

Learning Preferences through Noisy Communication in the Agentic Economy

Samir Huseynov
Auburn University
szh0158@auburn.edu

October 8, 2026

Abstract

As humans increasingly delegate economic decisions to AI agents, market outcomes depend on how accurately Agents learn and represent their Principals’ preferences. Using a controlled exchange environment, we study how exogenously manipulated noise in principal-agent communication causally affects preference learning and the gains from trade. We use Gemini Flash, GPT-Sol, and Claude Fable LLM models as AI agents in different market conditions, whereas synthetic Principals are endowed with objects with induced valuations. Agents report Principals’ valuations of objects after conversations with them. Then, the Top Trading Cycles market allocates objects based on Agents’ estimates. We find that LLM models differ in information-seeking behavior, but mostly exhibit decisions aligning with Bayesian benchmarks. However, this alignment can still lead to substantial welfare losses. Raising noise from zero to 35% reduces total market pay-off, removing 65–83% of the zero-noise-communication gains from trade. We conclude that the economic value of AI delegation depends on the quality of preference information available to Agents.

1 Introduction

As Artificial Intelligence (AI) adoption across a broad spectrum of economic activities becomes increasingly common, humans delegate more economic decisions to AI agents (Shahidi et al., 2026). The transition from AI assistance to AI delegation gives rise to a new economic system—the *Agentic Economy*—in which AI agents can act on humans’ behalf in economic transactions. The emergence of the Agentic Economy necessitates understanding human-AI agent interactions and how they translate into economic decisions and transactions. To a large extent, the success of this new economic system will depend on how well AI agents understand and map human economic preferences, and on how accurately they represent those preferences in market transactions (Hitzig et al., 2026). Consequently, systematic miscalibration of human preferences by AI agents could have substantial negative welfare

consequences (Priyanshu et al., 2026).

Using a controlled setting, this article studies how noise in human-AI agent interactions translates into market outcomes. We construct Top Trading Cycles (TTC) markets with synthetic human preferences (hereafter, *Principals*). A market consists of six Principals and six objects. Each Principal is endowed with one unique object, and then we induce Principals’ valuations over all objects, creating *true values*. Since we observe the full valuation vector of each Principal, preference-learning error is directly measurable. Each Principal is assigned an AI agent (hereafter, *Agent*), to which the Principal’s true valuations are not known. Agents communicate with Principals and estimate how much each Principal values the objects. Across different market conditions, we deploy three LLM configurations—Gemini Flash, GPT-Sol, and Claude Fable—to gauge their performance as Agents. Market allocations occur based on Agents’ reported valuations.

We exogenously vary the noise in Principals’ communications with Agents. To our knowledge, this is the first controlled exchange-market experiment to manipulate the communication noise in a synthetic principal-AI agent framework and map its effects on LLM preference learning and gains from trade. Communication noise is governed by a known answer-error probability, which is reported to Agents and allows us to construct Bayesian benchmarks for evaluating their decisions. By holding preferences and initial ownership fixed while varying communication noise and LLM model, this experimental setting allows us to study how noise affects Agents’ mapping of human preferences and, ultimately, the market outcomes generated by those decisions.

Our research inquiry is organized around three core questions. First, does noise in human-AI communication reduce the value created through economic transactions? Second, when communicating with Principals, how do Agents choose both the extensive margin—whether and how much additional information to seek—and the intensive margin—which questions to ask? Third, to what extent do Agents’ decisions and the resulting market outcomes deviate from the Bayesian benchmark?

We find that noisier communication reduces gains from

trade. Compared with the 0%-noise baseline, total payoff at 35% noise falls by 6.22 points for Gemini, 6.83 for Sol, and 7.89 for Fable. These losses remove 65%, 74%, and 83% of the gains generated by trading at zero noise. At 25% noise, Sol loses around 4.5 fewer total payoff points than Fable, suggesting models might behave differently at different noise levels. Compared to the 0%-noise baseline, trading increases total payoff at all noise levels, albeit not all Principals always conclude the market with a gain. At 35% noise, approximately 14–18% of Principals receive an object worth less to them than their initial endowment, compared with approximately 1–2% at zero noise.

Agents receive one opening prompt from Principals and can ask up to four follow-up questions before concluding their value estimates. However, we do not disclose this question budget in advance. Sol and Gemini exhaust nearly all available questions, whereas Fable generally stops before the limit and asks fewer questions as noise increases. Thus, models do not exhibit the same information-seeking behavior at the extensive margin.

We construct a Bayesian benchmark based on the follow-up questions Agents pose to Principals. The benchmark allows us to measure how closely Agents’ value estimates correspond to the information available to them. As communication noise rises, Gemini’s numerical deviations from this benchmark increase significantly. However, we do not detect statistically significant changes in these deviations for Sol or Fable when the noise level rises.

We also construct an adaptive Bayesian question policy to explore whether different follow-up question choices by Agents could improve preference learning. We do not detect statistically different trade gains when Agents follow this policy, suggesting that LLM models behave close to the optimal at the intensive margin of information gathering.

1.1 Related work and contribution

We build on [Hitzig et al. \(2026\)](#), who study how agents learn human preferences and represent them in centralized and decentralized markets. The authors create a barter economy in which agents infer preferences from five-minute conversations with human subjects and then execute trades. Agents’ rankings agree with participants’ rankings on around 61% of objects. Agentic representation improves with greater participant effort in human-AI conversations. However, the authors do not control for the noise or information quality in those conversations to establish causality. Our contribution is that, by exogenously varying communication noise while observing the full underlying preference vector, we measure agentic preference learning errors and trace their consequences in subsequent market outcomes.

However, recent work by [Priyanshu et al. \(2026\)](#) shows that providing Agents with more context can also backfire. The authors show that wealth-related information can

steer agentic recommendations toward more expensive items even when principals request a cheaper alternative. Our study focuses on a different source of misrepresentation. In our setting, Principals’ preferences remain fixed, but communication garbling reduces the reliability of their answers. Thus, we investigate how the communication noise leads to agentic misrepresentation.

Another strand of the literature studies LLM proxies for human preferences ([Huang et al., 2026](#); [Soumalias et al., 2025](#); [Dong et al., 2026](#)). These approaches use limited information from people to construct proxies and to support subsequent preference queries in economic decisions. We complement this work by examining how errors in preference communication might affect the quality of proxies and the following market allocations. Moreover, our induced-preference design also separates known preferences from observed agentic estimates, complementing revealed-preference approaches to evaluating AI alignment ([Suleymanov, 2026](#)).

Finally, [Handa et al. \(2024\)](#) combine language models with Bayesian optimal experimental design to elicit preferences. Our unassisted Agents are not instructed to apply Bayes’ rule or to follow an optimal query policy. We observe their choices within a scripted communication interface and compare their reports with benchmarks. Thus, we contribute by showing that AI agents’ reasoning is closely aligned with Bayesian benchmarks.

2 Study environment and experimental design

Figure 1 shows the general outline of our research design. We first generate six synthetic Principals with known individual valuations for six indivisible objects, A–F. Each Principal is endowed with one unique object. We hold the six preference types and initial ownership fixed and call this configuration a *world*. We construct 18 new worlds across study conditions by drawing new preference types and ownership assignments.

Each Principal is assigned one Agent with an isolated conversation channel with a scripted format. The Principal activates the channel first by sending a scripted prompt indicating a preference between A and B. Then Agents can ask follow-up questions in the fixed format of “Between object X and object Y, which would you choose?” We allow the Agent to select an object pair from the fifteen possible pairs with replacement. Thus, our study design fixes the form of the question, while leaving the selection of information (i.e., object pair) to the Agent. This design approach also prevents Agents from asking directly for numerical utilities or preference weights.

After each answer, the Agent can ask another question or stop the information gathering and submit six value estimates. Agents are allowed up to four follow-up ques-

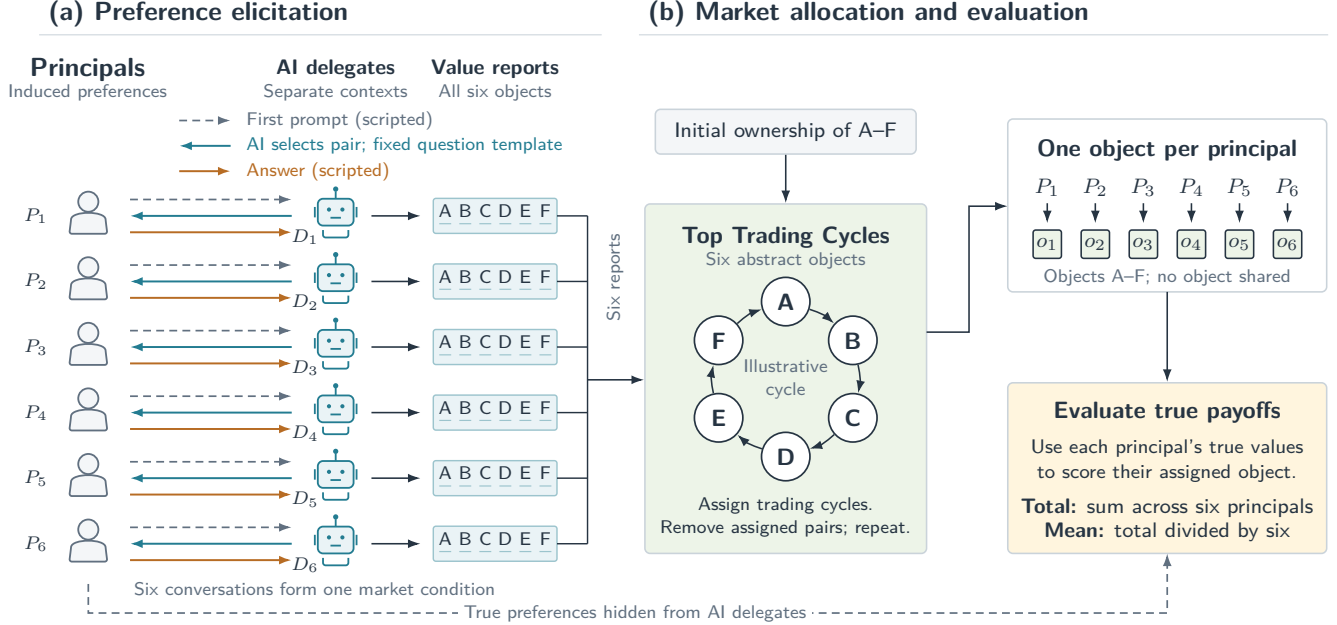


Figure 1: From a Principal’s preferences to the market allocation. In our framework, six synthetic Principals with induced preferences each interact with a separate AI delegate. The opening prompt by Principals and subsequent answers are scripted. Agents select follow-up object pairs using a fixed question template and decide when to stop and submit their estimates. Each Agent can ask up to four follow-up questions, and this question budget is ex-ante unknown to Agents. After Agents submit six value reports, the TTC executes based on the initial ownership. We then calculate each Principal’s final utility using their induced true preferences to assess how the communication noise level affects market welfare. The online supplement provides more details on TTC procedures.

tions in addition to the opening prompt, and they are not informed about this limit. If Agents exhaust their question budget, our protocol closes the communication channel and requires them to report their estimates of how much the Principal values each object. After all six Agents conclude their value estimates, the market reassigns objects using the Top Trading Cycles format, where we break ties using alphabetical priority. (Shapley and Scarf, 1974; Roth, 1982).

We deploy Gemini 3.8 Flash, GPT-6 Sol, and Claude Fable 5.1 LLM models across study worlds. For each world and model combination, we assign communication-noise probabilities of 0%, 10%, 25%, and 35%. Thus, each model-by-noise combination yields a distinct market condition. As Figure 2 illustrates, one complete world contains twelve conditions and seventy-two conversations. Agents are only informed about the utility framework, object attributes, prior distribution, and the communication noise level and are not instructed to act based on Bayesian updating.

2.1 Utility framework

In our setting, we fix object-level attributes across all worlds as follows:

Object	A	B	C	D	E	F
Attribute 1	3	0	0	2	1	0
Attribute 2	0	3	0	1	0	2
Attribute 3	0	0	3	0	2	1

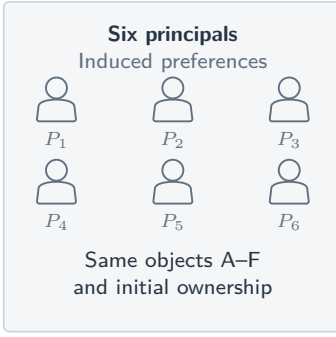
For a Principal with type $\theta = (w_1, w_2, w_3, \gamma)$, the true utility of object j is

$$u_{\theta}(j) = w_1 x_{j1} + w_2 x_{j2} + w_3 x_{j3} + \gamma x_{j1} x_{j2}. \quad (1)$$

Here x_{jk} is attribute k of object j , and each weight w_k can take values 1, 2, or 3, while γ can take values -2 , 0, or 2. There are $3^3 \times 3 = 81$ possible preference types. For each world, we independently and randomly draw preference types, creating synthetic Principals. The interaction parameter allows the first two attributes to matter beyond their separate contributions. This approach mimics more complex real life choice situations where Principals with identical weights on individual attributes can rank objects differently.

Agents know the objects and their attribute levels. Fixing objects creates a common economic environment, allowing us to attribute differences between worlds to distinct preferences and ownership statuses. It also enables posteriors to be comparable across the worlds.

(a) Hold one world fixed



(b) Repeat under twelve conditions

		Gemini	GPT Sol	Claude Fable
Answer-error probability	0%	6 conversations	6 conversations	6 conversations
	10%	6 conversations	6 conversations	6 conversations
	25%	6 conversations	6 conversations	6 conversations
	35%	6 conversations	6 conversations	6 conversations

Each cell is one market; all six delegates use that column's model.

One complete world = 12 markets = 72 conversations

3 models $\times$ 4 error rates $\times$ 6 principals

Figure 2: How one world generates the experimental conditions. Each study cell contains six principal–delegate conversations with the indicated model and communication-noise probability. All twelve cells are nested in the world.

2.2 Our Sample

Our sample contains 1,342 complete conversations and 216 complete market conditions across nineteen worlds. Due to failed conversations, the Sol and Fable models have eighteen complete comparison pairs of zero- and 35%-noise markets, whereas Gemini has seventeen valid pairs at 10% and 25% noise levels. We collected data through two sessions on September 30 and October 2. Appendix C.3 reports session-sensitivity analyses and Appendix C addresses missing markets.

3 Empirical approach

3.1 Total market welfare

Our primary empirical approach compares market conditions with different communication-noise levels while holding the world and model fixed. Let w identify a world, m a model, and r a non-zero noise rate. Let $a_{wmr}(i)$ be the object assigned to Principal i , whose true utility is u_{wi} . Total market welfare is

$$W_{wmr} = \sum_{i=1}^6 u_{wi}(a_{wmr}(i)). \quad (2)$$

For each non-zero noise level, we calculate the within-world market welfare change compared to the 0%-noise level with $d_{wmr} = W_{wmr} - W_{wm0}$ and average it over worlds where both relevant market pairs are complete.

Our primary regression analyses follow this model specification:

$$W_{wms} = \alpha_{wm} + \tau_{mr} \mathbf{1}\{s = r\} + \varepsilon_{wms}, \quad s \in \{0, r\}. \quad (3)$$

The dependent variable is the total true market payoff. We control for the world fixed effect (α_{wm}), and τ_{mr} is

the mean welfare change from zero noise to $r > 0$. The residual captures payoff variation not explained by these terms.

3.2 Bayesian benchmark

For each conversation h , we compute posterior probabilities over the eighty-one types and the implied value of object j :

$$\begin{aligned} p(\theta | h) &= \frac{p(h | \theta)p(\theta)}{\sum_{\theta'} p(h | \theta')p(\theta')}, \\ b_j(h) &= \sum_{\theta} p(\theta | h)u_{\theta}(j). \end{aligned} \quad (4)$$

The term $p(\theta)$ represents the known uniform prior, that is, the Principal has the same probability of possessing one of the available 81 preference types. The likelihood $p(h | \theta)$ function shows the probability of observing the conversation history conditional on each preference type. Thus, $b_j(h)$ is the exact posterior expected utility of object j conditional on the Agent's information set.

To construct a Bayesian benchmark, we replace all six Agents' reports with these posterior values and re-run TTC, holding ownership and tie-breaking rules fixed. This provides a *same-evidence Bayesian benchmark* or a benchmark conditional on asked actual follow-up questions. We separately examine optimal question selection by Agents in Section 4.4.

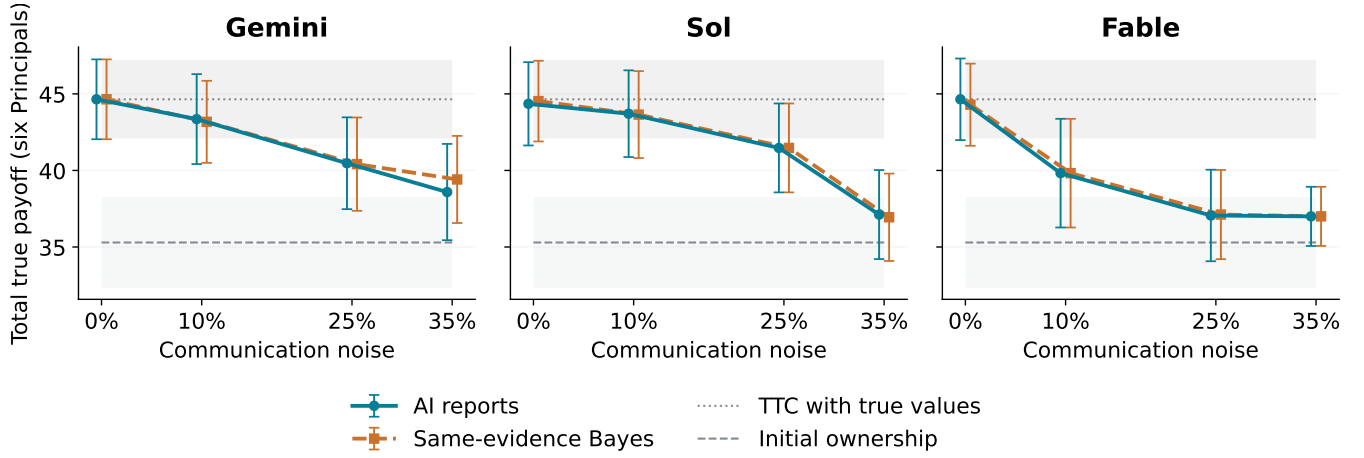


Figure 3: AI allocations and the same-evidence Bayesian benchmark. Solid teal and dashed orange curves show total true payoff under AI and exact posterior reports from the same conversations.

4 Results: communication noise, preference learning and trade

4.1 How much value is lost when communications become noisy?

We first investigate how communication noise affects total market payoff. Table 1 reports the average within-world welfare change for each model at each positive noise rate. At 10%, we do not detect a welfare change for Gemini or Sol, while Fable loses 4.17 points. At 25% and 35%, market payoff falls significantly. At the 35% noise level, the losses are 6.22, 6.83, and 7.89 points for Gemini, Sol, and Fable, respectively.¹

Table 1: Welfare loss at each communication-noise rate

Model / comparison	N	Change	95% CI	Holm p
Gemini / 10%–0%	17	−1.29	[−2.87, 0.28]	0.2002
Gemini / 25%–0%	17	−4.18**	[−6.49, −1.86]	0.0075
Gemini / 35%–0%	18	−6.22**	[−8.99, −3.45]	0.0012
Sol / 10%–0%	18	−0.33	[−1.80, 1.13]	0.6378
Sol / 25%–0%	18	−3.22*	[−5.61, −0.84]	0.0332
Sol / 35%–0%	18	−6.83**	[−9.84, −3.83]	0.0012
Fable / 10%–0%	18	−4.17*	[−6.72, −1.61]	0.0125
Fable / 25%–0%	18	−7.67***	[−10.08, −5.25]	< 0.0001
Fable / 35%–0%	18	−7.89***	[−9.95, −5.83]	< 0.0001

N is the number of worlds where both markets are complete. Intervals are pointwise 95% paired-world t intervals. Holm p -values: * $p < .05$, ** $p < .01$, *** $p < .001$.

To put these losses in perspective, we compare the value of initial ownership with payoff after zero-noise trading. Their difference measures the initial gain from trade. We then calculate how much of this gain is lost at 35% noise. Table 2 shows that the gain from trade at the zero-noise baseline is approximately nine to ten points. Between 65% and 83% of this gain disappears at the 35% noise

level, leaving only 1.61–3.28 points of trade gain above initial ownership. Figure 3 shows the corresponding payoff levels alongside the Bayesian benchmarks. Nevertheless, total payoff remains above no trade at the 35% noise level. In Section 4.5, we show that the gain from trade is not uniformly distributed among Principals.

Table 2: From ownership to trade: where the welfare loss occurs

Total true payoff / calculation	Gemini	Sol	Fable
Initial ownership (A)	34.61	34.61	34.61
After trading, 0% noise (B)	44.11	43.83	44.11
After trading, 35% noise (C)	37.89	37.00	36.22
Gain from trade at 0% (B–A)	9.50	9.22	9.50
Payoff lost to noise (B–C)	6.22	6.83	7.89
Gain remaining at 35% (C–A)	3.28	2.39	1.61
Fraction of initial gain lost	65%	74%	83%

Values are mean total utility points. We define Fraction lost as $(B-C)/(B-A)$.

4.2 When do agentic value-estimate differences affect welfare?

We next ask whether noisier communication makes Agents’ numerical reports depart further from the Bayesian benchmark constructed based on the conversation history. For each object, we take the absolute difference between the Agent’s reported value and the exact posterior mean, then average across the six objects and six Principals. Put differently, this measure informs how agentic value estimates deviate from the Bayesian benchmark.

Table 3 reports how this deviation changes relative to zero noise. Gemini’s deviation increases by 0.131, 0.303, and 0.487 points at 10%, 25%, and 35% noise, and our results are robust to multiple-testing adjustments. However, we do not detect corresponding changes for Sol or Fable.²

¹Appendix A, Table A.1 and Table A.2 present complete regression specifications.

²Table B.1 in Appendix B provides formal regression results for

Table 3: Does noise increase the distance from Bayesian values?

Model / comparison	Change	95% CI	Holm p
Gemini / 10%–0%	0.131***	[0.097, 0.165]	< 0.0001
Gemini / 25%–0%	0.303***	[0.236, 0.371]	< 0.0001
Gemini / 35%–0%	0.487***	[0.420, 0.553]	< 0.0001
Sol / 10%–0%	−0.047	[−0.108, 0.015]	0.5083
Sol / 25%–0%	−0.062	[−0.120, −0.004]	0.1903
Sol / 35%–0%	−0.042	[−0.116, 0.032]	0.6361
Fable / 10%–0%	0.010	[−0.029, 0.049]	0.6361
Fable / 25%–0%	−0.021	[−0.054, 0.013]	0.6361
Fable / 35%–0%	−0.038	[−0.069, −0.006]	0.1281

At each noise level, we compare Agents’ reported object values with Bayesian benchmark estimates constructed from the same conversations. Then we calculate the absolute difference for each object and average these differences across the six objects and six Principals within each world. Each row reports how this average discrepancy changes from zero noise to the indicated noise level. A positive estimate means that Agents’ reports move farther from their corresponding Bayesian estimates; a negative estimate means they move closer. We report world-paired pointwise 95% intervals. Holm p -values: * p < .05, ** p < .01, *** p < .001.

It must be noted that agentic value misestimates affect the TTC allocation only when they change relevant ordinal comparisons. Even when rankings change, they can still have no material impact on TTC allocations, as they may concern objects that are no longer available in remaining TTC rounds. Table 4 follows reporting differences through all complete markets. In 107 markets, every Principal has the same full ranking under AI and Bayesian reports. In another 73 markets, some rankings differ, but the choices among remaining objects are identical in every round. In 12 markets, temporary pointers differ, but final assignments remain unaffected. Thus, TTC procedures partly alleviate the impact of agentic value misestimates.

Table 4: Following a reporting difference through the market

Comparison with same-evidence Bayes	Markets
Same full rankings; same objects assigned	107
Different rankings; same round-by-round pointers	73
Different temporary pointers; same final objects	12
Different final objects; total gains offset losses	4
Different final objects; higher total true payoff	12
Different final objects; lower total true payoff	8

Thus, 192 of 216 markets produce identical object assignments, although only 107 preserve all rankings. For Gemini at the 35% noise level, none of the eighteen markets has identical full rankings due to value misestimates, yet thirteen markets still yield identical allocations to true preferences (Table B.6). This is why, in our setting, greater numerical divergence from true preferences does not necessarily translate into a welfare penalty.

However, in the remaining twenty-four markets, object assignments change, where twelve markets see total payoff increase under AI agent reports, while eight experience these deviations.

market payoff loss. In four markets, gains and losses offset each other.

We next examine whether Agents’ deviations from Bayesian value estimates amplify the welfare consequences of noisy communication. For each world, we calculate market payoff under two reporting rules. The first uses Agents’ actual value reports. The second replaces these reports with Bayesian estimates constructed from the same questions and answers. We perform this comparison at both zero and 35% noise. Consequently, we ask whether increasing noise produces a larger payoff decline under Agents’ reports than under the corresponding Bayesian reports.

For Gemini, increasing noise from zero to 35% reduces total market payoff by 6.22 points under Agents’ reports, compared with 5.44 points under Bayesian reports. This difference becomes larger by 0.78 points under Gemini’s reports. For Sol, the decline under Agents’ reports is approximately 0.33 points smaller than under Bayesian reports, whereas for Fable, it is approximately 0.33 points larger.

Table B.5 reports these additional payoff changes and their uncertainty. The estimates are −0.78 points for Gemini, +0.33 for Sol, and −0.33 for Fable, with 95% confidence intervals of [−2.35, 0.79], [−1.06, 1.73], and [−1.04, 0.37], respectively. None is statistically distinguishable from zero after adjustment for multiple comparisons. Thus, the substantial welfare losses documented above also arise when the market uses Bayesian estimates based on the available information. We find no statistically detectable additional loss associated with Agents’ departures from those estimates.

4.3 How do Agents learn and decide when to stop?

Agents can ask follow-up questions or stop information gathering and submit their estimates. They select which object pair to include in the scripted message should they decide to ask the follow-up question. We therefore observe both how much information Agents request (i.e., extensive margin) and which object pair comparisons they select in their follow-ups (i.e., intensive margin).

Figure 4 shows that Gemini and Sol use nearly all available follow-up questions across noise rates. At 35%, both reach four follow-ups in every conversation. Fable instead averages approximately 2.55 follow-ups at zero noise and 0.78 at 35% on the seventeen common worlds, suggesting that its information seeking falls as communication becomes less reliable.³ However, as the middle and right plots of Figure 4 show, the extensive margin differences in Fable’s information-seeking behavior do not harm the accuracy of its value estimates.

Table 5 examines whether the Agent identifies the truly

³Appendix B.2, Table B.2, presents formal available-pair regressions for each noise rate relative to zero.

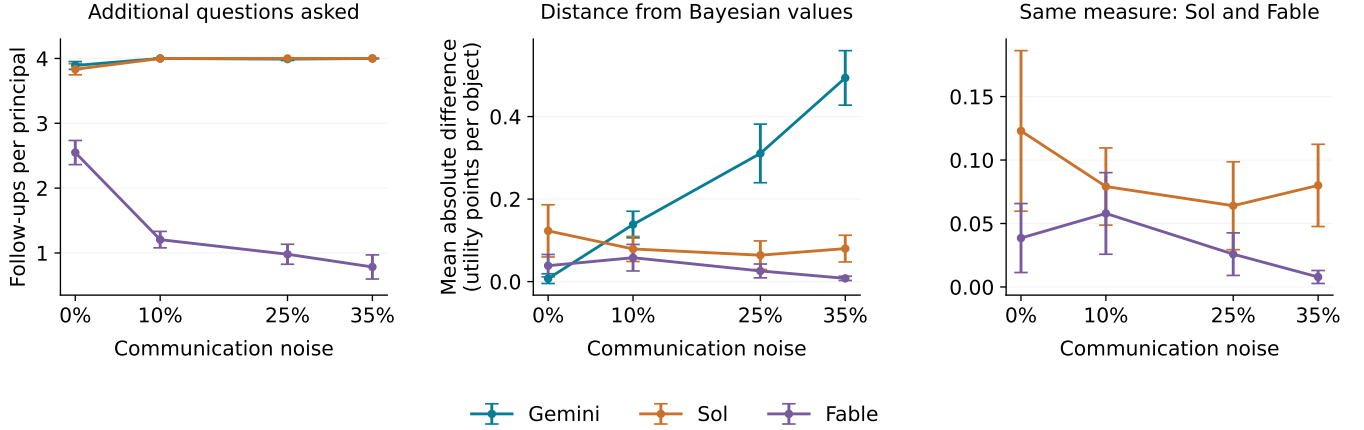


Figure 4: Question counts and distance from Bayesian reports. Left plot: optional follow-ups per Principal, excluding the opening prompt. Middle plot: mean absolute report-minus-Bayes discrepancy per object. Right plot: the same discrepancy for Sol and Fable on an enlarged scale.

most-preferred object of Principals. At zero noise, all Agents identify the true favorite, whereas at 35%, this share falls to approximately 45% for Gemini, 46% for Sol, and 41% for Fable, respectively. However, correctly identifying a favorite does not guarantee receiving it, as other Principals may compete for that object, making lower-ranked comparisons also matter for TTC.⁴

Table 5: Correct identification of a true favorite: changes from zero noise

Model / comparison	0% level	Change	Holm p
Gemini / 10%–0%	100.00	−10.78*	0.0194
Gemini / 25%–0%	100.00	−34.31***	< 0.0001
Gemini / 35%–0%	100.00	−54.63***	< 0.0001
Sol / 10%–0%	100.00	−13.89*	0.0114
Sol / 25%–0%	100.00	−36.11***	< 0.0001
Sol / 35%–0%	100.00	−53.70***	< 0.0001
Fable / 10%–0%	100.00	−30.56***	< 0.0001
Fable / 25%–0%	100.00	−52.78***	< 0.0001
Fable / 35%–0%	100.00	−59.26***	< 0.0001

Change shows percentage point differences. Holm p -values:

* $p < .05$, ** $p < .01$, *** $p < .001$.

4.4 Could a different question sequence improve the outcome?

Our first Bayesian benchmark is conditional on the opening prompt and follow-up questions and answers in Agents’ original conversations. But it raises the question of whether an Agent could have learned more by selecting different object pairs in the questions it posed. In Online Supplement Section S4, we describe an adaptive question selection policy that chooses each object pair after observing previous replies and accounts for possible future answers. Its objective is, for each communication

noise level, to minimize expected squared error in the six final object value estimates, conditional on the known prior and question budget. This policy updates after each answer the Principal provides and plans the optimal object pairs for the remaining questions. We evaluate this policy through simulations.

Changing which questions are asked: We first allow the policy to pose exactly as many follow-up questions as each Agent had in its original conversation. For instance, if an Agent asked one follow-up question, our policy also gets only one question to pose. The important distinction is that the policy is informed about this question budget in advance, whereas the original Agent chose when to stop without knowing the numerical cap.

For each market, we construct Bayesian object value estimates from the Agents’ original conversations and the adaptive policy’s simulated conversations. We then run TTC separately on these two sets of reports and compare total true market payoff. Using Bayesian reports in both cases ensures that our comparison concerns the information collected, rather than differences in Agents’ numerical reporting. We construct these comparisons within the same world and noise level to make them reasonably comparable.

Online Supplement Table S4.3 reports that, for Fable at 25% noise, the adaptive policy produces an average market payoff of 36.82 points, compared to 37.12 points with our Bayesian benchmark estimates. The estimated difference is therefore −0.29 points, with a 95% confidence interval of [−2.04, 1.46]. We repeat this procedure for all twelve model–noise comparisons, and none is statistically significant after Holm adjustment for multiple hypothesis testing. We therefore find no statistically detectable welfare improvement from replacing Agents’ question choices within the limit of their actual question counts.

Allowing more questions: We next give the adaptive policy four follow-up questions for every Principal, even for the conversations that Agents stopped earlier. For

⁴Appendix B.2, Table B.3, reports the corresponding world-fixed-effect regressions, standard errors, intervals and multiple-testing adjustments.

Fable at 25% noise, the estimated payoff improvement becomes 3.24 points relative to Bayesian reports based on its original conversations (Online Supplement Table S4.4).

Note that even this optimal learning policy faces substantial information gathering constraints in our study design. At 35% noise, with four optimally selected follow-up questions, our policy can only reduce expected squared value-estimation error by 7.5% relative to the information available after the opening prompt (Online Supplement Table S4.1). Consequently, a short conversation can leave considerable uncertainty about preferences even when questions are selected to maximize learning accuracy.

Choosing questions for market payoff: One drawback of optimizing question choices in our Bayesian adaptive policy is that this approach does not necessarily improve market welfare. Some objects may be unavailable to that Principal because of other participants’ reports and initial ownership. We therefore construct a second adaptive policy that selects questions to maximize the Principal’s expected payoff from the object assigned through TTC. For this comparison, we change one Principal’s conversation at a time while holding the other five Bayesian reports fixed. We give both policies the same question count and compare the represented Principal’s payoff under the market-payoff policy with our previous value-learning policy. Both policies submit Bayesian value estimates.

Table 6: Does targeting the assigned object improve the Principal’s payoff?

Model / noise	Expected gain	Realized gain	95% CI
Gemini / 0%	0.01	0.03	[-0.03, 0.09]
Gemini / 10%	0.17	0.29	[0.05, 0.54]
Gemini / 25%	0.20	0.10	[-0.06, 0.26]
Gemini / 35%	0.15	0.06	[-0.25, 0.37]
Sol / 0%	0.01	0.07	[-0.01, 0.16]
Sol / 10%	0.16	0.17	[-0.06, 0.39]
Sol / 25%	0.18	0.09	[-0.29, 0.48]
Sol / 35%	0.18	0.01	[-0.36, 0.37]
Fable / 0%	0.11	0.10	[0.02, 0.18]
Fable / 10%	0.14	-0.06	[-0.30, 0.18]
Fable / 25%	0.08	-0.03	[-0.18, 0.12]
Fable / 35%	0.05	0.04	[-0.22, 0.30]

Table 6 reports model-implied average expected gains of approximately 0.01–0.20 points per Principal.⁵ None of the twelve realized payoff differences is statistically significant after Holm adjustment. We conclude that, within the limitations of our study design, models’ information-seeking behavior was close to the optimal at the intensive margin.

⁵Appendix D, Tables D.1 and D.2 present the supporting regression analyses.

4.5 Can aggregate trade gains conceal individual harm?

We define *harm* when a Principal is allocated an object that has lower true value than the object initially endowed. At zero noise, the harmed shares are 0.93% for Gemini, 0.93% for Sol, and 1.85% for Fable. At 35%, these proportions rise to 14.81%, 17.59%, and 13.89%, respectively. This coincides with the share of harmed Principals increasing by 13.89, 16.67, and 12.04 percentage points (Table B.4).

Our findings warn that positive aggregate trade gains do not guarantee that AI delegation benefits each Principal. A trade can benefit several participants and yield enough positive gains to offset losses incurred by others. The TTC mechanism protects participants relative to their initial objectives according to the AI-submitted rankings. However, when those rankings misrepresent true preferences, an apparently beneficial trade can leave the Principal worse off. Future studies might explore how AI-delegated economic transactions could affect income inequality.

We next investigate whether these are the same Principals who would be harmed if every Agent reported the exact Bayesian values implied by its conversation. To conduct this exercise, we hold all questions and answers fixed and change only the AI reports submitted to TTC. Table 7 divides the Principal allocations into four mutually exclusive categories. Of the 126 allocations that leave a Principal worse off under AI reports, 119 also do so under Bayesian reports. Seven Principal allocations are harmed only under AI reports. Conversely, eight allocations produce harm only under Bayesian reports. Thus, the distribution of harmed Principals slightly changes when we replace AI agent reports with Bayesian reports.⁶

Table 7: Do the same Principals lose under AI and Bayesian reports?

Model	Allocations	Both	AI only	Bayes only	Neither
Gemini	420	33	2	6	379
Sol	432	34	4	1	393
Fable	444	52	1	1	390
All	1296	119	7	8	1162

The 1,296 allocations belong to 216 markets and exclude conversations without a complete six-person market. “AI only” means harmed under AI reports but not Bayesian reports; “Bayes only” reverses this comparison.

4.6 Do models differ in their response to communication noise?

We next examine whether communication noise affects market payoff differently across models. For each world, we first calculate each model’s payoff change relative to its

⁶Appendix D.3, Tables D.4–D.6 extend these analyses by exploring harm under different communication noise levels.

own zero-noise baseline. We then compare these changes between Gemini and Sol, Gemini and Fable, and Sol and Fable at each positive noise level.

At 25% noise, Sol loses 4.71 fewer points on the seventeen worlds across all models and rates, with a 95% interval of [2.52, 6.89] and Holm-adjusted $p = 0.0029$. Using all eighteen available Sol–Fable pairs gives 4.44 points, with interval [2.32, 6.57] and adjusted $p = 0.0034$ (Tables A.4 and A.5). However, Gemini does not exhibit differential behavior.

5 Discussion

To the best of our knowledge, we present the first controlled study that exogenously varies communication noise in interactions between synthetic Principals and AI agents. We trace how different communication noise levels affect agentic preference learning and subsequent market outcomes. Our study speaks to an emerging economic phenomenon—the Agentic Economy—in which humans increasingly delegate their economic decisions and activities to AI agents. We discuss how noise in human–AI agent interactions might affect economic outcomes in this emerging economy.

We document that increasing communication noise erodes a larger share of the gains from trade achieved under the zero-noise condition, and that this loss is not uniformly distributed among Principals. The TTC market structure prevents welfare losses to some extent. Therefore, future studies can explore how market welfare losses associated with communication noise are amplified or reduced under different trading systems.

We also show that different LLMs generate different trading outcomes, although their performance is mostly similar. Fable is more sensitive to small increases in communication noise. However, at the 35% noise level, all models behave similarly. This finding suggests that different LLMs can completely prevent losses in economic outcomes up to a certain communication-noise threshold. The practical implication is that reducing noise in human–AI agent interactions below a certain level might ensure that the Agentic Economy operates efficiently.

We are also the first to formally show, in an economic principal–agent setting, that LLM behavior closely aligns with Bayesian benchmarks. Our study design does not allow us to observe whether the internal processes of the deployed LLMs follow Bayesian reasoning. However, we show that Agents’ reported values are very close to Bayesian estimates.

One of the most appealing advantages of our study design is that we can distinguish between the extensive and intensive margins of agentic information-seeking behavior. We find that models differ significantly in how many follow-up questions they ask. Sol and Gemini almost always exhaust the available question budget, whereas Fable asks fewer questions as communication becomes

noisier. We do not find that this difference at the extensive margin translates into detectable changes in market welfare. This is partly due to the TTC mechanism. Future studies might investigate which trading mechanisms in the Agentic Economy are more vulnerable to differences in agentic information-seeking behavior.

We also construct different Bayesian policies showing that models mostly behave optimally at the intensive margin of information gathering. However, because we use a scripted communication protocol, our results might be affected by restrictions imposed by the study design.

Overall, we contribute to this emerging literature by establishing that communication noise in human–AI agent interactions causally affects market outcomes. Using a well-controlled experimental environment allows us to minimize confounds and trace how different noise levels influence gains from trade, information-seeking behavior, and the distribution of welfare. However, we provide AI agents with limited contextual information about their Principals. Future studies can extend this dimension and investigate how AI agents treat human Principals with different backgrounds.

Disclosure of AI Assistance

During this project, we used OpenAI models through Codex and Anthropic Claude to assist with methodological discussions, the development and checking of research code, and the preparation of tables and figures.

These tools also assisted with manuscript drafting and editing, including explanatory text, table notes, and figure captions in the appendix and online supplement.

AI tools were extensively used in compiling Appendix and the Online Supplement. However, we edited their language to improve readability.

We developed the research questions, directed the study, and reviewed and revised the manuscript, including the main text. We retain full responsibility for the methods, results, interpretation, references, and final content.

References

- L. Dong, K. Luo, and F. Xu. From product search to preference articulation: The economics of agentic commerce. arXiv:2608.08395, 2026. URL <https://arxiv.org/abs/2608.08395>.
- K. Handa, Y. Gal, E. Pavlick, N. Goodman, J. Andreas, A. Tamkin, and B. Z. Li. Bayesian preference elicitation with language models. arXiv:2403.05534, 2024. URL <https://arxiv.org/abs/2403.05534>.
- Z. Hitzig, S. Carr, T. Cotter, K. Troy, K. Turman, M. Massenkoff, and P. McCrory. Project swap: What happens when agents trade for us? Anthropic, 2026. URL <https://www.anthropic.com/research/project-swap>. September 24.
- D. Huang, E. Lock, F. Marmolejo-Cossío, and D. C. Parkes. Accelerated preference elicitation with LLM-based proxies. In *Web and Internet Economics: WINE 2025*, pages 631–648. Springer, 2026. URL <https://parkes.seas.harvard.edu/sites/g/files/omnuum8711/files/2025-12/accelerated.pdf>. Conference held in 2025; proceedings published in 2026.
- A. Priyanshu, S. Vijay, B. Jabarian, and N. Mireshghallah. Et tu, brute? economic misalignment in personal AI agents. arXiv:2609.24927v2, 2026. URL <https://arxiv.org/abs/2609.24927v2>. Version 2, September 25.
- A. E. Roth. Incentive compatibility in a market with indivisible goods. *Economics Letters*, 9:127–132, 1982. URL https://stanford.edu/~alroth/papers/1982_EL_IncentiveCompatibility.pdf.
- P. Shahidi, G. Rusak, B. S. Manning, A. Fradkin, and J. J. Horton. The coasean singularity? demand, supply, and market design with AI agents. In A. K. Agrawal, E. Brynjolfsson, and A. Korinek, editors, *The Economics of Transformative AI*. University of Chicago Press, 2026. URL <https://www.nber.org/books-and-chapters/economics-transformative-ai/coasean-singularity-demand-supply-and-market-design-ai-agents>.
- L. Shapley and H. Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1:23–37, 1974. doi: 10.1016/0304-4068(74)90033-0.
- E. Soumalias, Y. Jiang, K. Zhu, M. Curry, S. Seuken, and D. C. Parkes. LLM-powered preference elicitation in combinatorial assignment. arXiv:2502.10308, 2025. URL <https://arxiv.org/abs/2502.10308>.
- E. Suleymanov. A revealed preference framework for AI alignment. arXiv:2603.27868, 2026. URL <https://arxiv.org/abs/2603.27868>.

A Regressions supporting the main welfare comparisons

In this appendix, we present formal regressions supporting the main results. We also examine whether our findings change under alternative samples and statistical methods. The online supplement provides further details of study protocol and additional analyses.

Main-text question	Supporting regression	Sample and statistical inference
How does noise affect total payoff?	Each positive noise rate versus zero, with world fixed effects; Table A.1.	We use worlds with both markets complete, cluster CR2 standard errors by world, and apply Holm across nine comparisons.
Do models respond differently to noise?	Pooled regression with world-by-model fixed effects; Tables A.3–A.5.	We compare seventeen common worlds with all available paired worlds and apply separate Holm adjustments across nine comparisons.
Do reports move farther from Bayesian estimates?	Three noise indicators and world fixed effects; Table B.1.	We retain worlds complete at all four rates within each model and use CR2 standard errors and Holm across nine comparisons.
How does noise affect questioning, favorite identification, and harm?	Each positive noise rate versus zero, with world fixed effects; Appendix B.2.	We use available paired worlds, CR2 standard errors, and Holm across thirty-six comparisons covering four outcomes.
Do reporting discrepancies add to the welfare loss?	Zero versus 35% noise, with world fixed effects; Table B.5.	We use eighteen paired worlds per model, CR2 standard errors, and Holm across three comparisons.
What do the question-policy and harm comparisons show?	Regressions of differences averaged within each world; Appendix D.	We use complete markets from 17–19 worlds per condition and apply Holm within four twelve-test families and across all forty-eight tests.

A.1 Within-model welfare effects

We first examine how increasing communication noise changes total true market payoff for each model. We compare a world’s market at the indicated positive noise level with its market at zero noise. We estimate this comparison by regressing payoff on an indicator for the positive noise level and world fixed effects. Consequently, the coefficient measures the average change in payoff while holding preferences and initial ownership fixed. Table A.1 presents these regressions and reproduces the estimates in the main welfare table.

We cluster standard errors by world using the CR2 small-sample correction. Each comparison uses $N - 1$ degrees of freedom, where N is the number of paired worlds. We allow questioning and reporting behavior to respond to noise, so we do not include these responses as control variables.

Table A.1: Regressions of market payoff on communication noise

Model / comparison	N	Coefficient	CR2 SE	Within R^2	Holm p	RW p
Gemini / 10%–0%	17	−1.29	0.741	0.160	0.2002	0.1831
Gemini / 25%–0%	17	−4.18**	1.092	0.478	0.0075	0.0050
Gemini / 35%–0%	18	−6.22**	1.312	0.569	0.0012	0.0019
Sol / 10%–0%	18	−0.33	0.695	0.013	0.6378	0.6533
Sol / 25%–0%	18	−3.22*	1.131	0.323	0.0332	0.0283
Sol / 35%–0%	18	−6.83**	1.424	0.575	0.0012	0.0019
Fable / 10%–0%	18	−4.17*	1.211	0.411	0.0125	0.0110
Fable / 25%–0%	18	−7.67***	1.146	0.725	< 0.0001	< 0.0001
Fable / 35%–0%	18	−7.89***	0.976	0.793	< 0.0001	< 0.0001

We estimate a separate regression for each row, comparing total true market payoff at the indicated noise level with payoff at zero noise in the same worlds. Each regression includes world fixed effects and one indicator for the positive noise level. It uses $2N$ market observations from N worlds and includes no other controls. Consequently, the coefficients reproduce Table 1. We cluster standard errors by world using CR2. Holm and Romano–Wolf (RW) adjustments each cover these nine comparisons. The next table examines whether the estimates change when we require complete markets at all four noise levels.

We next include the three positive-noise indicators in one regression for each model. This specification uses only worlds with complete markets at all four noise levels, allowing us to examine whether the estimates change when every comparison uses the same sample. Table A.2 reports the results. Gemini’s estimated effect at 35% noise becomes −6.06

points, compared with -6.22 in the main table. This difference arises because the complete-four-rate sample excludes world 05, where payoff fell by nine points between zero and 35% noise. The payoff measure remains unchanged. The reported within- R^2 describes how much variation the noise indicators explain after accounting for world fixed effects; the experimental design provides the basis for causal interpretation.

Table A.2: Dependent variable: Total true market payoff (six-principal utility points)

Regressor / statistic	Gemini	Sol	Fable
10% vs 0%	-1.29 (0.74)	-0.33 (0.70)	-4.17* (1.21)
Romano-Wolf p	0.1831	0.6533	0.0110
25% vs 0%	-4.18** (1.09)	-3.22* (1.13)	-7.67*** (1.15)
Romano-Wolf p	0.0050	0.0283	< 0.0001
35% vs 0%	-6.06** (1.38)	-6.83** (1.42)	-7.89*** (0.98)
Romano-Wolf p	0.0025	0.0018	< 0.0001
Mean at 0%	44.65	43.83	44.11
World fixed effects	Yes	Yes	Yes
Other controls	None	None	None
Market observations	68	72	72
Worlds (clusters)	17	18	18
Within R^2	0.356	0.408	0.484

We estimate one OLS regression for each model using worlds with complete markets at all four noise levels. The dependent variable is total true market payoff. The regressors are indicators for 10%, 25%, and 35% noise and world fixed effects, with zero noise as the reference condition. Parentheses report world-clustered CR2 standard errors. Tests use a t distribution with $n - 1$ degrees of freedom, where n is the number of worlds, under the independent normal world-contrast assumption. Stars use Holm-adjusted p -values across the nine coefficients: * $p < .05$, ** $p < .01$, and *** $p < .001$. They indicate within-model changes, rather than differences between models. Requiring complete markets at every rate excludes one additional Gemini world, changing its 35% estimate from -6.22 on eighteen pairs to -6.06 . Within- R^2 measures fit after removing world means. RW adjusts the same nine coefficients in this balanced sample, separately from the nine available-pair estimates in the main table.

A.2 Between-model noise effects

We also examine whether the same increase in communication noise produces different payoff changes across models. We estimate the following pooled specification:

$$W_{wmr} = \alpha_{wm} + \sum_{k \in \{.10, .25, .35\}} \sum_{\ell \in \{G, S, F\}} \beta_{\ell k} \mathbf{1}\{m = \ell, r = k\} + e_{wmr}.$$

Here, w identifies the world, m identifies the model, and r is the communication-noise rate. The dependent variable, W_{wmr} , is total true market payoff. The world-by-model fixed effect, α_{wm} , gives each model its own baseline within each world. The coefficient $\beta_{\ell k}$ measures the payoff change for model ℓ at noise rate k relative to zero noise; G , S , and F denote Gemini, Sol, and Fable. The indicator equals one for the specified model and noise rate, and e_{wmr} is the regression residual. We cluster standard errors by world because the models face the same underlying preferences and ownership assignments.

Table A.3 reports the pooled regression. We then subtract one model's noise-effect coefficient from another's to compare how much payoff each model loses relative to its own zero-noise baseline. Table A.4 presents all three model pairs at each positive noise level.

Table A.3: Dependent variable: total true market payoff

Noise indicator / statistic	Gemini	Sol	Fable
10%–0%	-1.29	-0.65	-4.82
CR2 SE	(0.74)	(0.66)	(1.08)
25%–0%	-4.18	-2.88	-7.59
CR2 SE	(1.09)	(1.14)	(1.21)
35%–0%	-6.06	-7.24	-7.65
CR2 SE	(1.38)	(1.45)	(1.00)
World-by-model fixed effects	Yes	Yes	Yes
Other controls	None	None	None
Market observations	204	204	204
World clusters	17	17	17
Within R^2	0.440	0.440	0.440

We estimate one pooled OLS regression using the worlds with complete markets for every model and noise level. The dependent variable is total true market payoff. The regression includes world-by-model fixed effects and separate noise indicators for each model, with zero noise as the reference condition. Each column reports that model’s coefficients from this single regression. We cluster standard errors by world using CR2 and include no controls determined after treatment. Within- R^2 measures fit after removing world-by-model means. Table A.4 compares these coefficients across models.

Table A.4: All model differences implied by the pooled welfare regression

First model minus second / rate	Difference	CR2 SE	95% CI	Holm p	RW p
Gemini–Sol / 10%–0%	–0.65	0.87	[–2.49, 1.20]	1.0000	0.8451
Gemini–Sol / 25%–0%	–1.29	1.32	[–4.10, 1.51]	1.0000	0.8451
Gemini–Sol / 35%–0%	1.18	1.32	[–1.63, 3.98]	1.0000	0.8451
Gemini–Fable / 10%–0%	3.53	1.32	[0.72, 6.33]	0.1012	0.0949
Gemini–Fable / 25%–0%	3.41	1.16	[0.95, 5.87]	0.0672	0.0569
Gemini–Fable / 35%–0%	1.59	1.68	[–1.97, 5.15]	1.0000	0.8451
Sol–Fable / 10%–0%	4.18*	1.28	[1.47, 6.88]	0.0381	0.0286
Sol–Fable / 25%–0%	4.71**	1.03	[2.52, 6.89]	0.0029	0.0023
Sol–Fable / 35%–0%	0.41	1.59	[–2.96, 3.78]	1.0000	0.8451

We subtract the second model’s noise effect from the first model’s effect in Table A.3. A positive estimate means that the first model loses less payoff, or gains more, relative to its own zero-noise condition. These comparisons use the same 204 observations and seventeen world clusters as the pooled regression, so they do not have a separate R^2 . We report world-clustered CR2 standard errors, two-sided tests using t_{16} , and individual 95% confidence intervals. Holm and RW each adjust across the nine comparisons, and stars use the Holm-adjusted values. The main 35% comparison separately uses all eighteen available paired worlds.

The pooled regression includes seventeen worlds with complete markets for every model and noise level. In Table A.5, we instead retain every world with the four markets needed for a particular comparison: both models at zero noise and at the positive noise level. The Sol–Fable difference at 25% remains positive under both sample definitions. Its significance at 10% depends on the sample, while the 35% comparison remains statistically inconclusive. These results suggest that the timing of payoff losses differs across models as communication becomes noisier. The evidence does not support a model ranking that applies at every noise level.

Table A.5: Model comparisons using every available paired world

First minus second / rate	N	Difference	95% CI	Holm p	RW p
Gemini–Sol / 10%–0%	17	−0.65	[−2.49, 1.20]	1.0000	0.8407
Gemini–Sol / 25%–0%	17	−1.29	[−4.10, 1.51]	1.0000	0.8154
Gemini–Sol / 35%–0%	18	0.61	[−2.28, 3.50]	1.0000	0.8407
Gemini–Fable / 10%–0%	17	3.53	[0.72, 6.33]	0.1012	0.1007
Gemini–Fable / 25%–0%	17	3.41	[0.95, 5.87]	0.0672	0.0633
Gemini–Fable / 35%–0%	18	1.67	[−1.68, 5.01]	1.0000	0.8076
Sol–Fable / 10%–0%	18	3.83	[1.20, 6.47]	0.0560	0.0515
Sol–Fable / 25%–0%	18	4.44**	[2.32, 6.57]	0.0034	0.0029
Sol–Fable / 35%–0%	18	1.06	[−2.39, 4.50]	1.0000	0.8407

For each model pair and noise level, we retain every world with both models’ markets complete at zero noise and at that positive noise level. We first calculate each model’s payoff change within the world and then subtract the second model’s change from the first. The estimate is the average of these differences, measured in total payoff points. We report individual confidence intervals using t_{N-1} , where N is the number of paired worlds. Holm and RW each adjust across these nine comparisons; RW preserves the dependence among observations from the same world. Because the available worlds differ across rows, these estimates do not come from one common-sample pooled regression and do not inherit its R^2 .

B Regressions supporting learning and representation results

B.1 Differences between reported and Bayesian values

We next examine whether communication noise changes the distance between Agents’ reports and Bayesian value estimates. At each noise level, we calculate the absolute difference between each reported object value and the Bayesian estimate constructed from that same conversation. We average these differences across the six objects and six Principals within each world. This average discrepancy is the dependent variable in Table B.1.

For each model, we regress the discrepancy on the three positive-noise indicators and world fixed effects. A positive coefficient means that reports move farther from their corresponding Bayesian estimates as noise rises from zero to the indicated level. This outcome measures numerical disagreement in reported values. Its economic consequences depend on whether that disagreement changes rankings and allocations, which we examine separately in the main text.

Table B.1: Dependent variable: Absolute AI-minus-Bayesian value discrepancy (points per object)

Regressor / statistic	Gemini	Sol	Fable
10% vs 0%	0.131*** (0.016)	-0.047 (0.029)	0.010 (0.018)
80% MDE	0.066	0.119	0.075
Romano-Wolf p	< 0.0001	0.3566	0.6009
25% vs 0%	0.303*** (0.032)	-0.062 (0.028)	-0.021 (0.016)
80% MDE	0.131	0.112	0.065
Romano-Wolf p	< 0.0001	0.1434	0.4819
35% vs 0%	0.487*** (0.031)	-0.042 (0.035)	-0.038 (0.015)
80% MDE	0.129	0.143	0.061
Romano-Wolf p	< 0.0001	0.4819	0.0920
Mean at 0%	0.007	0.123	0.045
World fixed effects	Yes	Yes	Yes
Other controls	None	None	None
Market observations	68	72	72
Worlds (clusters)	17	18	18
Within R^2	0.820	0.102	0.190
Linearity test: adjusted p	0.8951	0.2961	0.8951

We measure the absolute difference between each Agent’s reported object value and its Bayesian estimate based on the same conversation, then average across objects and Principals. Each column regresses this discrepancy on indicators for 10%, 25%, and 35% noise and world fixed effects. Zero noise is the reference condition. We use worlds complete at all four rates within each model. Parentheses report world-clustered CR2 standard errors; tests use t_{n-1} under independent normal world contrasts. Stars use Holm-adjusted p -values across the nine changes: $*p < .05$, $**p < .01$, and $***p < .001$. RW adjusts the same nine coefficients. Within- R^2 measures fit after removing world means. The MDE uses $\alpha = .05/9$, the estimated standard deviation of world differences, and a normal working model. It describes detectable changes in discrepancy points per object, rather than an equivalence margin. The linearity test asks whether changes are proportional to the actual noise-rate spacing, with a separate Holm adjustment across three models. Failure to reject leaves linearity unresolved. Appendix C.1 explains the inference methods.

B.2 Questioning, favorite identification and individual harm

We also examine how communication noise affects Agents’ questioning, their ability to identify a Principal’s favorite object, and the share of Principals harmed by trade. The dependent variables in the next three tables are, respectively, average follow-up questions per Principal, the percentage of Principals whose Agent identifies a truly best object, and the percentage harmed by the final allocation. A favorite is correctly identified if it has the highest true value, allowing for ties. Harm occurs when the assigned object has lower true value than the Principal’s initial object.

For each model and outcome, we compare each positive noise rate with zero noise using world fixed effects. Each coefficient therefore measures a change within the same underlying world. Holm and Romano-Wolf adjustments cover thirty-six comparisons across four outcomes, three models, and three positive noise levels. The fourth outcome, error relative to true object values, is reported in the supplement and remains included in these adjustments.

Table B.2: Dependent variable: Optional follow-up questions (questions per Principal)

Model / rate	N	Coefficient	CR2 SE	95% CI	Holm p	RW p	Within R^2
Gemini / 10%–0%	17	0.11*	0.028	[0.05, 0.17]	0.0126	0.0153	0.475
Gemini / 25%–0%	17	0.10*	0.029	[0.04, 0.16]	0.0194	0.0250	0.420
Gemini / 35%–0%	18	0.12*	0.030	[0.06, 0.18]	0.0114	0.0137	0.494
Sol / 10%–0%	18	0.19*	0.047	[0.09, 0.29]	0.0114	0.0137	0.500
Sol / 25%–0%	18	0.19*	0.047	[0.09, 0.29]	0.0114	0.0137	0.500
Sol / 35%–0%	18	0.19*	0.047	[0.09, 0.29]	0.0114	0.0137	0.500
Fable / 10%–0%	18	-1.38***	0.106	[-1.60, -1.16]	< 0.0001	< 0.0001	0.909
Fable / 25%–0%	18	-1.59***	0.116	[-1.84, -1.35]	< 0.0001	< 0.0001	0.917
Fable / 35%–0%	18	-1.80***	0.129	[-2.07, -1.52]	< 0.0001	< 0.0001	0.919

The dependent variable is the average number of follow-up questions per Principal, excluding the opening answer. A positive coefficient means that Agents ask more questions as noise increases. Each row compares the indicated positive noise level with zero noise using an OLS regression with world fixed effects and one noise indicator. The coefficient equals the average within-world change. The regression uses $2N$ market observations from N worlds and includes no other controls. We report CR2 standard errors and individual 95% confidence intervals using t_{N-1} . Holm and RW each adjust across the same thirty-six comparisons covering four outcomes. Stars use Holm-adjusted values. Appendix C.1 explains the RW procedure, which resamples world-level contrasts under the null hypothesis.

Table B.3: Dependent variable: Correct identification of a true favorite (percentage points)

Model / rate	N	Coefficient	CR2 SE	95% CI	Holm p	RW p	Within R^2
Gemini / 10%–0%	17	−10.78*	3.177	[−17.52, −4.05]	0.0194	0.0250	0.419
Gemini / 25%–0%	17	−34.31***	3.635	[−42.02, −26.61]	< 0.0001	< 0.0001	0.848
Gemini / 35%–0%	18	−54.63***	5.524	[−66.28, −42.98]	< 0.0001	< 0.0001	0.852
Sol / 10%–0%	18	−13.89*	3.369	[−21.00, −6.78]	0.0114	0.0137	0.500
Sol / 25%–0%	18	−36.11***	4.519	[−45.65, −26.58]	< 0.0001	0.0001	0.790
Sol / 35%–0%	18	−53.70***	5.646	[−65.61, −41.79]	< 0.0001	< 0.0001	0.842
Fable / 10%–0%	18	−30.56***	3.369	[−37.66, −23.45]	< 0.0001	< 0.0001	0.829
Fable / 25%–0%	18	−52.78***	6.211	[−65.88, −39.67]	< 0.0001	< 0.0001	0.809
Fable / 35%–0%	18	−59.26***	4.711	[−69.20, −49.32]	< 0.0001	< 0.0001	0.903

The dependent variable is the percentage of Principals whose Agent identifies a truly most-preferred object, allowing for ties in true values. A negative coefficient means that fewer Agents identify a true favorite as noise increases. Each row compares the indicated positive noise level with zero noise using an OLS regression with world fixed effects and one noise indicator. The coefficient equals the average within-world change. The regression uses $2N$ market observations from N worlds and includes no other controls. We report CR2 standard errors and individual 95% confidence intervals using t_{N-1} . Holm and RW each adjust across the same thirty-six comparisons covering four outcomes. Stars use Holm-adjusted values. Appendix C.1 explains the RW procedure, which resamples world-level contrasts under the null hypothesis.

Table B.4: Dependent variable: Principals worse off than initial ownership (percentage points)

Model / rate	N	Coefficient	CR2 SE	95% CI	Holm p	RW p	Within R^2
Gemini / 10%–0%	17	4.90*	1.899	[0.88, 8.93]	0.0401	0.0368	0.294
Gemini / 25%–0%	17	10.78*	3.765	[2.80, 18.77]	0.0337	0.0320	0.339
Gemini / 35%–0%	18	13.89*	3.628	[6.23, 21.54]	0.0121	0.0153	0.463
Sol / 10%–0%	18	4.63	2.257	[−0.13, 9.39]	0.0560	0.0495	0.198
Sol / 25%–0%	18	10.19*	2.741	[4.40, 15.97]	0.0126	0.0180	0.448
Sol / 35%–0%	18	16.67*	4.261	[7.68, 25.66]	0.0114	0.0153	0.474
Fable / 10%–0%	18	13.89**	3.087	[7.38, 20.40]	0.0054	0.0092	0.543
Fable / 25%–0%	18	15.74*	3.922	[7.47, 24.02]	0.0114	0.0143	0.487
Fable / 35%–0%	18	12.04*	3.515	[4.62, 19.45]	0.0194	0.0250	0.408

The dependent variable is the percentage of Principals assigned an object worth less than their initial endowment. A positive coefficient means that more Principals are harmed as noise increases. Each row compares the indicated positive noise level with zero noise using an OLS regression with world fixed effects and one noise indicator. The coefficient equals the average within-world change. The regression uses $2N$ market observations from N worlds and includes no other controls. We report CR2 standard errors and individual 95% confidence intervals using t_{N-1} . Holm and RW each adjust across the same thirty-six comparisons covering four outcomes. Stars use Holm-adjusted values. Appendix C.1 explains the RW procedure, which resamples world-level contrasts under the null hypothesis.

B.3 Does reporting add to the noise loss?

We next ask whether increasing noise produces a larger welfare loss under Agents’ actual reports than under Bayesian reports based on the same conversations. For each world, model, and noise rate, we calculate $G_{wmr} = W_{wmr}^{AI} - W_{wmr}^B$, where W_{wmr}^{AI} is total payoff under Agents’ reports and W_{wmr}^B is total payoff under the corresponding Bayesian reports. Thus, G_{wmr} measures how replacing Bayesian reports with Agents’ reports changes market payoff at that noise level.

Table B.5 regresses this payoff difference on an indicator for 35% noise and world fixed effects, using the zero- and 35%-noise markets. The coefficient measures $G_{wm,35} - G_{wm,0}$. A negative coefficient means that payoff falls more under Agents’ reports than under Bayesian reports as noise increases; a positive coefficient means that it falls less. The confidence intervals include both additional losses and improvements. Consequently, we do not detect a systematic additional loss associated with reporting discrepancies, although the estimates leave uncertainty about its size. Table B.6 provides the accompanying counts of ranking, allocation, and payoff differences between the two reporting rules.

Table B.5: Change in the AI-minus-Bayesian allocation welfare gap

Model	Worlds	Change	CR2 SE	95% interval	Holm p
Gemini	18	-0.78	0.74	[-2.35, 0.79]	0.9303
Sol	18	0.33	0.66	[-1.06, 1.73]	0.9303
Fable	18	-0.33	0.33	[-1.04, 0.37]	0.9303

At each noise level, we subtract market payoff under Bayesian reports from payoff under Agents’ reports, using the same conversations for both. We then compare this payoff gap at 35% noise with its value at zero noise within each world. The coefficient comes from a regression of the gap on a 35%-noise indicator and world fixed effects. A negative estimate means that reporting discrepancies add to the payoff decline as noise increases. Holm adjusts across the three model comparisons. We also examine sign-flip inference because many world-level changes are zero. Holding the conversations fixed allows us to compare reporting rules, but does not separately identify the causal effects of asking questions and updating beliefs.

Table B.6: How reporting differences affect rankings, assignments, and payoff

Model	Noise	Markets	Same rank	Rank only	Offset	Higher	Lower
Fable	0%	18	10	7	0	1	0
Fable	10%	19	7	12	0	0	0
Fable	25%	18	9	8	0	0	1
Fable	35%	19	15	3	1	0	0
Gemini	0%	18	17	1	0	0	0
Gemini	10%	17	8	6	0	2	1
Gemini	25%	17	2	11	0	2	2
Gemini	35%	18	0	13	1	2	2
Sol	0%	18	10	5	1	1	1
Sol	10%	18	10	7	0	1	0
Sol	25%	18	13	5	0	0	0
Sol	35%	18	6	7	1	3	1

We rerun each complete market using Agents’ reports and Bayesian estimates from the same conversations. “Same rank” means that complete rankings and final assignments coincide. “Rank only” means that some rankings differ but final assignments remain unchanged, even if temporary TTC pointers differ. “Offset” means that assignments change but individual gains and losses sum to zero. “Higher” and “Lower” indicate whether total true payoff is higher or lower under Agents’ reports. These are counts of complete markets. We account for dependence across a world’s markets when reporting assignment-change uncertainty in the online supplement.

C Robustness and precision for the main findings

C.1 Dependence and multiple testing

Markets within a world share preferences and initial ownership. Their outcomes are therefore related, and TTC also connects the six Principals’ assignments within each market. We therefore use the world as the sampling unit when calculating uncertainty. Holm-adjusted p -values account for multiple comparisons within each stated testing family. The reported 95% confidence intervals apply to individual coefficients. Since those intervals are not adjusted jointly, an interval can exclude zero even when the corresponding Holm-adjusted test is not statistically significant.

We also examine whether our conclusions depend on the method used to calculate statistical significance. The Romano–Wolf procedure uses 49,999 wild-bootstrap draws centered under the null hypothesis. Within each draw, every model and outcome from the same world receives the same sign weight. We then recalculate each contrast divided by its standard error. This preserves the dependence across observations from the same world while adjusting for multiple tests. We use this procedure to assess whether the inference changes with a small number of worlds. Its validity relies on statistical assumptions and does not follow from randomized reassignment of the observed treatments.

The exact sign-flip tests enumerate sign reversals of the world-level differences and require those differences to be independent and symmetrically distributed under the null. The whole-world bootstrap repeatedly draws worlds from the sample while keeping each world’s observations together. These methods rely on different assumptions, but the main welfare conclusions are consistent across the reported checks.

Table C.1: Uncertainty sensitivity for welfare regressions

Coefficient	Rate	CR2 p	Sign-flip p	World-bootstrap CI	Wild- t p	Holm CR2 p
Gemini (primary)	35%	0.0002	< 0.0001	[-8.78, -3.78]	0.0004	0.0003
Gemini	10%	0.1001	0.1191	[-2.71, 0.12]	0.1011	0.2002
Gemini	25%	0.0015	0.0016	[-6.29, -2.12]	0.0010	0.0075
Gemini	35%	0.0005	0.0002	[-8.76, -3.53]	0.0005	0.0028
Sol (primary)	35%	0.0002	0.0003	[-9.61, -4.11]	0.0007	0.0003
Sol	10%	0.6378	0.6992	[-1.67, 1.00]	0.6665	0.6378
Sol	25%	0.0111	0.0031	[-5.61, -1.33]	0.0046	0.0332
Sol	35%	0.0002	0.0003	[-9.56, -4.11]	0.0010	0.0012
Fable (primary)	35%	< 0.0001	< 0.0001	[-9.72, -6.06]	< 0.0001	< 0.0001
Fable	10%	0.0031	0.0048	[-6.39, -1.78]	0.0038	0.0125
Fable	25%	< 0.0001	< 0.0001	[-9.89, -5.50]	< 0.0001	< 0.0001
Fable	35%	< 0.0001	< 0.0001	[-9.78, -6.00]	< 0.0001	< 0.0001
Gemini effect (pooled)	10%	0.1001	0.1191	[-2.76, 0.12]	0.0978	0.5004
Gemini effect (pooled)	25%	0.0015	0.0016	[-6.29, -2.12]	0.0009	0.0120
Gemini effect (pooled)	35%	0.0005	0.0002	[-8.71, -3.53]	0.0006	0.0041
Sol minus Gemini (pooled)	10%	0.4678	0.5156	[-1.06, 2.24]	0.4845	1.0000
Sol minus Gemini (pooled)	25%	0.3423	0.3604	[-1.18, 3.82]	0.3345	1.0000
Sol minus Gemini (pooled)	35%	0.3872	0.4109	[-3.59, 1.53]	0.3901	1.0000
Fable minus Gemini (pooled)	10%	0.0169	0.0206	[-6.06, -1.00]	0.0188	0.1012
Fable minus Gemini (pooled)	25%	0.0096	0.0104	[-5.59, -1.24]	0.0077	0.0672
Fable minus Gemini (pooled)	35%	0.3587	0.3736	[-4.65, 1.71]	0.3554	1.0000

We retain each world’s complete contrast when resampling. Sign-flip tests assume that these contrasts are symmetric under the null. The percentile intervals use 10,000 whole-world resamples. The wild- t tests use 10,000 draws with independent positive or negative world weights under the null and a correction for the finite number of simulations. The alternative p -values and intervals shown here are unadjusted, whereas the final column reports Holm-adjusted CR2 values. Consequently, the columns do not apply the same multiple-testing threshold. The accompanying `v4_uncertainty_sensitivity.csv` records adjusted alternative values, seeds, and samples. Small differences between simulation runs reflect Monte Carlo variation.

Table C.2: How sensitive are the conclusions to a broader testing inventory?

Outcome / model	Original-family p	Holm-15 p	Holm-60 p
Welfare / Gemini	0.0003	0.0012	0.0091
Welfare / Sol	0.0003	0.0012	0.0084
Welfare / Fable	< 0.0001	< 0.0001	< 0.0001
Harm share / Gemini	0.0036	0.0036	0.0579
Harm share / Sol	0.0036	0.0036	0.0494
Harm share / Fable	0.0036	0.0036	0.1293

We examine whether the main findings remain statistically significant when we adjust across broader groups of tests. The underlying tests are two-sided paired-world t tests. Holm-15 combines the three welfare comparisons with twelve secondary comparisons. Holm-60 uses the full inventory recorded in the replication files. These groups overlap and reuse observations. The main tables retain their original testing families; this table provides an additional check using broader adjustments. The accompanying CSV lists all fifteen core comparisons and all sixty inventory entries.

We further examine whether the conclusions change when we adjust across broader groups of fifteen and sixty tests from the earlier analysis. These groups do not include every subsequent extension. The new question-policy and harm analyses use their own testing families, together with a combined forty-eight-test sensitivity. These adjustments address multiple testing within the specified groups. Our findings remain exploratory because some analyses were selected, and collection was continued, after interim results had been examined.

C.2 Missing markets and sample composition

The main comparisons include worlds in which both relevant markets were completed successfully. In Table C.3, we examine whether the estimates change when we restrict the sample to common worlds, exclude conversations requiring repairs, or omit one world at a time. Table C.4 instead assigns the lowest and highest feasible payoff values to missing allocations. We retain failed conversations in the collection record and do not replace them. These analyses show how

sensitive the findings are to sample composition and missing outcomes; they do not establish that collection failures are unrelated to market outcomes.

Table C.3: Welfare-effect sensitivity to sample composition

Model	Shared pairs(n)	All balanced(n)	Repair-free(n)	Leave-one-out range
Gemini	-6.22 (18)	-6.06 (17)	-6.22 (18)	[-6.71,-5.59]
Sol	-6.83 (18)	-7.24 (17)	-6.83 (18)	[-7.41,-6.18]
Fable	-7.89 (18)	-7.65 (17)	-7.89 (18)	[-8.29,-7.47]

We compare the estimated 35%-versus-zero payoff change under alternative sample restrictions. “All balanced” requires complete markets for every model and noise level within a world. “Repair-free” excludes conversations requiring a repair. For the final column, we omit each usable world in turn and recalculate the estimate. The reported range shows how much the estimate changes under these omissions; it is not a confidence interval.

Table C.4: How much can incomplete markets change the observed 35%-minus-0% contrast?

Model	Touched worlds	Lower feasible mean	Upper feasible mean
Fable	19	-8.32	-7.11
Gemini	19	-7.11	-4.68
Sol	19	-7.68	-5.26

We hold completed market outcomes fixed and allow each incomplete market to take any feasible assignment. The known true values determine the lowest and highest possible payoff for each missing allocation. We use these values to bound the average 35%-versus-zero payoff change across worlds where collection had begun. The bounds do not cover the uncollected remainder of the twenty-four planned worlds. They describe uncertainty due to missing allocations in this sample, rather than sampling confidence intervals, and require no claim that the missing outcomes are random.

C.3 Collection dates

Data collection took place in six operating sessions across September 30 and October 2. Two markets contain conversations from both dates. Table C.5 reports estimates excluding pairs that contain mixed-day markets and separately compares the two collection cohorts. Most worlds were collected on only one date, so world fixed effects generally absorb collection-date indicators. Consequently, differences between the cohorts may reflect differences in underlying worlds as well as collection dates. These comparisons cannot establish whether the hosted models changed between sessions.

Table C.5: Collection-cohort sensitivity and its identification limit

Model	Pure-day pairs	Effect	95% interval	Cohort Holm p
Gemini	17	-5.59	[-8.17, -3.00]	0.9665
Sol	18	-6.83	[-9.84, -3.83]	0.7244
Fable	18	-7.89	[-9.95, -5.83]	0.2489

The effect column excludes any pair containing a market whose conversations span both collection dates. The cohort test separately compares payoff changes in September-only worlds with those in October-only worlds. We adjust its p -values across the three models using Holm. These two calculations examine different aspects of sample composition. Since collection dates and worlds largely coincide, neither can establish that the models improved between dates.

C.4 Power and the size of an unresolved effect

We report minimum detectable effects to assess which payoff changes our sample can reliably distinguish from zero. In Table C.6, the minimum detectable effect is the true change in total market payoff that would be detected with 80% probability under the stated assumptions. The calculation uses the observed standard deviation of paired-world differences, a normal-distribution assumption for those differences, and the conservative significance level $\alpha = .05/9$ for nine comparisons.

The power column instead evaluates a specified three-point effect. It does not use the estimated effect as if it were the true population effect. Both calculations use an estimated variance that is itself uncertain in a small sample. They therefore show which effects the sample could detect under the stated assumptions. The planned boundary remains twenty-four worlds; each estimate uses the paired worlds actually available.

Table C.6: Precision for the expanded nine-comparison welfare family

Model / rate	Worlds	80% MDE	Power for 3 points
Gemini / 10%–0%	17	3.05	78%
Gemini / 25%–0%	17	4.50	36%
Gemini / 35%–0%	18	5.36	23%
Sol / 10%–0%	18	2.84	85%
Sol / 25%–0%	18	4.62	34%
Sol / 35%–0%	18	5.81	18%
Fable / 10%–0%	18	4.94	28%
Fable / 25%–0%	18	4.68	33%
Fable / 35%–0%	18	3.99	48%

We calculate the minimum detectable effect (MDE) using a two-sided noncentral- t test, the observed standard deviation of paired-world differences, and the conservative significance level $\alpha = .05/9$. The MDE is the change in total true market payoff that this calculation detects with 80% probability. The power column instead evaluates a specified three-point change, rather than the estimated coefficient. Both quantities depend on the estimated variance and the normal working model. They do not account for the decision to continue collection after examining interim results or guarantee the precision of a larger sample.

D Supporting analysis for the two offline extensions

D.1 How the two policies select questions

We construct two adaptive Bayesian question policies with different objectives. The value-learning policy minimizes expected squared error in the six final object value estimates. The market-payoff policy maximizes the represented Principal’s expected payoff from the object assigned through TTC. Both policies update their beliefs after each answer and account for possible future answers when selecting the next question.

To construct the market-payoff policy, we hold initial ownership and the other five Bayesian reports fixed. These reports determine which objects Principal i can obtain by changing its own reported ranking. We denote this set by M_i . For any strict ranking, TTC assigns Principal i its highest-ranked object within that set. We verify this relationship using a separate TTC implementation for all $6! = 720$ rankings in each of the 1,296 individual contexts. When posterior values are equal, we use the original alphabetical tie-breaking rule.

Let p denote the posterior probabilities over the eighty-one preference types, $u_\theta(j)$ the value of object j for type θ , and h the number of questions remaining. If no questions remain, the policy selects the obtainable object with the highest expected value:

$$V_i(p, 0) = \max_{j \in M_i} \sum_{\theta} p(\theta) u_\theta(j).$$

The expression $V_i(p, 0)$ is Principal i ’s expected payoff at submission. If questions remain, the policy compares all fifteen object pairs:

$$V_i(p, h) = \max_{q \in \mathcal{Q}} \sum_{a \in q} \Pr(a \mid p, q) V_i(p^{q,a}, h - 1).$$

Here, $\mathcal{Q}$ is the set of possible pairs, q is a selected pair, and a is the answer naming one object in that pair. The probability $\Pr(a \mid p, q)$ follows from the current beliefs and the known answer-error process. The term $p^{q,a}$ denotes the updated beliefs after receiving that answer. For each question, we calculate expected payoff across its two possible answers and the subsequent questions. The policy selects the pair with the highest expected payoff.

The policy does not observe the Principal’s true preference type or future answer draws. The simulator uses the true type to generate answers and evaluate the final allocation. At the same answer position, the original conversation and both replays use the same uniform random draw. Consequently, if the second answer is reversed in the original conversation, it is also reversed in a replay, even if the replay asks about a different pair. Each policy still faces the stated answer-error probability. Sharing the draws makes the comparisons dependent and does not create additional independent observations.

We solve the finite question sequence by storing previously calculated belief-state values and reusing them whenever that state is reached again. Calculations use double-precision arithmetic, with alphabetical tie-breaking when values differ by no more than 10^{-11} . Independent exact-rational calculations check the recursion for short remaining question sequences.

The market-payoff policy receives the other five reports and each Agent’s realized question count. These inputs give it information that the original Agent did not have. In particular, that question count was chosen after answers

in the original conversation, which may differ from the answers received in the replay. Consequently, the expected gains describe the benchmark under these information and question-budget assumptions. They need not be unbiased conditional on the Agent’s original stopping decision. We compare the two policies within these restrictions. The calculation does not establish when Agents should voluntarily stop, how all six Agents should jointly maximize market payoff, or how an instruction to follow either policy would affect hosted-model behavior.

D.2 Regression form of the question-policy comparisons

For each model and noise rate, we calculate how much the represented Principal’s payoff changes when we replace the value-learning policy with the market-payoff policy. We make this replacement separately for each of the six Principals and average their payoff differences within the world. Let D_w denote this average for world w . We estimate $D_w = \beta + e_w$, where β is the mean payoff difference across worlds and e_w is the residual.

The same coefficient can be obtained by regressing the two policy outcomes on a policy indicator and world fixed effects. Its CR2 standard error equals $s_D/\sqrt{N}$ in this balanced comparison, where s_D is the sample standard deviation of the world-level differences and N is the number of worlds. Table D.1 reports the estimates, standard errors, confidence intervals, and multiple-testing adjustments. We omit R^2 for the intercept-only regression because it provides no useful measure of explanatory fit. We do not control for questioning or reporting choices made after treatment.

Table D.1: Dependent variable: Principal payoff change under the market-payoff policy relative to the value-learning policy

Model / noise	Worlds	Coefficient	SE	95% CI	Holm p	Flip Holm p
Gemini / 0%	18	0.028	0.028	[-0.03, 0.09]	1.0000	1.0000
Gemini / 10%	17	0.294	0.115	[0.05, 0.54]	0.2298	0.2344
Gemini / 25%	17	0.098	0.076	[-0.06, 0.26]	1.0000	1.0000
Gemini / 35%	18	0.056	0.147	[-0.25, 0.37]	1.0000	1.0000
Sol / 0%	18	0.074	0.041	[-0.01, 0.16]	0.8796	1.0000
Sol / 10%	18	0.167	0.108	[-0.06, 0.39]	1.0000	1.0000
Sol / 25%	18	0.093	0.182	[-0.29, 0.48]	1.0000	1.0000
Sol / 35%	18	0.009	0.173	[-0.36, 0.37]	1.0000	1.0000
Fable / 0%	18	0.102	0.038	[0.02, 0.18]	0.2022	0.3438
Fable / 10%	19	-0.061	0.113	[-0.30, 0.18]	1.0000	1.0000
Fable / 25%	18	-0.028	0.070	[-0.18, 0.12]	1.0000	1.0000
Fable / 35%	19	0.044	0.123	[-0.22, 0.30]	1.0000	1.0000

We compare the represented Principal’s realized payoff under the market-payoff question policy with payoff under the value-learning policy. A positive coefficient means that targeting the allocation produces a higher payoff. We average the six Principal-level differences within each world and estimate $D_w = \beta + e_w$, where D_w is that world average. Equivalently, we regress the two outcomes on world fixed effects and one policy or reporting-rule indicator, using $2N$ observations and N world clusters. The standard error is $s_D/\sqrt{N}$; the balanced CR2 calculation uses $N - 1$ degrees of freedom. We include no controls determined after treatment. Holm adjusts across the twelve rows for this outcome. Exact sign-flip tests enumerate the world-level sign assignments under a symmetry assumption. An all-zero sample receives $p = 1$ and a zero-width descriptive interval, which does not establish population equivalence. We omit R^2 because the intercept-only regression provides no informative measure of fit.

Table D.2 compares the market-payoff policy with Bayesian reports constructed from the Agents’ original conversations. Both sides use posterior value estimates, removing the Agents’ numerical reporting discrepancies from this comparison. However, the comparison changes the questions and their objective while also giving the evaluator access to the other five reports. It therefore evaluates the full benchmark replacement rather than identifying the causal effect of question choice alone. Table D.3 reports the corresponding precision and sample-sensitivity results.

Table D.2: Dependent variable: Principal payoff change under the market-payoff policy relative to Bayesian reports from the original conversation

Model / noise	Worlds	Coefficient	SE	95% CI	Holm p	Flip Holm p
Gemini / 0%	18	0.037	0.022	[-0.01, 0.08]	1.0000	1.0000
Gemini / 10%	17	0.118	0.124	[-0.14, 0.38]	1.0000	1.0000
Gemini / 25%	17	0.157	0.121	[-0.10, 0.41]	1.0000	1.0000
Gemini / 35%	18	-0.056	0.164	[-0.40, 0.29]	1.0000	1.0000
Sol / 0%	18	0.037	0.022	[-0.01, 0.08]	1.0000	1.0000
Sol / 10%	18	0.065	0.082	[-0.11, 0.24]	1.0000	1.0000
Sol / 25%	18	0.000	0.144	[-0.30, 0.30]	1.0000	1.0000
Sol / 35%	18	0.056	0.151	[-0.26, 0.37]	1.0000	1.0000
Fable / 0%	18	0.102	0.057	[-0.02, 0.22]	1.0000	1.0000
Fable / 10%	19	-0.009	0.088	[-0.19, 0.18]	1.0000	1.0000
Fable / 25%	18	0.213	0.118	[-0.04, 0.46]	1.0000	1.0000
Fable / 35%	19	-0.026	0.116	[-0.27, 0.22]	1.0000	1.0000

We compare the represented Principal’s realized payoff under the market-payoff policy with payoff under Bayesian reports based on the Agent’s original conversation. This comparison changes question selection and gives the policy access to the other five reports. A positive coefficient means that this benchmark replacement produces a higher payoff. We average the six Principal-level differences within each world and estimate $D_w = \beta + e_w$, where D_w is that world average. Equivalently, we regress the two outcomes on world fixed effects and one policy or reporting-rule indicator, using $2N$ observations and N world clusters. The standard error is $s_D/\sqrt{N}$; the balanced CR2 calculation uses $N - 1$ degrees of freedom. We include no controls determined after treatment. Holm adjusts across the twelve rows for this outcome. Exact sign-flip tests enumerate the world-level sign assignments under a symmetry assumption. An all-zero sample receives $p = 1$ and a zero-width descriptive interval, which does not establish population equivalence. We omit R^2 because the intercept-only regression provides no informative measure of fit.

Table D.3: Question-policy precision and sample sensitivity

Model / noise	Expected gain	80% MDE	Power: 0.5	Bootstrap CI	17-world gain	Holm p
Gemini / 0%	0.006	0.12	100%	[0.00, 0.08]	0.03	1.0000
Gemini / 10%	0.167	0.49	82%	[0.08, 0.52]	0.29	0.2507
Gemini / 25%	0.200	0.32	100%	[-0.04, 0.25]	0.10	1.0000
Gemini / 35%	0.150	0.62	55%	[-0.21, 0.34]	0.10	1.0000
Sol / 0%	0.010	0.17	100%	[0.00, 0.16]	0.08	0.8792
Sol / 10%	0.164	0.46	89%	[-0.03, 0.38]	0.20	0.8792
Sol / 25%	0.179	0.77	33%	[-0.29, 0.42]	0.09	1.0000
Sol / 35%	0.176	0.73	38%	[-0.33, 0.33]	0.06	1.0000
Fable / 0%	0.105	0.16	100%	[0.04, 0.18]	0.10	0.3078
Fable / 10%	0.140	0.47	85%	[-0.28, 0.15]	-0.03	1.0000
Fable / 25%	0.081	0.30	100%	[-0.17, 0.10]	-0.01	1.0000
Fable / 35%	0.046	0.52	76%	[-0.19, 0.28]	0.05	1.0000

Expected gain is the average improvement predicted by the market-payoff policy relative to the value-learning policy. It is smaller than the 80% MDE in every row. This expected gain describes the specified model; it is neither a confidence bound nor a bound for every Principal. We calculate MDEs and power using the observed world-level standard deviation and a conservative two-sided significance level of $\alpha = .05/12$. Power evaluates a specified 0.5-point improvement in the represented Principal’s payoff. Bootstrap intervals use 10,000 resamples of whole worlds. The last two columns restrict the analysis to seventeen worlds complete across all models and noise levels. Very small sample variance can make estimated power optimistic, so these calculations do not establish that small effects are absent. The accompanying CSV also reports harm-share MDEs for a five-percentage-point effect and a Holm adjustment across all forty-eight new comparisons. No comparison is significant under that broader adjustment.

D.3 Who is harmed, and by how much?

We compare the true value of each Principal’s assigned object with the value of their initial object under both reporting rules. Table D.4 separates four outcomes. “Both” means the Principal loses under both AI and Bayesian reports. “AI only” means a loss occurs only under AI reports, while “Bayes only” means a loss occurs only under Bayesian reports. “Neither” means the Principal is not harmed under either rule. Receiving an object of equal value counts as no harm.

This comparison identifies whether changing the reports changes who bears a loss. It does not identify the cause of the original communication error. Harm can also occur at zero noise because a few correct pairwise answers may leave uncertainty about the ranking of the objects a Principal could receive.

Table D.4: Harm classification and total loss severity by noise rate

Model / noise	Allocations	Both	AI only	Bayes only	Neither	AI loss	Bayes loss
Gemini / 0%	108	1	0	0	107	1	1
Gemini / 10%	102	6	0	1	95	18	20
Gemini / 25%	102	11	1	3	87	38	45
Gemini / 35%	108	15	1	2	90	53	55
Sol / 0%	108	0	1	0	107	1	0
Sol / 10%	108	5	1	0	102	14	13
Sol / 25%	108	12	0	0	96	29	29
Sol / 35%	108	17	2	1	88	66	57
Fable / 0%	108	2	0	0	106	5	5
Fable / 10%	114	17	0	0	97	53	53
Fable / 25%	108	18	1	0	89	63	61
Fable / 35%	114	15	0	1	98	46	49

We classify each Principal according to whether the final object is worth less than the initial endowment under both reporting rules, only Agents' reports, only Bayesian reports, or neither. The loss columns sum $\max\{0, u_i(\text{owned}) - u_i(\text{assigned})\}$ across the displayed allocations. Consequently, another Principal's gain does not cancel a loss. At noise levels of 0%, 10%, 25%, and 35%, Gemini has 18, 17, 17, and 18 complete markets; Sol has eighteen at each level; and Fable has 18, 19, 18, and 19. Fable's additional markets in world 19 have no zero-noise counterpart. They enter these within-market reporting comparisons but not the main paired noise-effect estimates. Tables D.5 and D.6 report uncertainty using world-level differences.

We next examine whether the reporting rule changes the percentage of Principals harmed. For each Principal, we record whether harm occurs under AI reports and under Bayesian reports, using one for harm and zero otherwise. We subtract the Bayesian indicator from the AI indicator, average across the six Principals in each world, and multiply by one hundred. The resulting measure is the difference in harmed shares, expressed in percentage points. It also equals the AI-only share minus the Bayes-only share. Table D.5 regresses this world-level difference on a constant, separately for each model and noise level. A positive coefficient means that more Principals are harmed under AI reports.

Table D.5: Dependent variable: AI-minus-Bayesian harmed share (percentage points)

Model / noise	Worlds	Coefficient	SE	95% CI	Holm p	Flip Holm p
Gemini / 0%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Gemini / 10%	17	-0.980	0.980	[-3.06, 1.10]	1.0000	1.0000
Gemini / 25%	17	-1.961	2.426	[-7.10, 3.18]	1.0000	1.0000
Gemini / 35%	18	-0.926	1.635	[-4.38, 2.52]	1.0000	1.0000
Sol / 0%	18	0.926	0.926	[-1.03, 2.88]	1.0000	1.0000
Sol / 10%	18	0.926	0.926	[-1.03, 2.88]	1.0000	1.0000
Sol / 25%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Sol / 35%	18	0.926	2.119	[-3.54, 5.40]	1.0000	1.0000
Fable / 0%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Fable / 10%	19	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Fable / 25%	18	0.926	0.926	[-1.03, 2.88]	1.0000	1.0000
Fable / 35%	19	-0.877	0.877	[-2.72, 0.97]	1.0000	1.0000

We compare the percentage of Principals harmed under Agents' reports with the percentage harmed under Bayesian reports. A positive coefficient means that more Principals receive an object worth less than their endowment under Agents' reports. The coefficient is measured in percentage points. We average the six Principal-level differences within each world and estimate $D_w = \beta + e_w$, where D_w is that world average. Equivalently, we regress the two outcomes on world fixed effects and one policy or reporting-rule indicator, using $2N$ observations and N world clusters. The standard error is $s_D/\sqrt{N}$; the balanced CR2 calculation uses $N - 1$ degrees of freedom. We include no controls determined after treatment. Holm adjusts across the twelve rows for this outcome. Exact sign-flip tests enumerate the world-level sign assignments under a symmetry assumption. An all-zero sample receives $p = 1$ and a zero-width descriptive interval, which does not establish population equivalence. We omit R^2 because the intercept-only regression provides no informative measure of fit.

The share of harmed Principals describes how widely losses are distributed, but it does not show their size. A Principal can lose under both reporting rules while suffering a larger loss under one of them. We therefore measure the size of each loss as $\max\{0, u_i(\text{owned}) - u_i(\text{assigned})\}$, where u_i is Principal i 's true utility. A Principal who gains or receives an equally valuable object has a loss of zero. Table D.6 compares the average loss under AI reports with that under Bayesian reports. We first average the differences within each world so that inference accounts for the dependence among the six assignments.

Some model-noise conditions contain few differences between the reporting rules, and some contain none. A zero-width sample interval in such a condition reflects the absence of observed variation, rather than establishing equality in the population. The replication files retain whole-world bootstrap intervals, estimates using seventeen common worlds, minimum detectable effects, and the combined forty-eight-test sensitivity.

Table D.6: Dependent variable: AI-minus-Bayesian loss severity (points per Principal)

Model / noise	Worlds	Coefficient	SE	95% CI	Holm p	Flip Holm p
Gemini / 0%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Gemini / 10%	17	-0.020	0.020	[-0.06, 0.02]	1.0000	1.0000
Gemini / 25%	17	-0.069	0.057	[-0.19, 0.05]	1.0000	1.0000
Gemini / 35%	18	-0.019	0.055	[-0.14, 0.10]	1.0000	1.0000
Sol / 0%	18	0.009	0.009	[-0.01, 0.03]	1.0000	1.0000
Sol / 10%	18	0.009	0.009	[-0.01, 0.03]	1.0000	1.0000
Sol / 25%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Sol / 35%	18	0.083	0.116	[-0.16, 0.33]	1.0000	1.0000
Fable / 0%	18	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Fable / 10%	19	0.000	0.000	[0.00, 0.00]	1.0000	1.0000
Fable / 25%	18	0.019	0.019	[-0.02, 0.06]	1.0000	1.0000
Fable / 35%	19	-0.026	0.026	[-0.08, 0.03]	1.0000	1.0000

We compare the average loss under Agents' reports with the average loss under Bayesian reports. A Principal's loss is the shortfall of the assigned object's true value below the value of the initial endowment, with gains and equal-value assignments recorded as zero loss. A positive coefficient means that Agents' reports produce a larger average loss, measured in points per Principal. We average the six Principal-level differences within each world and estimate $D_w = \beta + e_w$, where D_w is that world average. Equivalently, we regress the two outcomes on world fixed effects and one policy or reporting-rule indicator, using $2N$ observations and N world clusters. The standard error is $s_D/\sqrt{N}$; the balanced CR2 calculation uses $N - 1$ degrees of freedom. We include no controls determined after treatment. Holm adjusts across the twelve rows for this outcome. Exact sign-flip tests enumerate the world-level sign assignments under a symmetry assumption. An all-zero sample receives $p = 1$ and a zero-width descriptive interval, which does not establish population equivalence. We omit R^2 because the intercept-only regression provides no informative measure of fit.

Online Supplement: Learning Preferences through Noisy Communication

Samir Huseynov
Auburn University
szh0158@auburn.edu

October 8, 2026 | Version 7.4.1

Online supplement: protocol, examples and extended diagnostics

In this supplement, we provide examples of the communication protocol and trading mechanism, additional diagnostics, and the sequential Bayesian question analysis. The paper’s appendix presents the regressions and robustness checks supporting the main results. Some tables below retain the samples and testing families used in earlier analyses. They examine the same collected data and should not be treated as additional independent experiments. The paper and its replication files also report the market-payoff question policy.

S1 Worked explanations and examples

S1.1 A simplified example of noisy answers

We first illustrate Bayesian updating using a constructed example with two equally likely preference types. This example is separate from the collected markets. One type prefers A and the other prefers B. At 35% noise, a Principal who prefers A answers A with probability 0.65, while a Principal who prefers B answers A with probability 0.35. After observing A, we apply Bayes’ rule:

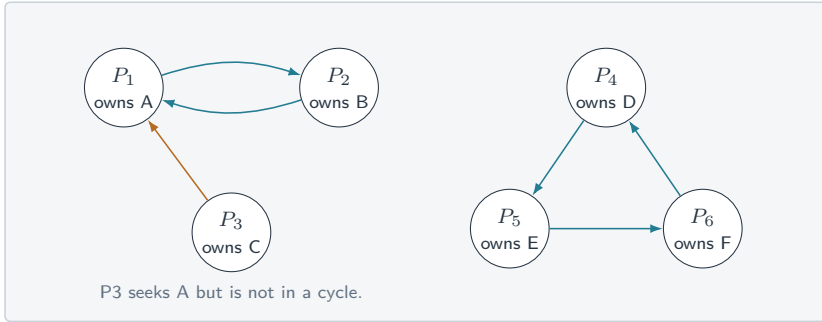
$$\begin{aligned} & \Pr(\text{A-preferring type} \mid \text{answer A}) \\ &= \frac{0.65 \times 0.5}{0.65 \times 0.5 + 0.35 \times 0.5} = 0.65. \end{aligned}$$

Thus, the probability that the Principal prefers A increases from 50% to 65%. The reply provides information but does not reveal the preference type with certainty. If a second independent reply also names A, this probability rises to $0.65^2 / (0.65^2 + 0.35^2) \approx 0.775$. If the second reply names B, the two replies offset and the probability returns to 0.5. Our experiment applies the same logic to 81 possible preference types. The replication files `worked_bayes_81_types.csv` and `worked_example.json` provide a complete example with exact fractions. We select that example by filename order rather than its outcome.

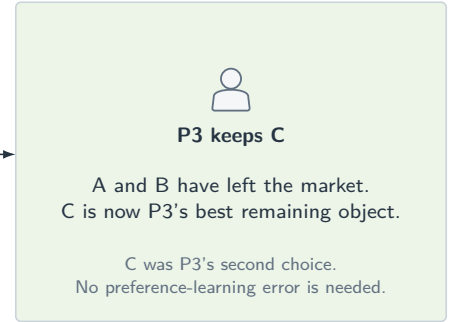
S1.2 Top Trading Cycles with six principals and six objects

(a) Round 1: two trading cycles

Constructed example. Each node shows a principal and the object initially owned.



(b) Round 2: one principal remains



Arrows point to the owner of the best remaining object; principals in each cycle receive the object they seek.

Final assignment: P1–B P2–A P3–C P4–E P5–F P6–D

Figure S1.1: How TTC assigns objects across trading rounds. We construct this example to illustrate the mechanism; it is not a collected market from the 81-type environment. An arrow points from each Principal to the owner of their most-preferred available object. Two cycles trade in the first round. P3 remains and keeps C in the second round, even though P3 initially prefers A.

We begin with P1 owning A, P2 owning B, P3 owning C, P4 owning D, P5 owning E, and P6 owning F. P1 prefers B and P2 prefers A, so each points to the other. These arrows form a trading cycle. P1 receives B and P2 receives A, after which both leave the market. P4, P5, and P6 form another cycle: P4 wants E, P5 wants F, and P6 wants D. They receive those objects and also leave. The two cycles can trade in the same round because they involve different Principals and objects.

P3 initially ranks A first and C second. In the first round, P3 points to A's owner but does not belong to a cycle. Once the other cycles trade, A is no longer available. In the second round, only P3 and C remain, so P3 keeps C. Consequently, receiving an initially second-ranked object need not reflect a preference-learning error. TTC assigns objects according to preferences over what remains available. The same reasoning applies when several preferred objects leave before a Principal trades. The file `ttc_teaching_example.csv` records the full rankings and independently verified assignment.

In the data, we distinguish an object's position after tie-breaking from its strict rank, which counts only objects reported as more valuable. We also record the objects available in each round, the executed cycles, and the assignments obtained from Bayesian reports based on the same conversations. These measures allow us to examine whether a lower-ranked assignment reflects unavailable objects, ties, or reporting discrepancies.

S1.3 Why different temporary choices can produce the same allocation

In each TTC round, every remaining Principal points to their favorite available object, and every object points to its owner. Principals who belong to a cycle receive their selected objects and leave. A Principal who points into another cycle waits until the next round. Consequently, two sets of reports can produce different temporary choices for a waiting Principal while producing the same completed trades. After the desired object leaves, that Principal selects again from the remaining objects.

Table S1.1: A saved market: temporary pointers can differ without changing assignments

Reports	Round	Principals remaining	Object pointers	Executed cycles
AI	1	[0, 1, 2, 3, 4, 5]	0→A, 1→A, 2→E, 3→B, 4→C, 5→B	5→5; 4→4
AI	2	[0, 1, 2, 3]	0→A, 1→A, 2→E, 3→F	0→3→0
AI	3	[1, 2]	1→E, 2→E	1→1
AI	4	[2]	2→D	2→2
Bayes	1	[0, 1, 2, 3, 4, 5]	0→A, 1→A, 2→B, 3→B, 4→C, 5→B	5→5; 4→4
Bayes	2	[0, 1, 2, 3]	0→A, 1→A, 2→F, 3→F	0→3→0
Bayes	3	[1, 2]	1→E, 2→E	1→1
Bayes	4	[2]	2→D	2→2

We show the recorded TTC sequence for Gemini in world 00 at 35% noise. Principal indices 0–5 follow the saved data order and are distinct from object names. Each cycle lists Principals, with each receiving the object to which they point. Some temporary pointers differ between the reporting rules, but the final object assignments are identical. The complete trace and initial endowments are recorded in `v7_pointer_traces.json`.

S1.4 A saved conversation and the submitted message structure

One saved conversation

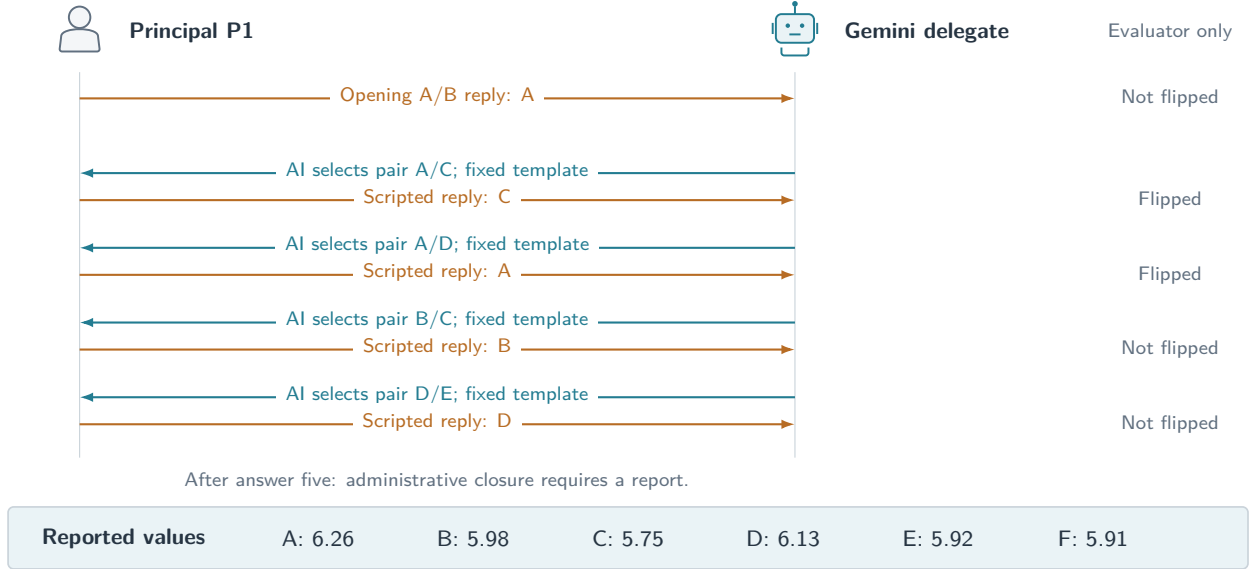


Figure S1.2: How an Agent selects follow-up questions. We show Gemini’s saved conversation with P1 in the first scheduled world at 35% noise. The rows summarize the actual object pairs and answers rather than reproducing the conversation verbatim. Teal arrows show pairs selected by the Agent within the fixed question template, and amber arrows show scripted Principal replies. The evaluator column identifies incorrect replies; the Agent does not see this information. The protocol supplies the opening pair, and Gemini selects the later pairs. Questioning ends after the fifth answer because the administrative limit is reached. We round the final six reported values to two decimal places for display and retain their exact values in the data.

The protocol fixes the wording of every question. Agents choose which pair to ask about and whether to ask again. Their setup provides the object attributes, the prior over 81 preference types, the utility formula, and the answer-error probability. It keeps the Principal’s realized type, true values, and random answer draws hidden. We first supply the opening A/B reply. The Agent can then submit an ASK action naming a pair or an ACT action reporting six estimated values. The script generates a reply from the specified error process and adds it to that Principal’s separate conversation. After the fifth answer, the protocol requires ACT. Agents are not told the numerical cap in advance, and the prompt does not instruct them to calculate a Bayesian posterior.

We select Figure S1.2 by its scheduled world, Principal, model, and noise rate. The two incorrect replies alter the evidence received by the Agent, while the Principal’s underlying values remain fixed. The final reports need not recover those values exactly. We reconstruct the Bayesian estimates separately and run TTC once all six Agents in the market have submitted valid reports.

S1.5 A worked paired-effect calculation

Consider a hypothetical world with total payoff of 45 at zero noise and 39 at 35% noise. Its noise effect is $39 - 45 = -6$ points for the market, equivalent to $-6/6 = -1$ point per Principal. We calculate the same difference in every available paired world and average those differences. Worlds can have different baseline payoff levels, but each comparison holds that world’s preferences and ownership fixed. The standard error measures variation across these world-level changes, accounting for the dependence among the six Principals within a market.

S2 Full diagnostic and robustness results

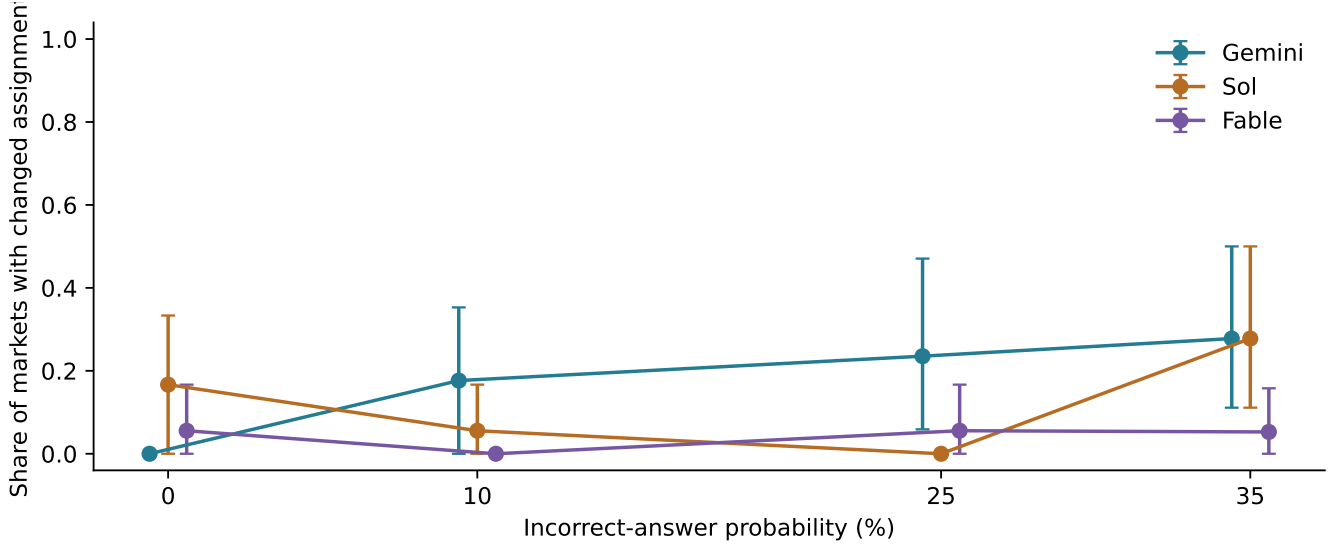


Figure S2.1: How often Agents’ reports change assignments relative to Bayesian reports. We compare TTC assignments under Agents’ reports with assignments under Bayesian estimates based on the same conversations. The figure uses all complete markets for each model and noise level. Intervals resample whole worlds, retaining the dependence among markets in a world. The paper’s appendix reports the corresponding counts.

S2.1 Numerical agreement and signed errors

Table S2.1: Numerical agreement with the same-transcript Bayesian values

Model	Noise (%)	n sessions	Within 0.005 (%)	MAE	Median max	90th max
Gemini	0	111	98.2	0.007	0.0000	0.004
Gemini	35	111	20.7	0.511	1.0126	1.754
Sol	0	111	84.7	0.123	0.0000	1.500
Sol	35	111	49.5	0.081	0.0054	0.579
Fable	0	113	85.8	0.045	0.0043	0.591
Fable	35	114	78.9	0.007	0.0044	0.005

We compare each reported object value with the exact Bayesian estimate from the same conversation. “Within 0.005” is the percentage of conversations in which all six differences are no larger than 0.005 utility points, the tolerance used in the historical rehearsal. “Median max” and “90th max” summarize the largest absolute difference within each conversation. MAE averages absolute differences using complete markets; the other columns use all valid conversations. The accompanying CSV retains the earlier 10^{-6} tolerance and additional cutoffs so that agreement can be examined under different definitions.

We examine numerical agreement under several tolerances. The table retains the historical 0.005 criterion, while the replication files also report the earlier 10^{-6} criterion and tolerances of 0.001, 0.01, and 0.1. Reporting these alternatives shows how the measured agreement depends on the chosen cutoff. Numerical closeness alone does not establish how an Agent reasons or whether two reporting rules have equivalent welfare effects.

Table S2.2: Signed differences between reported and Bayesian values at 35% noise

Model	Worlds	Signed mean	95% CI	Holmp
Gemini	19	0.008	[-0.05, 0.06]	1.0000
Sol	19	-0.002	[-0.02, 0.02]	1.0000
Fable	19	0.001	[-0.00, 0.00]	1.0000

We subtract the Bayesian estimate from each Agent’s reported value and average these signed differences within each world using its available valid conversations. A positive estimate means that reports are higher on average than the corresponding Bayesian values. The sample includes worlds where collection began, even if some markets remained incomplete, and therefore differs from the main complete-market sample. Holm adjusts across twelve comparisons covering three models and four noise levels. Positive and negative errors can cancel in this average, so we also report absolute discrepancies.

We also retain object-level signed and absolute errors, the shares of positive and negative errors, results excluding true ties, squared errors relative to truth and Bayesian estimates, strict-pair agreement, and Kendall’s tau-b in the analysis files. Signed errors reveal whether reports tend to be higher or lower, whereas absolute errors measure the size of discrepancies regardless of direction. Since absolute error is nonnegative, a nonsignificant test against zero would not establish exact Bayesian behavior.

S2.2 Question frequencies and information paths

Table S2.3: Follow-up questions and the information they provide

Model	Follow-ups 0/35%	Cap % 0/35%	Best IG %	Best choice %	IG gap
Gemini	3.88 / 4.00	88.9 / 100.0	3.6	38.5	0.0072
Sol	3.81 / 4.00	82.4 / 100.0	4.1	49.5	0.0059
Fable	2.57 / 0.81	15.7 / 0.9	1.1	27.2	0.0070

Follow-ups exclude the opening answer, and the cap column reports the percentage of conversations reaching the administrative question limit. At 35% noise, “Best IG” reports the percentage of selected questions tied for the greatest expected information gain among fifteen pairs. “Best choice” counts questions tied for the largest expected improvement in the value of the best object a Principal could select directly from the six objects. This one-step diagnostic does not account for TTC restrictions imposed by other Principals. The IG gap measures the shortfall from the greatest available information gain, in bits. These query-level percentages include observed follow-ups from failed conversations, which are marked in the data. They describe question choices and do not treat follow-ups as independent observations.

We describe how many follow-up questions Agents ask, how often they reach the administrative cap, and how their next question compares with available alternatives. These diagnostics examine information gathering within the protocol. They do not establish that voluntary stopping is optimal.

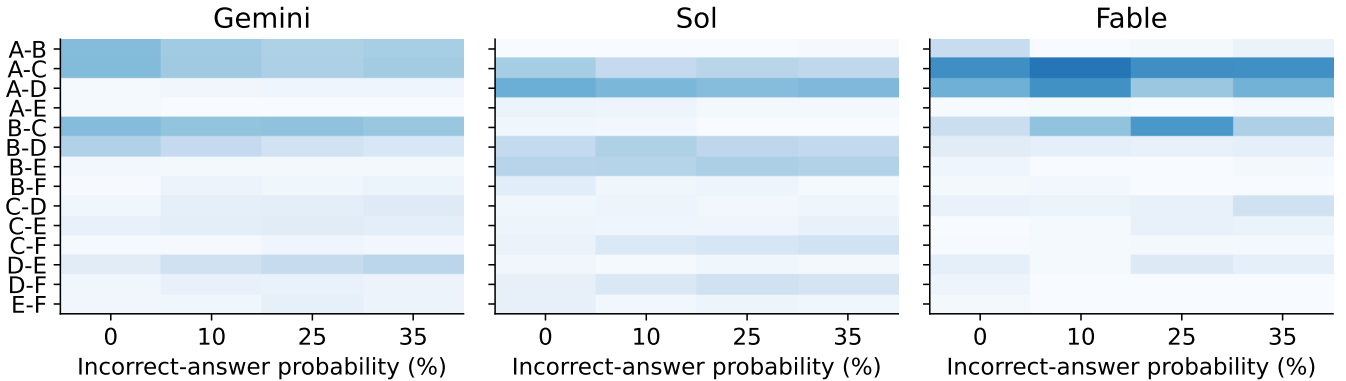


Figure S2.2: Which object pairs do Agents ask about? Each row is an object pair. Within each model and noise level, we divide the number of times each pair is selected by the total number of observed follow-ups. Darker cells therefore indicate more frequent choices. We exclude the fixed opening A/B comparison. These frequencies describe the observed queries, which remain dependent within conversations.

The analysis files also record rankings and selected learning paths. Estimates after later answers include only conversations in which questioning continued. Consequently, changes along those paths do not identify the causal effect of adding a question.

S2.3 Allocation benchmarks, ties, and effects on other Principals

Table S2.4: Payoff by the assigned object’s reported rank

Model	Reported rank	n	Gain vs own object	Gap vs Bayes
Fable	first	208	1.32	0.00
Fable	second	124	0.53	-0.01
Fable	third or lower	112	-0.06	0.05
Gemini	first	208	1.68	0.05
Gemini	second	105	0.91	0.01
Gemini	third or lower	107	0.08	-0.21
Sol	first	212	1.67	-0.01
Sol	second	121	0.80	0.00
Sol	third or lower	99	0.26	0.04

We group final allocations in complete markets by the assigned object’s rank in the Agent’s original report. Rank equals one plus the number of objects reported as strictly better, allowing equally valued objects to share a rank. The payoff columns compare the assigned object’s true value with the initial endowment and with the object assigned using Bayesian reports from the same conversation. Since the assigned rank is itself an outcome, these comparisons describe associations rather than causal effects of receiving a lower-ranked object.

We compare the assigned object’s reported rank with its payoff gain relative to ownership and its payoff difference relative to the Bayesian assignment. As the TTC example shows, receiving a lower-ranked object can reflect the departure of preferred objects rather than a reporting error.

Table S2.5: Market payoff under alternative reporting and allocation rules at 35% noise

Model	Delegate	Bayes	Truth TTC	No trade	Max sum	Max true IR
Fable	36.00	36.00	43.79	34.47	47.89	47.32
Gemini	37.89	38.67	44.11	34.61	47.89	47.28
Sol	37.00	36.83	44.11	34.61	47.89	47.28

We compare total true payoff across complete markets at 35% noise. “Delegate” uses Agents’ reports, “Bayes” uses posterior estimates from the same conversations, and “Truth TTC” uses true values in the same TTC mechanism. “No trade” retains initial ownership. “Max sum” selects the highest total payoff among all 720 assignments. “Max true IR” restricts that calculation to assignments that leave every Principal at least as well off as their endowment. The last two columns describe feasible payoff benchmarks calculated by the evaluator, rather than market mechanisms tested on Agents.

We also rerun each complete market after reversing the common alphabetical tie-breaking priority, separately under Agents’ reports, Bayesian reports, and true values. This calculation changes the tie rule for all six Principals together. It examines the sensitivity of allocations to ties using the existing reports and adds no new model responses.

Table S2.6: Offline replacement of one report with true values at 35% noise

Model	Replacements	Own change	Others'change	Total change
Fable	114	1.29	-0.29	1.00
Gemini	108	1.38	-0.28	1.10
Sol	108	1.31	-0.02	1.29

We replace one Principal’s report with their true values while keeping the other five reports fixed, then rerun TTC. We repeat this calculation for each Principal in every complete market. The columns report the average payoff change for the represented Principal, the other five Principals combined, and the market as a whole. These six replacements within a market are dependent, and their effects need not add. The calculation examines the consequences of report accuracy; it is not a remedy deployed in the experiment.

S3 Sequential Bayesian question selection

S3.1 The economic question and the objective

An Agent can estimate values correctly from the information received but still learn little because it selects uninformative questions or stops early. We therefore construct a policy that chooses questions from the same fifteen object pairs to minimize expected squared value-estimation error. The policy considers how possible answers affect subsequent questions. We calculate this benchmark using the existing data and the known response process, without collecting additional model responses.

The policy uses the same uniform prior over 81 types, the same opening A/B comparison, and the same known answer-error probability. It updates after every reply and then selects the next pair. Repeated questions receive fresh answer draws. Unlike the original Agents, the policy knows its remaining question budget. We assign no cost to an additional question. These assumptions define what the policy optimizes and must be considered when interpreting Agents’ stopping choices.

Let p denote the posterior distribution over types after the observed replies. Let u_j be the Principal’s true value for object j , where j indexes the six objects. If the policy stops and reports posterior means, its expected squared error per object is

$$R(p) = \frac{1}{6} \sum_{j=1}^6 \text{Var}_p(u_j).$$

A lower value of $R(p)$ means more accurate value estimates on average. It does not directly measure gains from trade. With h follow-ups remaining, we calculate

$$V(p, 0) = R(p), \quad V(p, h) = \min_{q \in \mathcal{Q}} \sum_{a \in q} \Pr(a \mid p, q) V(p^{q,a}, h - 1).$$

Here, $\mathcal{Q}$ contains the fifteen pairs, a names the object in the reply, and $p^{q,a}$ is the updated posterior after that reply. For each possible question, we consider both answers and the best subsequent question choices. We select the pair that produces the lowest expected final error. Thus, the calculation solves the finite sequence of choices rather than selecting only the best immediate question. We use the recorded alphabetical convention to resolve ties. The solver uses double-precision arithmetic with an explicit tie tolerance, and independent exact-rational checks reproduce the frontier values.

Squared error is one possible objective for preference learning. Other objectives include reducing entropy, reducing absolute error, improving the Principal’s TTC payoff, or raising total market payoff. Posterior means minimize squared error, but other objectives can imply different reports or questions. Because questions are free in this benchmark, the policy can use its entire budget. That assumption does not establish that an Agent with different beliefs or perceived question costs stopped irrationally.

Table S3.1: How much uncertainty can optimal questions remove?

Wrong-answer probability	Opening only	Two follow-ups	Four follow-ups	Four-step greedy	Risk retained
0%	5.556	3.129	2.039	2.099	36.7%
10%	5.900	4.367	3.438	3.473	58.3%
25%	6.241	5.580	5.091	5.097	81.6%
35%	6.373	6.113	5.896	5.899	92.5%

We calculate expected squared value-estimation error per object under the known preference-type prior and all possible answer sequences. “Opening only” already incorporates the fixed A/B reply. “Risk retained” divides the error after four optimally chosen follow-ups by the opening-only error at the same noise level. These are quantities calculated from the specified model, rather than estimates from sampled worlds, so sampling confidence intervals do not apply. The greedy policy selects the question with the largest immediate expected error reduction at each step, whereas the optimal policy also considers subsequent questions.

At 35% noise, expected squared error falls from 6.373 after the opening answer to 5.896 after four optimally selected follow-ups. Thus, 92.5% of the opening-only error remains. This percentage describes value-estimation uncertainty after the opening reply; it is not the share of welfare lost or of all prior information left unknown. The result shows how much uncertainty can remain after a short, noisy conversation even when question selection is optimized. Comparing the optimal and greedy policies also shows that planning the full sequence improves learning only slightly relative to repeatedly selecting the best immediate question in this environment.

S3.2 What do the Agents leave unlearned?

For every observed follow-up, we compare two policies starting from the same conversation history. The first selects the best next pair and then continues optimally. The second uses the Agent’s actual next pair and continues optimally afterward. The difference in expected final error measures the cost of that next question choice. Separately, when the Agent submits its reports, we calculate the additional learning available from the questions that remain before the administrative cap.

Table S3.2: The learning value left by question choice and stopping

Model / noise	Follow-ups	Choice cost	95% interval	Unused at stop	95% interval
Gemini / 0%	397	0.159	[0.149, 0.170]	0.000	[0.000, 0.000]
Gemini / 10%	408	0.152	[0.141, 0.163]	0.000	[0.000, 0.000]
Gemini / 25%	407	0.068	[0.065, 0.071]	0.003	[0.000, 0.009]
Gemini / 35%	408	0.031	[0.030, 0.033]	0.000	[0.000, 0.000]
Sol / 0%	391	0.119	[0.108, 0.130]	0.040	[0.008, 0.089]
Sol / 10%	408	0.132	[0.119, 0.146]	0.000	[0.000, 0.000]
Sol / 25%	408	0.061	[0.058, 0.064]	0.000	[0.000, 0.000]
Sol / 35%	408	0.029	[0.028, 0.030]	0.000	[0.000, 0.000]
Fable / 0%	260	0.181	[0.164, 0.200]	1.031	[0.896, 1.182]
Fable / 10%	123	0.106	[0.098, 0.114]	1.680	[1.601, 1.753]
Fable / 25%	100	0.064	[0.057, 0.072]	0.875	[0.831, 0.918]
Fable / 35%	80	0.032	[0.028, 0.035]	0.387	[0.368, 0.407]

We use seventeen common complete worlds, giving 102 conversations in each model–noise condition. “Choice cost” measures the expected squared-error increase from using the Agent’s chosen next pair rather than the best pair, allowing optimal questions afterward in both cases. We average this quantity over observed follow-ups, so it describes choices conditional on asking. “Unused at stop” averages across all conversations and measures the additional error reduction available from using the administratively remaining questions at zero cost. Both measures are in squared utility points per object, rather than market payoff points. Individual 95% intervals use percentiles from 10,000 whole-world resamples. The follow-up counts are numbers of decisions, not independent sample sizes.

At 35% noise, the average cost of the next question choice is about 0.03 squared points per object for each model. Consequently, selecting a pair that is not exactly optimal can have a small learning cost because several pairs provide similar information. Fable leaves more potential learning at submission because it often stops before using four follow-ups. When an Agent reaches the cap, the remaining learning value is zero by construction because no further questions are available. Neither calculation assigns a monetary cost to questions or establishes optimal voluntary stopping. The full calculations cover all 1,342 valid conversations. The table uses seventeen worlds complete across all models and rates to maintain a common sample.

S3.3 Replay with matched counts and with the full budget

We next examine market payoff using Bayesian reports generated by the Bayesian learning policy. We retain the original preferences, endowments, noise probabilities, and uniform answer draws at each answer position. Selecting a different pair can produce a different answer because the same draw is applied to a different true preference comparison. The simulator uses hidden preferences to generate and evaluate replies, while the policy observes only its prior and the replies it receives.

In the matched-count replay, each Principal’s policy receives exactly as many follow-ups as the original Agent used. In the full-budget replay, every policy receives four follow-ups. We compare both with TTC using Bayesian estimates from the Agents’ original conversations. Using Bayesian reports throughout removes differences in numerical reporting from these comparisons. Payoff is measured across six Principals. Dividing an estimate and its interval by six expresses the same result per Principal.

Table S3.3: Matching each Agent's realized question count

Model / noise	Follow-ups	Welfare change	95% interval	Holm t p	Holm flip p	80% MDE	Power (3)
Gemini / 0%	3.89	+0.29	[-0.18, 0.77]	1.0000	1.0000	0.95	100%
Gemini / 10%	4.00	-1.59	[-3.43, 0.26]	0.8681	0.9937	3.71	56%
Gemini / 25%	3.99	-0.06	[-2.51, 2.39]	1.0000	1.0000	4.92	28%
Gemini / 35%	4.00	-1.59	[-4.13, 0.96]	1.0000	1.0000	5.12	25%
Sol / 0%	3.83	-0.41	[-1.39, 0.57]	1.0000	1.0000	1.97	100%
Sol / 10%	4.00	-2.06	[-4.24, 0.13]	0.6937	0.8030	4.39	37%
Sol / 25%	4.00	-1.12	[-3.78, 1.54]	1.0000	1.0000	5.34	22%
Sol / 35%	4.00	+0.88	[-1.57, 3.33]	1.0000	1.0000	4.93	28%
Fable / 0%	2.55	-0.65	[-2.60, 1.31]	1.0000	1.0000	3.93	49%
Fable / 10%	1.21	+2.65	[0.38, 4.92]	0.3005	0.3457	4.56	34%
Fable / 25%	0.98	-0.29	[-2.04, 1.46]	1.0000	1.0000	3.52	62%
Fable / 35%	0.78	-0.82	[-2.83, 1.19]	1.0000	1.0000	4.04	46%

We give the policy the same number of follow-ups as each Agent used in the original conversation. We subtract TTC payoff under Bayesian reports based on the Agent’s original conversation from payoff under the policy’s Bayesian reports. Each row uses seventeen paired worlds, and follow-ups are measured per Principal. A positive estimate means higher total true payoff across the six Principals. We report individual 95% paired-world intervals using t_{16} . Exact sign-flip tests enumerate all signs under independent, symmetric world differences. Holm adjusts across the twelve comparisons within this question-budget condition. MDEs and power use a separate conservative two-sided significance level of $\alpha = .05/12$, a normal working model, and the observed paired standard deviation. “Power (3)” evaluates a specified three-point payoff change, rather than the estimated effect. These calculations do not account for selecting policies after examining the data.

None of the matched-count comparisons is statistically significant after either Holm adjustment. For example, Fable’s estimated improvement at 10% noise is 2.65 points, with a Holm-adjusted t -test p -value of .301. We therefore do not detect a welfare improvement from changing the question sequence while preserving question counts. The reported MDEs show that seventeen paired worlds leave substantial uncertainty about modest gains.

Table S3.4: Giving every Principal four follow-up questions

Model / noise	Follow-ups	Welfare change	95% interval	Holm t p	Holm flip p	80% MDE	Power (3)
Gemini / 0%	4.00	+0.00	[-0.26, 0.26]	1.0000	1.0000	0.52	100%
Gemini / 10%	4.00	-1.59	[-3.43, 0.26]	0.8681	0.9937	3.71	56%
Gemini / 25%	4.00	-0.06	[-2.51, 2.39]	1.0000	1.0000	4.92	28%
Gemini / 35%	4.00	-1.59	[-4.13, 0.96]	1.0000	1.0000	5.12	25%
Sol / 0%	4.00	+0.12	[-0.05, 0.29]	1.0000	1.0000	0.34	100%
Sol / 10%	4.00	-2.06	[-4.24, 0.13]	0.6937	0.8030	4.39	37%
Sol / 25%	4.00	-1.12	[-3.78, 1.54]	1.0000	1.0000	5.34	22%
Sol / 35%	4.00	+0.88	[-1.57, 3.33]	1.0000	1.0000	4.93	28%
Fable / 0%	4.00	+0.35	[-0.61, 1.31]	1.0000	1.0000	1.93	100%
Fable / 10%	4.00	+1.76	[-0.42, 3.95]	0.9561	1.0000	4.39	37%
Fable / 25%	4.00	+3.24	[1.10, 5.37]	0.0645	0.0425	4.28	40%
Fable / 35%	4.00	+0.82	[-1.53, 3.17]	1.0000	1.0000	4.72	31%

We give the policy four follow-ups for every Principal, including those whose Agents stopped earlier. We subtract TTC payoff under Bayesian reports based on the Agent’s original conversation from payoff under the policy’s Bayesian reports. Each row uses seventeen paired worlds, and follow-ups are measured per Principal. A positive estimate means higher total true payoff across the six Principals. We report individual 95% paired-world intervals using t_{16} . Exact sign-flip tests enumerate all signs under independent, symmetric world differences. Holm adjusts across the twelve comparisons within this question-budget condition. MDEs and power use a separate conservative two-sided significance level of $\alpha = .05/12$, a normal working model, and the observed paired standard deviation. “Power (3)” evaluates a specified three-point payoff change, rather than the estimated effect. These calculations do not account for selecting policies after examining the data.

The twelve full-budget comparisons use the same 68 simulated policy markets and compare them with different original-conversation benchmarks. They are therefore dependent. Each table applies its own Holm adjustment across twelve tests; the adjustment does not jointly cover both tables. At zero noise, very small paired standard deviations produce small estimated MDEs and power close to 100%.

For Fable at 25% noise, the estimated full-budget improvement is 3.24 points, compared with -0.29 when question counts are matched. Their difference is 3.53 points. This comparison changes the number of questions available to the policy and therefore cannot attribute the full-budget gain solely to better question selection. The relation is also not uniform across noise levels: at 10%, Fable’s matched-count estimate exceeds its full-budget estimate. More information can reduce expected squared error without improving the realized TTC assignment in every market.

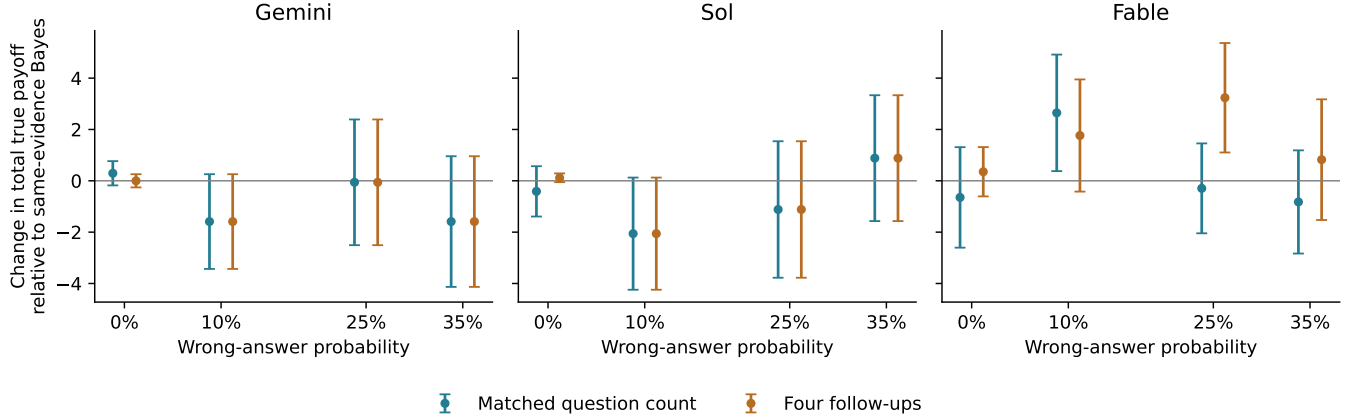


Figure S3.1: How question choice and question count affect the payoff comparison. We subtract total payoff under Bayesian reports from the Agents’ original conversations from payoff under the value-learning policy. Blue points give the policy each Agent’s original follow-up count; orange points allow four follow-ups. Each estimate uses seventeen paired worlds, and bars show individual 95% paired-world t intervals. We offset points horizontally for readability. Tables S3.3 and S3.4 report multiple-testing adjustments and power. Overlap between the two intervals does not test the difference between policies.

S3.4 A worked market: why the comparison matters

We illustrate the comparison using Fable in recorded world 07 at 25% noise. Initial ownership is worth 33 points across the six Principals. Trading on Fable’s reports produces 23 points, with three Principals receiving less valuable objects than their endowments. Bayesian reports from the same conversations produce 24 points. Thus, changing only the numerical reports raises total payoff by one point in this market.

The alternative Bayesian learning policy produces 26 points when it receives each Agent’s original question count and 37 points when every Principal receives four follow-ups. Relative to the original-conversation Bayesian payoff of 24, the full-budget gain is 13 points, consisting of $26 - 24 = 2$ and $37 - 26 = 11$. The full-budget payoff of 37 also equals truthful-report TTC payoff in this world.

The original replies help explain the loss. Principal 1 receives two incorrect answers out of three, and Principal 5 receives three incorrect answers out of three. Fable asks two follow-ups for each of them and none for Principals 4 and 6. Bayesian updating therefore operates on limited and sometimes misleading information.