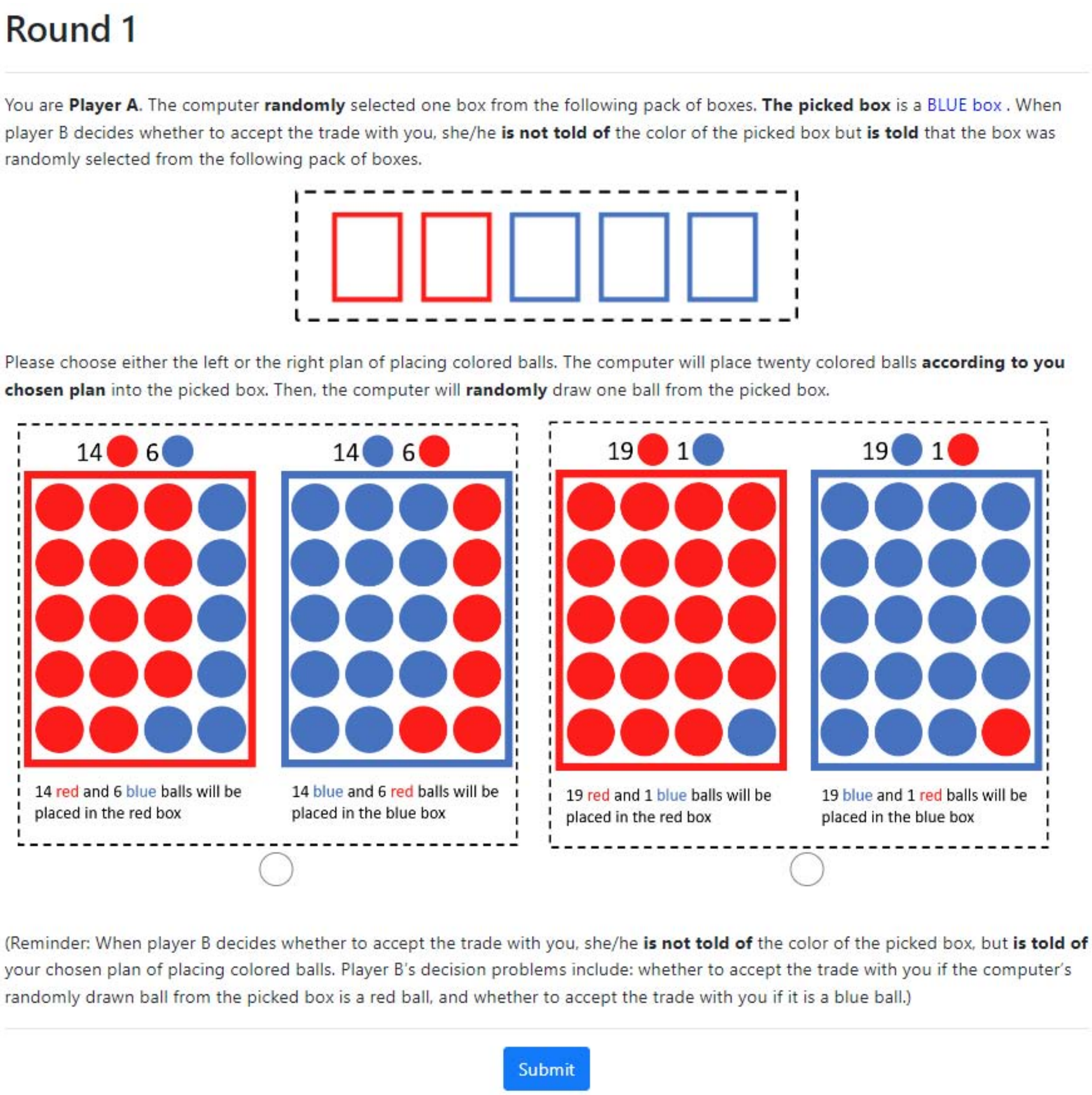


FIGURE 32. FEEDBACK IN THE ENDOGENOUS-REVEALING SETTING *Notes: The feedback page was shown to both persuader and receiver participants.*


Round 1

You are **Player A**. The computer **randomly** selected one box from the following pack of boxes. **The picked box** is a BLUE box . When player B decides whether to accept the trade with you, she/he **is not told of** the color of the picked box but **is told** that the box was randomly selected from the following pack of boxes.

Please choose either the left or the right plan of placing colored balls. The computer will place twenty colored balls **according to you chosen plan** into the picked box. Then, the computer will **randomly** draw one ball from the picked box.

(Reminder: When player B decides whether to accept the trade with you, she/he **is not told of** the color of the picked box, but **is told of** your chosen plan of placing colored balls. Player B's decision problems include: whether to accept the trade with you if the computer's randomly drawn ball from the picked box is a red ball, and whether to accept the trade with you if it is a blue ball.)

Submit

Figure 33. The Persuader's Tasks in the Exogenous-Not Revealing Setting

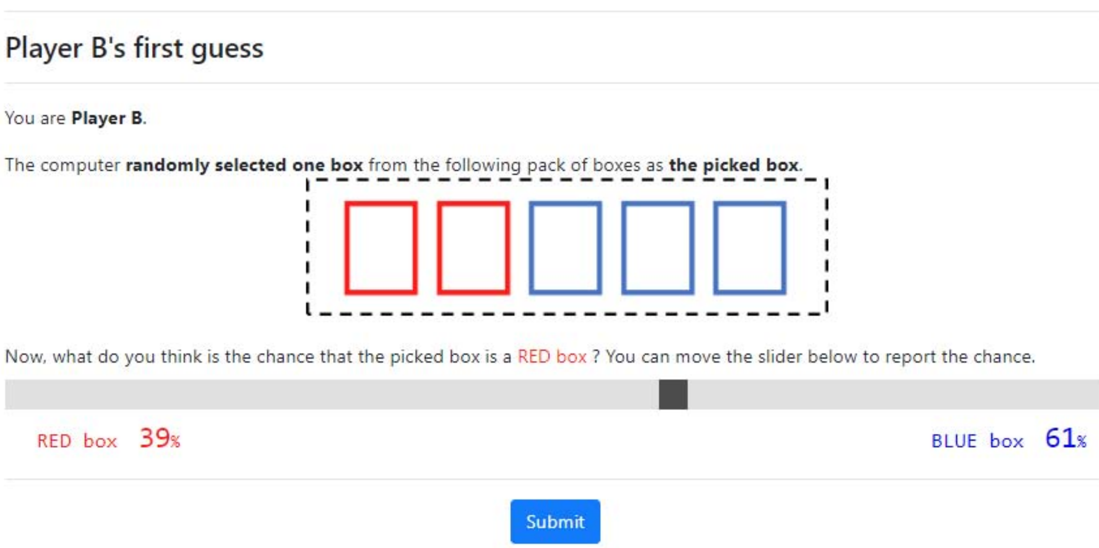


Figure 34. The Receiver's Initial Belief Elicitation in the Exogenous-Not Revealing Setting

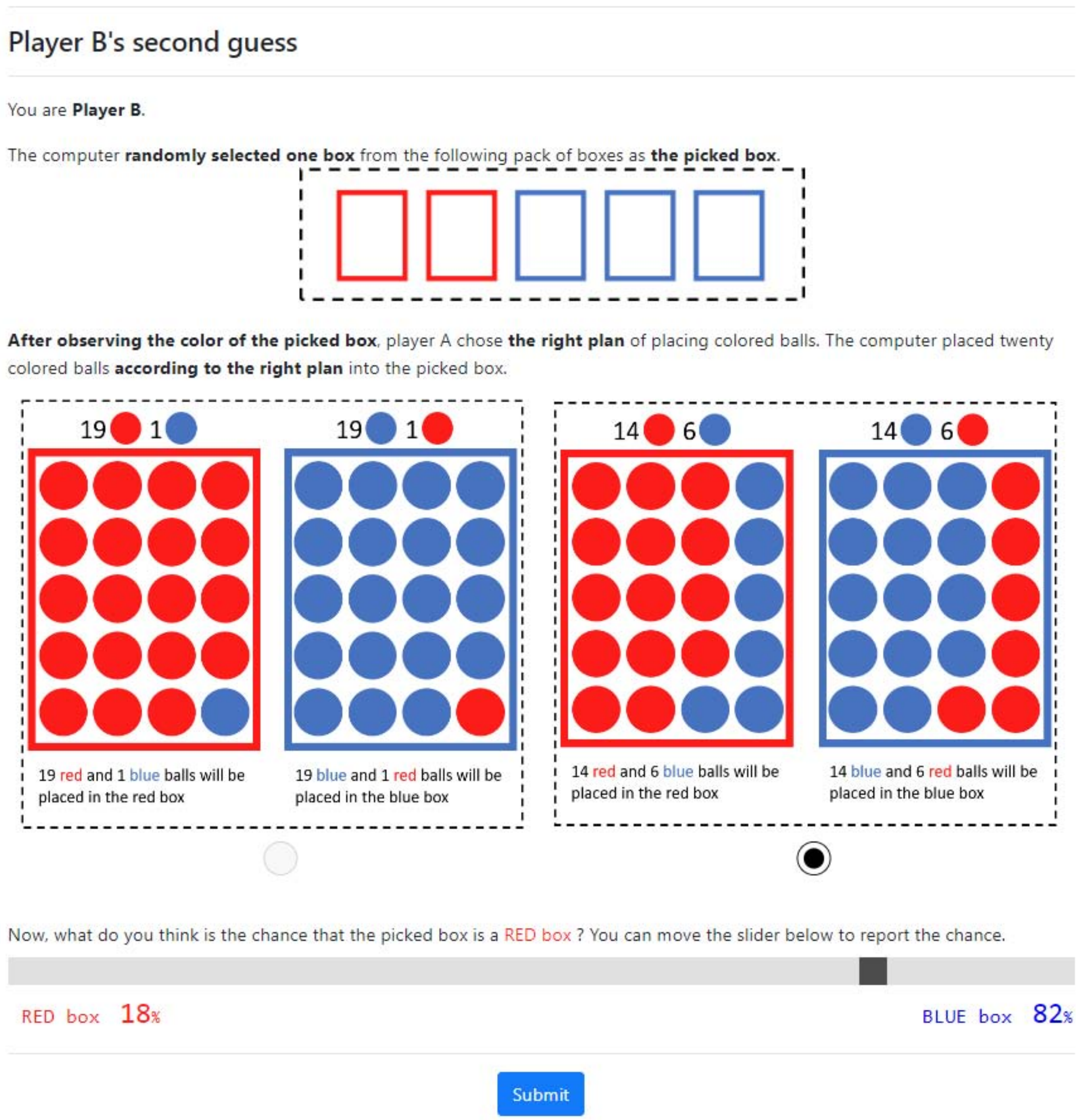


Figure 35. The Receiver's Interim Belief Elicitation in the Exogenous-Not Revealing Setting

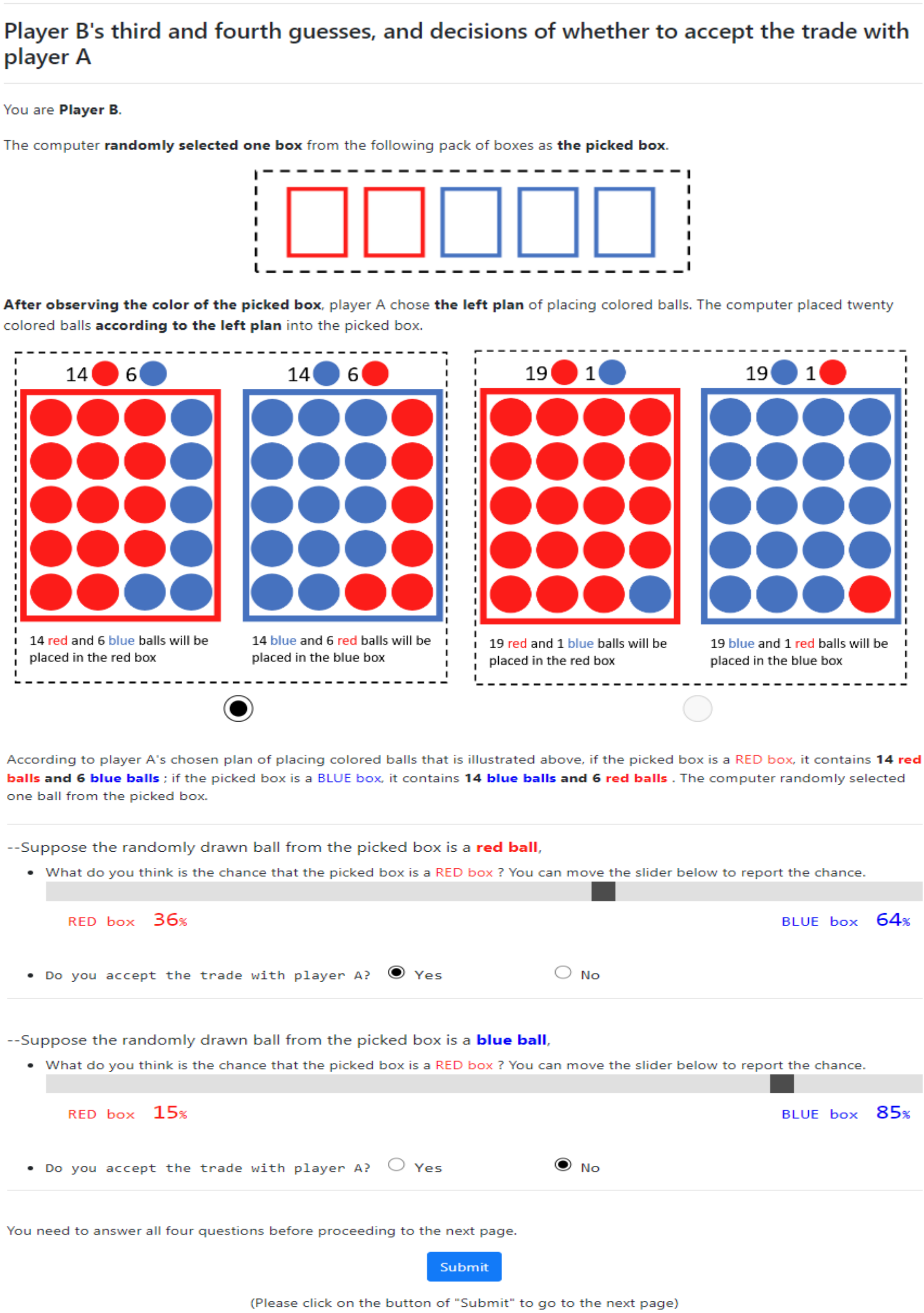


Figure 36. The Receiver's Posterior Belief Elicitation and Choice Task in the Exogenous-Not Revealing Setting

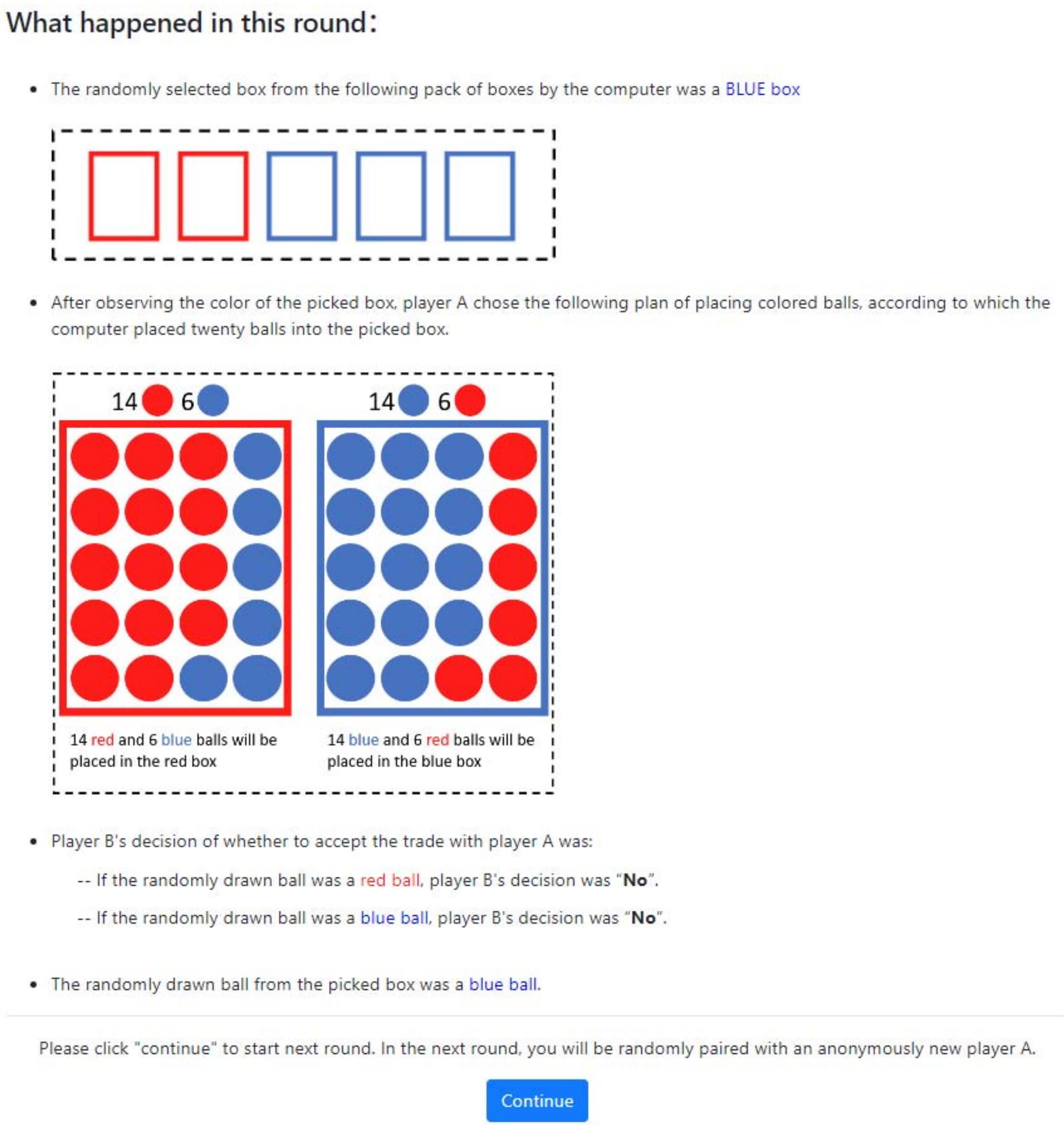


FIGURE 37. FEEDBACK IN THE EXOGENOUS-NOT REVEALING SETTING *Notes: The feedback page was shown to both persuader and receiver participants.*