Communicating with Data Generating Processes: An Experimental Analysis

Agata Farina
Princeton University

Clément Herman
Princeton University

April 29, 2026
Click here for the latest version

Abstract

In many applications, agents can more easily influence how data are generated than manipulate the data themselves. For example, students choose which classes to take but cannot modify their grades once assigned, and firms determine how performance indicators are calculated rather than directly altering results. This paper experimentally studies information transmission when data are generated through an unobservable, strategically chosen process. We focus on settings where an informed sender—such as a student or a firm—privately selects a data-generating process (DGP)—a portfolio of classes or an accounting methodology—to shape the beliefs of an uninformed receiver, such as an employer or an investor. Across treatments, we vary which DGPs are feasible and whether some come at a cost. These variations span different levels of information verifiability and capture, within a unified framework, the core insights of disclosure, cheap-talk, and signaling models. Our findings reveal three main patterns. First, while senders select their DGPs strategically in line with theoretical predictions, receivers often fail to account for this strategic selection. Second, introducing differential costs across feasible DGPs mitigates this DGP selection neglect. Third, while receivers' biases are robust across environments, their consequences vary: information transmission falters most when the evidence receivers observe is highly sensitive to the sender's DGP choices. Finally, increasing transparency about the selected DGP—a policy often advocated in practice—does not, in general, improve information transmission.

We are deeply grateful to Guillaume Fréchette, Alessandro Lizzeri, and Leeat Yariv for their guidance and support throughout this project. We are also deeply indebted to Maher Said for numerous insightful discussions and constructive feedback that substantially improved this work. We thank Chiara Aina, Jeremy Bertomeu, Guido Deana, Joyee Deb, Mira Frick, Ryota Iijima, Alessandro Ispano, Mario Leccese, Marta Mojoli, Paola Moscariello, Pietro Ortoleva, Jacopo Perego, Wolfgang Pesendorfer, Kirtivardhan Singh, and Sevgi Yuksel for insightful comments and suggestions. We also thank seminar and conference participants at Princeton University, New York University, Penn State Smeal College of Business, University of Chicago Booth School of Business, University of North Carolina at Charlotte, Paris School of Economics, Sciences Po, Ohio State University, Carnegie Mellon University, Carlos III de Madrid, University of Oxford, LUISS Guido Carli, Olin Business School at Washington University in Saint Louis, and the 2025 Caltech Behavioral and Experimental Workshop, the 2025 BEEMA Conference, and the 2025 North America ESA Meeting for helpful feedback.

1 Introduction

Understanding how people extract information from data is fundamental to economics. In strategic environments, there are many ways in which informed agents—*senders*—can affect the evidence on which others—*receivers*—form their beliefs. In some settings, data can be misrepresented; in others, information can be withheld or selectively disclosed. These forms of strategic manipulation, extensively studied in the literature, often distort receivers' inference and motivate regulatory efforts to limit the sender's discretion.[1] However, in many applications, senders can more easily influence how data are generated than manipulate the data themselves. For instance, students choose whether to take more or less challenging classes but cannot change their grades once assigned; firms select accounting methodologies that shape how performance indicators are constructed rather than directly altering results; and pharmaceutical companies decide on trial parameters but cannot withhold trial outcomes.[2] In such cases, employers, investors, and regulators form beliefs based on data—GPA, performance indicators, and trial outcomes—which are generated through an unobservable, strategically chosen *data-generating process*. Such strategic choices create uncertainty about the interpretation of the available evidence. How agents resolve this uncertainty determines the effectiveness of information transmission.[3] Recognizing the resulting challenges can inform policy design and motivate a shift of attention from *ex-post* data manipulation to *ex-ante* strategic data generation.

This paper develops an experimental framework to study how senders strategically select data-generating processes (DGPs) and how this affects belief formation. A sender endowed with some private information (his *type*) seeks to influence the beliefs of an uninformed receiver through the unobservable choice of a DGP, which may come at a cost. Once data are produced according to the chosen DGP, the sender has no discretion over their disclosure. To infer the sender's type, the receiver must reason about which data-generating process likely produced the evidence—a task that combines *strategic* and *probabilistic* reasoning. We systematically vary the sender's technological and strategic constraints by changing the set of feasible DGPs and the costs of their selection. These variations allow us to study senders' choices and receivers' inferences in environments that differ in the *verifiability of information*—that is, in how much discretion the sender has over the realized data distribution[4]—and in the type of DGP selection

[1] For example, restaurants are required to display their hygiene grade cards (Jin and Leslie, 2003).

[2] See, for instance, Ting and Lee (2012), Ognjanovic et al. (2016), Fields et al. (2001), and Ioannidis (2005).

[3] Recent experimental work provides mixed evidence on belief formation under exogenous uncertainty about the data-generating process (e.g., Liang, 2025, Musolff and Zimmermann, 2025, Aina and Schneider, 2025).

[4] The literature typically uses *verifiable* and *unverifiable* information to describe whether the sender can mis-

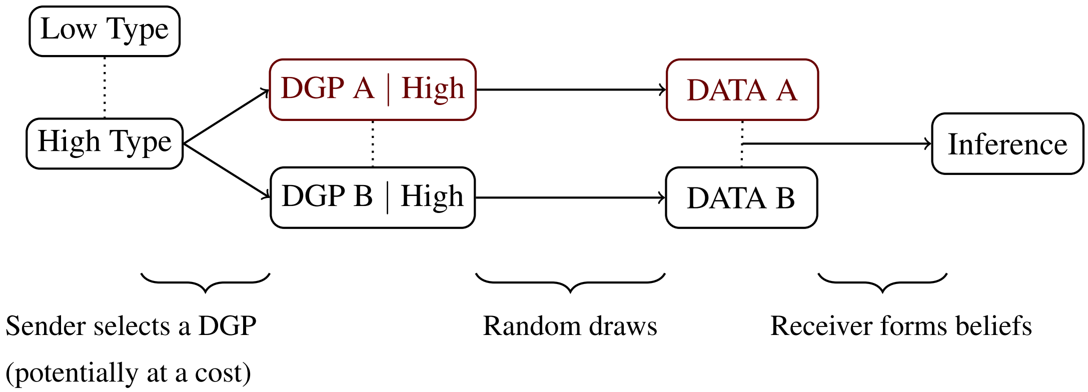


Figure 1: Overview of the theoretical framework.

predicted in equilibrium. By jointly varying the feasible DGPs and the selection costs, our framework nests classic information transmission games—including cheap talk, disclosure, and signaling—within a unified structure.

Our findings reveal three main patterns. First, senders' DGP choices align with theoretical predictions, but receivers often fail to account for this strategic selection. Second, introducing differential costs across feasible DGPs mitigates this "DGP selection neglect." Third, while receivers' biases are robust across environments, their consequences vary: information transmission is most impaired when the evidence receivers observe is highly sensitive to the senders' DGP choice. Finally, we study the effect of an increase in DGP transparency by making the senders' DGP choices observable to the receivers, and find that it does not, in general, improve information transmission.

Theoretical Framework and Experimental Design. Our model captures the insights of several classical information transmission games within a unified framework and guides our experimental design (see Figure 1 for an overview of the framework). The sender's type is private information and can be *high* or *low*. The sender seeks to convince the receiver that his type is high. For instance, a student knows his own ability and wants to positively impact the employer's perception. While the sender cannot directly communicate his type, he can provide the receiver with hard data—such as grades—that depend on the underlying type and on an element of randomness. The sender can influence the distribution of the data through the choice of a data-generating process (DGP). For simplicity, we assume a binary set of feasible DGPs, representing, for example, two courses available to the student. Once realized, the data must

report evidence. Here, we use the term more broadly to capture discretion over the mapping from private information to data, tracing a continuum of verifiability.

be disclosed and cannot be manipulated, capturing the constraint that prevents a student from altering his transcript after grades are assigned. Choosing among DGPs may involve exogenous selection costs that can vary across types—for instance, some courses have prerequisites that are harder to meet for certain types of students.

Two assumptions are key. First, the feasible data distributions *depend on the sender's type*: for any class, higher ability students are more likely to achieve higher grades. Second, the chosen DGP is *unobservable* to the receiver—e.g., employers cannot assess how challenging a class is. The receiver only sees the realized data to form a joint inference about the sender's type and the relevant DGP. The interaction of feasible DGPs and selection costs determines the sender's incentives, shaping his strategic choice, the resulting data distribution, and the receiver's inference.

Two polar cases illustrate the role of the feasible DGPs on inference. When all feasible DGPs induce the same distribution of data conditional on the type, the sender cannot affect the final distribution. This corresponds to a *verifiable information* environment: the mapping between type and data is fixed. At the opposite extreme, when all DGPs induce the same distribution regardless of the type, data convey no information about the sender's private information. This yields an *unverifiable information* environment: data are informative about the chosen DGP, while inference about the type relies entirely on strategic reasoning.[5]

While the available DGPs define the *information environment*, selection costs shape incentives within it. Together, these features allow the framework to capture the core mechanisms of both communication and signaling games. When DGP selection is costless, the sender's choice resembles that in a standard *communication* game, where each type can freely select among feasible messages to influence the receiver's belief. Introducing positive selection costs, instead, can make certain DGPs too expensive, limiting some types' ability to shape inference and generating *signaling* opportunities through the choice of DGP.[6]

Our experiment implements this framework in the laboratory. A sender of high or low type can choose between two available DGPs. Each feasible DGP maps the sender's type into a probability distribution over binary signals and satisfies a monotone likelihood ratio property: a

[5] Under extreme parametrizations of available DGPs, our framework can nest standard disclosure games (Grossman, 1981; Milgrom, 1981; Jovanovic, 1982; Okuno-Fujiwara et al., 1990) and standard cheap talk games (Crawford and Sobel, 1982).

[6] With extremely unverifiable information and the appropriate costs, the setting nests a standard signaling game (Spence, 1978). It is important to highlight that, differently from standard noisy signaling games (Matthews and Mirman, 1983; Carlsson and Dasgupta, 1997), the data resulting from the sender's choice contain objective information about the sender's type.

positive signal is more likely when the type is high. DGPs differ in their likelihood of producing positive signals, and such a ranking is consistent across types. Hence, we can think of the sender's options as comprising an *Easy* DGP that generates more positive signals in expectation and a *Hard* DGP that generates fewer. After the DGP is chosen, a fixed number of signals is drawn according to the selected process and shown to the receiver, who makes an incentivized inference about the sender's type and the chosen DGP.

Across our treatments, we vary the information environment, namely the set of feasible DGPs available to the sender. We implement, between subjects, two possible information environments that capture the relevant insights of the two polar cases. In our "*verifiable information environment*", we assume that any DGP available to the high-type sender produces, in expectation, more positive signals than those available to the low type. The sender's actions can change the magnitude of the receiver's updates, but not the direction: an additional positive signal is always stronger evidence of a high type. This implies that the data are intrinsically informative about the type, regardless of strategic considerations. In our "*unverifiable information environment*", we assume that the Hard DGP for the high type induces a lower probability of a positive signal than the Easy DGP for the low type—a high-ability student taking a very hard class may, in expectation, receive a lower grade than a low-ability student taking an easy class. In this case, the sender's choice can affect not only the magnitude but also the direction of the receiver's update: strategic reasoning becomes essential for inference.[7]

Regarding selection costs—which vary within subjects—we study three settings. First, we consider a case in which all DGPs are costless. Then, we examine two cases where the low-type sender must pay a cost to choose either the Easy or the Hard DGP, while the high type faces no cost. These costs stand in for real-world frictions: for example, a low-ability student may face higher costs to meet a class prerequisite, and so be more constrained in the class choice. Such costs limit the low type's ability to mimic the data distribution of the high type, affecting equilibrium predictions.

The theoretical framework yields several predictions that depend on the information environment and the cost of DGP selection. In the absence of costs, both sender types choose the Easy DGP to maximize, in expectation, the number of positive signals that trigger more optimistic receiver beliefs. When costs are introduced, however, separation may be sustained in equilibrium, and the choice of DGP can lead to more informative outcomes. In equilibrium, receivers should account for this strategic selection and form beliefs about both the sender's type and the

[7] Throughout the paper, we use the terms verifiable and unverifiable information to describe environments that approximate—but do not exactly coincide with—the polar cases discussed above.

chosen DGP as if they knew the relevant mapping from types and DGPs to data.

Empirical Findings. Our data reveal several patterns. First, senders' DGP choices are qualitatively consistent with the theory: subjects respond to both exogenous incentives—introduced through costs—and to endogenous incentives—induced by receivers' behavior. Quantitatively, however, selection is less extreme than predicted. This deviation is most pronounced when information is unverifiable and costs are absent, a setting in which senders face greater uncertainty about receivers' responses. Overall, these results indicate that strategic DGP selection is a key driver of sender behavior. As a consequence, the distribution of evidence depends critically on these choices and must be accounted for by receivers when interpreting the data.

Receivers' guesses vary both with the realized signals and across treatments, suggesting that subjects respond to the primitives of the inference problem as well as to the observed data. Despite this positive result, we observe large and systematic deviations from equilibrium behavior. Exploiting the within-subject variation in cost parameters, we find that in both verifiable and unverifiable environments, only about 15% of subjects behave optimally throughout the session. In contrast, 35 – 40% of subjects almost fully neglect the senders' strategic choices and act as if DGPs were randomly selected. This bias leads subjects to interpret data as having an "objective meaning"—one that depends on the information environment rather than on the sender's selection. Finally, we identify a large group of receivers who partially neglect DGP selection but respond to exogenous costs in ways consistent with the theory.

The patterns in receivers' behavior highlight three main points. First, although subjects respond to the probabilistic features of the environment, they do not consistently condition their choices on the senders' DGP selection. Second, receivers' ability to account for strategic selection may vary across environments. Third, exogenous incentives may be easier to incorporate into their reasoning than endogenous ones. We confirm the limited role of probabilistic challenges through non-strategic treatments in which receivers are informed about the DGP selection rule. Despite the potentially complex inference task, we find no systematic biases in aggregate inferences in these cases.[8]

Although receivers' behavior is generally consistent across environments, the implications for inference differ. In the verifiable information setting, most deviations concern beliefs about the DGP, while subjects effectively extract information about the type, for which the data are intrinsically informative. As a result, inference improves when exogenous costs are introduced.

[8]The experimental literature has extensively studied how subjects perform in Bayesian updating tasks and documents recurrent failures. See Benjamin (2019) for an overview.

Receivers become more accurate about the chosen process, both because costs make DGP selection more predictable and because separation generates more extreme data realizations. In contrast, in the unverifiable information setting, the data are highly informative about the DGP but not about the sender's type. While inference about the process remains accurate, introducing costs that induce separation worsens inference about the type, even among receivers who partially react to costs. This pattern provides insights into effective information extraction. It suggests that information transmission about the sender's private type—often the main outcome of interest—is particularly impaired when the realized evidence is highly sensitive to the senders' DGP choice. Nevertheless, despite these biases, receivers tend to extract more information when more is transmitted by the senders.

Implications of DGP Transparency. Our analysis uncovers a source of inference biases that traditional disclosure regulations—like mandated disclosure policies—cannot resolve. These policies focus on what information is released, not on how that information is generated. Yet, in many settings, the strategic choice of data-generating processes provides a natural channel for evidence manipulation. Regulating this production stage, however, may be much more complex in practice.

Recent regulatory efforts in science and accounting emphasize transparency in the processes that generate data—for example, through pre-registration in research or clear disclosure of how performance indicators are constructed. In our framework, such policies correspond to making the sender's DGP choice observable to the receiver. However, greater transparency does not necessarily enhance information transmission, as observability may alter senders' incentives and reduce the informativeness of the data. Our design includes an auxiliary treatment in which DGP choices are observable, and the results confirm that transparency alone does not improve information transmission.

1.1 Related Literature

Asymmetric information is central to many economic interactions. Communication and signaling games provide canonical frameworks to study how such asymmetries can be resolved and have been extensively studied both theoretically and experimentally. Our framework departs from these models in three key respects. First, unlike standard communication games without commitment, the sender cannot manipulate the message observed by the receiver—through concealment, selection, or misreporting—but can only influence the distribution from which data are drawn.[9] Second, unlike standard information design models, the choice of data-generating process (DGP) is *unobserved* by the receiver and may depend on the sender's private information.[10] Third, although potential asymmetries in action costs, our framework cannot be reduced to a standard (noisy) signaling game: the DGP choice is unobservable, and the resulting data contain objective information about the sender's type.[11] Even with these distinctions, extreme cost and DGP parametrizations can replicate standard disclosure, cheap talk, and signaling games within our setting.

Our paper contributes to the experimental literature on information transmission games. The varying degrees of *information verifiability* induced by the feasible DGPs connect our design to both disclosure and cheap-talk experiments. In disclosure games, several studies document the failure of the unraveling principle. For example, in the laboratory, Jin et al. (2021) shows that receivers tend to be insufficiently skeptical when senders conceal information, making it profitable to withhold unfavorable states.[12] In contrast, experiments on cheap-talk games find that senders tend to overcommunicate relative to theoretical predictions, while receivers exhibit excessive trust in disclosed messages (e.g., Cai and Wang, 2006).[13] A few studies examine settings where senders select verifiable evidence from a larger pool, and find that receivers insufficiently adjust for selected samples in their inference (Brown and Fragiadakis, 2019; Degan et al., 2023; Farina et al., 2024).[14] The sender's task in our framework—choosing a DGP—also relates to experimental studies of information design (e.g., Fréchette et al., 2022; Zhou, 2023; Rentschler and Samad, 2025). Unlike these studies, our senders do not commit to a DGP before learning their private type. Finally, the existence of asymmetric frictions in DGP selection connects our setting to the experimental signaling literature, which has primarily focused on the robustness of equilibrium selection criteria and on the impact of signaling technologies

[9] Seminal papers include Grossman (1981), Milgrom (1981), Jovanovic (1982), Verrecchia (1983), Dye (1985), Jung and Kwon (1988), Okuno-Fujiwara et al. (1990), Crawford and Sobel (1982), and Green and Stokey (2007).

[10] See Kamenica and Gentzkow (2011), Kamenica (2019), and Bergemann and Morris (2019). Recent work allows informed choice of the information structure, e.g. Perez-Richet (2014) and Koessler and Skreta (2023)).

[11] See Riley (2001) for a survey. Key contributions include Spence (1978), Leland and Pyle (1977), Kihlstrom and Riordan (1984), Matthews and Mirman (1983), and Carlsson and Dasgupta (1997).

[12] Forsythe et al. (1989), King and Wallin (1991), Dickhaut et al. (2003), Benndorf et al. (2015), Hagenbach and Perez-Richet (2018), de Clippel and Rozen (2024), Jin et al. (2022), Fréchette et al. (2022), Farina and Leccese (2024), Hagenbach and Saucet (2024), and Farina et al. (2024) also document excessive concealment. Similar evidence arises in field (Mathios, 2000; Jin and Leslie, 2003; Brown et al., 2012; Bertomeu et al., 2020).

[13] Blume et al. (2020) surveys this literature. Influential studies include Dickhaut et al. (1995), Forsythe et al. (1999), Blume et al. (1998), Sánchez-Pagés and Vorsatz (2007), Wang et al. (2010), and Wilson and Vespa (2020).

[14] Penczynski et al. (2023) studies selective disclosure in the presence of sender competition. Burdea et al. (2023) study a setting in which a sender sends a two-dimensional message, but only one dimension can be verified.

on information transmission.[15]

Because receivers in our setting form beliefs from data generated by an *unobserved* process, our work also connects to the emerging literature on inference under uncertainty about the data-generating process. Aina and Schneider (2025) study how subjects form beliefs when data may come from conflicting DGPs—i.e., DGPs that would lead to contrasting updating directions—and find that individuals tend to overweight the most likely one.[16] Instead, Liang (2025) studies belief updating when the accuracy of the data source is uncertain and finds systematic underreaction to information. Musolff and Zimmermann (2025) show that when individuals face multiple explanatory models of varying complexity, they display extreme model selection in actions, while their beliefs remain sensitive to model uncertainty. Our setting introduces an additional, *strategic*, dimension of uncertainty: receivers do not know which DGP was chosen or how that choice depends on the sender's type. One of the main contributions of our paper is to document how subjects cope with this additional uncertainty and to identify systematic biases in the relative weights they assign to different DGPs and the resulting effects on inference.

Our environment allows for both probabilistic and strategic inference errors, each of which can impair information extraction. Relatedly, Geng and Guan (2025) uses a persuasion game—where the choice of an experiment is observable—to structurally separate these two types of mistakes and find them to be orthogonal across subjects and comparable in magnitude. While we also detect both types of errors, the general pattern of deviations mostly reflects difficulties in accounting for the sender's strategic DGP choice. Similar to Ali et al. (2021), we design treatments with exogenous DGP selection that serve as benchmarks to isolate the effect of strategic uncertainty.

Receivers in our experiment often fail to account for the sender's strategic choices and, therefore, for the informational content those choices convey. This connects our work to the broader literature on mistakes in strategic inference and failures to extract information from others' actions (for surveys, see Kagel and Levin, 2016 and Niederle and Vespa, 2023). While the patterns in our data connect to several bounded-rationality models, our findings also reveal novel features. First, the inability to condition on type-dependent strategies (e.g., Eyster and Rabin,

[15]For instance, Brandts and Holt (1992), Brandts and Holt (1993), Kübler et al. (2008), and Güth and Winter (2013) study deterministic signaling games, while de Haan et al. (2011) and Jeitschko and Normann (2012) incorporate noise in the signaling technology.

[16]Ambuehl and Thysen (2024) and Charles and Kendall (2024) experimentally explore related questions in the context of causal reasoning. Fréchette et al. (2024) and Kendall and Oprea (2024) study how individuals form mental models from raw data. From a theoretical perspective, Ortoleva (2012) studies possible reactions to events whose probability is sufficiently low under the prior beliefs.

2005; Jehiel, 2005; Jehiel and Koessler, 2008; Esponda, 2008; Fong et al., 2025; Cohen and Li, 2022) should not generate biases when senders pool over the same actions. However, we find that such biases persist even in pooling environments, and, if anything, the introduction of type-specific costs—which induce partial separation—tends to improve inference about the opponent's strategy. Models based on level-k or cognitive-hierarchy reasoning (e.g., Nagel, 1995; Stahl II and Wilson, 1994; Camerer et al., 2004; Crawford and Iriberri, 2007; Agranov et al., 2012; Alaoui and Penta, 2016; Alaoui et al., 2020) could rationalize some of these patterns. Yet, even with extensive feedback during the experiment, subjects continue to neglect how strategic behavior affects inference. Finally, our results complement evidence on selection neglect. While prior studies document difficulties in accounting for data selection—both in strategic and non-strategic settings (Esponda and Vespa, 2018; Enke, 2020; Araujo et al., 2021; Barron et al., 2024; Brown and Fragiadakis, 2019; Degan et al., 2023; Farina et al., 2024)—our receivers face a comparable challenge and tend to neglect DGP selection. A key distinction is that accounting for exogenous DGP selection appears substantially easier than reasoning about strategic DGP choices.

2 Theoretical Framework

2.1 Model

We study a communication game between a privately informed sender (he) and an uninformed receiver (she). The sender's type $\theta \in \{\theta_H = 1,\ \theta_L = 0\}$ is drawn with prior $\Pr(\theta = \theta_H) = \mu_0$. The sender's objective is to maximize the receiver's posterior belief that $\theta = \theta_H$; the receiver's objective is to form accurate beliefs.

The sender can communicate with the receiver by selecting a data-generating process (DGP), from which N signals will be produced (i.e. the data observed by the receiver). The sender can choose between two DGPs, A and B. Both DGPs map the state space $\Theta = \{\theta_H, \theta_L\}$ to an outcome $s \in \{0, 1\}$. This implies that, for each type θ, the choice of a DGP induces a probability distribution over s: $Prob(s = 1|\theta, A) = p_A(\theta)$ and $Prob(s = 1|\theta, B) = p_B(\theta)$. We assume that each DGP (A or B) satisfies the monotone likelihood ratio property: $p_i(\theta_H) > p_i(\theta_L)$ for each $i \in \{A, B\}$. For this reason, we refer to $s = 1$ as a positive signal. We also assume that for each $\theta \in \{\theta_H, \theta_L\}$, $p_A(\theta) > p_B(\theta)$, i.e. each type is strictly more likely to draw a positive signal using DGP A compared to DGP B, which allows us to refer to DGP A as the *Easy DGP* and to DGP B as the *Hard DGP*: producing a positive signal is more likely when

DGP A is chosen, no matter the type. Combining the two assumptions, the highest likelihood of drawing a positive signal corresponds to type θ_H choosing the Easy DGP, and the lowest one to type θ_L choosing the Hard DGP.

The other two probabilities can be ranked in either order, but to avoid trivial indifference, we assume that $p_A(\theta_L) \neq p_B(\theta_H)$. This leaves us with two possible rankings. First, $p_B(\theta_H) > p_A(\theta_L)$: no matter the chosen DGP, type θ_H is always more likely to draw a positive signal than type θ_L. Second, $p_A(\theta_L) > p_B(\theta_H)$: given the choice of DGPs, it may be that type θ_H is less likely to draw a positive signal than type θ_L, reversing the potential meaning of the signals about θ. In the rest of the paper and in our experimental analysis, we will refer to these different sets of feasible DGPs as different *information environments*. In addition, we will slightly abuse terminology and refer to cases with $p_B(\theta_H) > p_A(\theta_L)$ as *verifiable information environments*, given the limited sender's discretion over the meaning of the signals, and, for the opposite reason, to cases with $p_A(\theta_L) > p_B(\theta_H)$ as *unverifiable information environments*.

The sender may face a cost of choosing a certain DGP. Such a cost can depend on the type θ and on the DGP i: $c(\theta, i)$. In particular, we will focus on three cases, which induce, respectively, the insights of communication and signaling games:

1. The DGPs have *no cost*, i.e. $c(\theta, i) = 0$ for all $\theta \in \{\theta_H, \theta_L\}$ and $i \in \{A, B\}$;

2. The *cost is asymmetric across types and DGPs*, and it is always higher for the low type. In particular, we assume that type θ_H never pays any cost, while type θ_L pays a positive cost to select DGP i and no cost to select experiment $i' \neq i$—$c(\theta_L, i) = 0$ and $c(\theta_L, i') = c$. We separately consider the case in which $i = A$ and $i = B$.

Once the sender chooses the DGP, N signals are independently drawn from it and directly shown to the receiver. Given the structure of the DGPs, the available evidence is the outcome of a Binomial distribution. For this reason, we can simplify the notation and describe such an outcome as the number of realized successes $\bar{s} \in \{0, ..., N\}$, which we will refer to as *data*.

After observing $\bar{s}$, the receiver makes four guesses about the likelihood of each pair of DGP i and type θ. We can define such guesses through the function $a(\cdot, \cdot|\bar{s}) : \{\theta_H, \theta_L\} \times \{A, B\} \to [0, 1]$, for each $\bar{s} \leq N$, with $\sum a(\theta, i|\bar{s}) = 1$. We can denote the set of functions satisfying this structure as $\mathcal{A}$. For a given realization $\bar{s}$, chosen DGP $\hat{i}$, type $\hat{\theta}$, and guess $a \in \mathcal{A}$, the receiver's payoff is equal to[17]

$$u_R(a|\hat{\theta}, \hat{i}, \bar{s}) = 1 - \frac{1}{2} \sum_{\theta \in \{\theta_H, \theta_L\}, i \in \{A, B\}} \left(\mathbb{1}\{\theta = \hat{\theta}, i = \hat{i}\} - a(\theta, i|\bar{s})\right)^2.$$

[17] This specification matches the implementation of the binarized scoring rule in the experimental design.

Finally, for a given realization $\bar{s}$ and guessing function $a \in \mathcal{A}$, the payoff of the sender of type θ choosing DGP i is equal to

$$u_S(a,\theta,i|\bar{s}) = a(\theta_H, A|\bar{s}) + a(\theta_H, B|\bar{s}) - c(\theta, i).$$

2.1.1 Measure of Informativeness

In our game, the receiver is making an inference over both the sender's type and the chosen DGP. Given the sender's choice and the receiver's guesses, we can measure the amount of information transmitted about any of these variables as the expected reduction of uncertainty induced by communication. We can obtain this measure by constructing two hypothetical expected payoffs for the receiver, which only account for the guesses about the type and about the DGP. Specifically, given $\bar{s}$, a sender's type θ and a chosen DGP $\hat{\imath}$, we can define the following payoff functions: $u_R(a|\hat{\theta},\bar{s}) = 1 - \frac{1}{2}\sum_{\theta\in\{\theta_H,\theta_L\}}\left(\mathbb{1}\{\theta=\hat{\theta}\} - a(\theta|\bar{s})\right)^2$ and $u_R(a|\hat{\imath},\bar{s}) = 1 - \frac{1}{2}\sum_{i\in\{A,B\}}\left(\mathbb{1}\{i=\hat{\imath}\} - a(i|\bar{s})\right)^2$, where $a(\theta|\bar{s}) = a(\theta, A|\bar{s}) + a(\theta, B|\bar{s})$ and $a(i|\bar{s}) = a(\theta_H, i|\bar{s}) + a(\theta_L, i|\bar{s})$. Given these payoff functions, we can derive the expected payoffs $\mathbb{E}_{\bar{s}}\left[\mathbb{E}_{\hat{\theta}}[u_R(a|\hat{\theta},\bar{s})]\right]$ and $\mathbb{E}_{\bar{s}}\left[\mathbb{E}_{\hat{\imath}}[u_R(a|\hat{\imath},\bar{s})]\right]$, which measure the overall informativeness about the sender's type and the chosen DGP.[18]

In our empirical analysis, we can decompose the informativeness into the *information transmitted* by the sender and the *information processed* by the receiver. To measure the information transmitted, we use the expected payoff of a perfectly Bayesian receiver, who correctly accounts for the sender's strategy in the inference (Fréchette et al., 2022; Farina et al., 2024). Instead, to measure the information processed by the receiver, we can use the expected payoff of the receivers in our data, who may, instead, have incorrect beliefs about the sender's strategy and deviate from correct Bayesian updating.

2.2 Equilibrium Analysis

Depending on the set of feasible DGPs, our game can admit multiple equilibria when information is unverifiable. To discipline the set of equilibria, we focus on the class of PBEs in which the receiver's beliefs about θ_H are strictly monotone in $\bar{s}$ (see Appendix A).[19] We derive our

[18] Appendix C provides more details about this measure of informativeness.

[19] While only roughly 50% of our senders hold monotonic beliefs in the treatments with equilibria multiplicity, our refinement selects the empirically optimal sender's behavior—given the receivers' guesses—and the optimal one given the senders' elicited beliefs. This suggests that our refinement is useful to organize equilibrium predictions in a meaningful way.

predictions for each relevant combination of the DGP selection costs.

No Cost. First, assume that $c(\theta, i) = 0$ for all $\theta \in \{\theta_H, \theta_L\}$ and $i \in \{A, B\}$. This case resembles a standard communication environment, where the only constraints faced by the sender are the technological ones imposed by the feasible DGPs.

Proposition 1. *In the absence of DGP selection costs, there exists a unique monotone PBE in which both sender types choose the Easy DGP (DGP A) with probability one.*

Proposition 1 provides us with our first prediction: in the absence of cost, any sender type should select the Easy DGP, to maximize the expected number of positive signals. As a consequence, receivers should interpret more favorable data realizations as stronger evidence of θ_H, and the magnitude of these updates depends on the exact DGP parametrizations.

Cost on the Hard DGP. We can now move to the case in which type θ_L needs to pay a positive cost to choose the Hard DGP (DGP B).

Proposition 2. *Assume that $c(\theta_L, B) = c$.*

1. *In any verifiable information environment, there exists a unique monotone PBE in which both sender types choose the Easy DGP (DGP A).*

2. *In any unverifiable information environment, there exists a monotone PBE in which both sender types choose the Easy DGP (DGP A). Given the specific DGP parameters, there exists a threshold $\bar{c}$, for which the following statements are true. When the cost is high ($c \geq \bar{c}$), there exists a separating monotone PBE in which type θ_H chooses the Hard DGP (DGP B) and type θ_L chooses the Easy DGP (DGP A). If the cost is low ($c < \bar{c}$), and for certain combinations of DGPs, there may exist a monotone PBE in which type θ_H chooses the Hard DGP (DGP B), and type θ_L randomizes.*

Proposition 2 shows that the introduction of strong exogenous frictions in the DGP choice, modeled as a high selection cost, can introduce new meaningful patterns in the unverifiable information environments. In this case, a separating equilibrium completely reverses how the receiver interprets the data, making more positive signals stronger evidence of θ_L. This result is consistent with the one in a signaling game, where the sender's action can provide information about the sender's type. The main difference here is that the sender's action is not observable, and the choice of the DGP can only be inferred through the data realization. Depending on the DGP parametrization, this additional separating equilibrium may be more informative than the pooling one and, as such, preferred by both the receiver and type θ_H.

Cost on the Easy DGP. Finally, we can consider the case in which type θ_L incurs a cost to choose the Easy DGP (DGP A).

Proposition 3. *Assume that $c(\theta_L, A) = c$. Given the specific DGP parameters, there exist two thresholds, $\underline{c} < \bar{c}$, such that the following statements are true. If the cost is low ($c \leq \underline{c}$), there exists a unique monotone PBE in which both sender types choose the Easy DGP (DGP A). If the cost is high ($c \geq \bar{c}$), there exists a unique separating monotone PBE in which type θ_H chooses the Easy DGP (DGP A) and type θ_L the Hard DGP (DGP B). Finally, for intermediate values of c there exists a monotone PBE in which type θ_H chooses the Easy DGP (DGP A) and type θ_L randomizes.*[20]

Similarly to Proposition 2, Proposition 3 shows that introducing an strong friction on the choice of the Easy DGP leads to separation in the sender's choices. Given our assumptions over the structure of the feasible DGPs, this kind of separation induces the most informative outcome about θ_H, with the high type maximizing the probability of a positive signal and the low type minimizing it. This implies that the receiver's posteriors about θ need to react more to extreme realizations of $\bar{s}$ than when no selection cost is introduced.

Summary of Predictions. In the rest of the paper, we focus on the *high cost cases* from Proposition 2 and Proposition 3. Indeed, we parameterize the feasible DGPs and the selection costs in ways that support separating equilibria. While Section 3 provides more details about our choices of the parameters and the corresponding predictions, we can start summarizing some of the main qualitative predictions derived from our theoretical framework.

On the sender's side, data are produced by *strategically selected data-generating processes*. The precise form this selection takes depends on the information environment and the existence of exogenous costs. In turn, receivers are privy to selected data that depends on the information environment and on the existence of exogenous costs. The sender's DGP selection affects the meaning of the data, requiring an *appropriate adjustment in the receiver's choices*.

3 Experimental Design

The experiment implements the theoretical framework from the previous section in a controlled laboratory experiment. The design also includes control sessions in which either the strategic interaction is removed, and receivers perform an updating task without observing the DGP, or

[20] More details about the relevant thresholds for Proposition 2 and Proposition 3 can be found in Appendix B.

in which the strategic interaction is present but the chosen DGP is observable. The role of these control sessions is to separately assess the role of strategic uncertainty and DGP unobservability on the information transmission outcomes.[21]

3.1 Main Strategic Treatments

Overview and Timeline. In our main strategic treatments, subjects are randomly assigned to one of two fixed roles—sender or receiver—which they keep for the entire session. In each round, the Computer randomly draws the sender's type, which in the experiment is represented as a "shape"—the *Circle* sender is type θ_L, while the *Triangle* sender is the type θ_H—with probability 50% each. Conditional on this type, the sender faces a menu of two DGPs (Bags), the DGP A (*Bag A*) and the DGP B (*Bag B*), each containing nine balls and differing in their proportions of purple and green balls. This implies that each combination (θ, DGP) induces a different probability of drawing a purple ball.

The sender selects a DGP, and the computer then draws three balls (signals) *with replacement* from the chosen bag (DGP). We allow for three draws to get a richer variation in the data realizations, which can help us identify different patterns in the subjects' behavior. The receiver observes only the three signal realizations and does not observe either the sender's type or the DGP choice. After observing the signals, the receiver reports a full posterior over the four type–DGP combinations. The report is entered in a 2×2 table whose entries must sum to 100% (see Figure 36 in Appendix I). The interface displays row and column sums to make explicit the implied marginals over the type and the DGP.

Beliefs are incentivized using the binarized scoring rule (BSR) applied to the four-cell report.[22] Upon submission, the report is converted into a score $S \in [0, 100]$, which is then used as the probability of receiving a 100-point bonus in a lottery conducted at the end of the round.[23] The sender's payoff in each round (in points) equals the receiver's stated probability

[21] The experiment was pre-registered on the *AsPredicted* platform. The main treatments and the non-strategic auxiliary can be found here: https://aspredicted.org/brsg-hccb.pdf. The observable DGP treatments have been separately pre-registered as follow-up treatments, and can be found here: https://aspredicted.org/y8av5a.pdf.

[22] Incentive compatibility of the binarized scoring rule is independent of a subject's risk attitude (Allen, 1987; McKelvey and Page, 1990; Schlag and Van der Weele, 2013; Hossain and Okui, 2013).

[23] The formula used to compute the score is provided in a cheat sheet that subjects can access at any time (see Appendix K). During the instruction period, we do not explain the details of the rule to compute the score, but we inform subjects that the way in which this computation is made makes it incentive compatible to honestly reveal their beliefs over the type-DGP pairs. See Danz et al. (2022) for a discussion about the different implementations of the binarized scoring rule in belief-elicitation tasks.

Table 1: Probability of a positive signal (purple ball) for each (θ, DGP) in each information environment.

Verifiable Information			Unverifiable Information		
	θ_H	θ_L		θ_H	θ_L
DGP A / Easy	$\frac{7}{9}$	$\frac{4}{9}$	DGP A / Easy	$\frac{7}{9}$	$\frac{6}{9}$
DGP B / Hard	$\frac{5}{9}$	$\frac{1}{9}$	DGP B / Hard	$\frac{3}{9}$	$\frac{1}{9}$

that the sender type is high—expressed on a 0–100 scale—minus the cost of the selected DGP. Payoffs may be negative in case the cost exceeds the elicited belief of the receiver.

Treatment Variations. For our main strategic treatments, we vary the available DGPs—i.e., our information environments—*between subjects*, and the DGP-selection costs *within subjects*.

Information environment. We vary the compositions of the two DGPs available to each sender type. We parametrize the different DGPs so that, referring to our theoretical framework, DGP A corresponds to the *Easy DGP*, DGP B to the *Hard DGP*, and the purple balls to the *positive signals*. Specifically, Table 1 reports the DGP compositions in each information environment, by summarizing the probability that each combination (θ, DGP) produces a positive signal. In the experiment, this information is presented to subjects as four bags containing a share of purple and green balls, which would be used by the computer to make three draws with replacement (see Figure 35 in Appendix I). As in the theoretical framework, for any DGP (Easy or Hard), type θ_H has a higher probability of drawing a positive signal. Similarly, for each θ, the Easy DGP is more likely to produce a positive signal than the Hard DGP.

As discussed in Section 2.2, the key distinction between our information environments lies in how much discretion the sender has over the induced distributions of data. In particular, we fix the probability of a positive signal induced by the pairs $(\theta_H, Easy)$ and $(\theta_L, Hard)$, and we vary the remaining probabilities to capture the main insights of our verifiable and unverifiable information environments.

Costs. DGP selection costs apply only to the low type and vary across rounds within subjects. All costs are publicly announced to both senders and receivers and reminded at the moment of the choice, together with the available DGPs. We consider three configurations: (i) no cost; (ii)

Table 2: Theoretical predictions for sender's behavior.

Verifiable Information		
No Cost	**Cost on Hard**	**Cost on Easy**
θ_H : Easy DGP	θ_H : Easy DGP	θ_H : Easy DGP
θ_L : Easy DGP	θ_L : Easy DGP	θ_L : Hard DGP
Unverifiable Information		
No Cost	**Cost on Hard**	**Cost on Easy**
θ_H : Easy DGP	θ_H : Easy / Hard DGP	θ_H : Easy DGP
θ_L : Easy DGP	θ_L : Easy DGP	θ_L : Hard DGP

Notes: Each cell reports the predicted behavior for the sender in our treatments. The beliefs of the receiver are computed by taking as given such behavior and applying Bayes' rule.

70 points on the Hard DGP (DGP B); and (iii) 70 points on the Easy DGP (DGP A). For both information environments, the costs we implement are high enough to lead to the separating equilibria discussed in Proposition 2 and Proposition 3. The three cost parametrizations serve the purpose of inducing different strategic incentives within the same environment. The no-cost treatments allow for pure communication, with each type being able to freely choose among DGPs. The introduction of positive and large costs *de facto* restricts the choices of the low type, leading to different equilibrium predictions and, at the same time, potentially reducing the strategic uncertainty receivers face.

Theoretical Predictions Given the Experimental Parameters. Given our experimental parameters, we can further discuss the equilibrium predictions induced by our model (which are summarized in Table 2). As discussed in Section 2.2, when there is no selection cost, both sender types should choose the Easy DGP. Given our DGP parametrizations, this implies that receivers can extract more information about the type in the verifiable information environment, where we should observe a stronger reaction to information.

Introducing a cost on the Hard DGP for θ_L does not change the theoretical prediction in the verifiable information environment. However, this same modification substantially affects the predictions in the unverifiable information case, where type θ_H can separate from θ_L by choosing the Hard DGP and maximizing the expected number of negative signals. This new prediction—which induces a more informative outcome than in the absence of the cost—allows for the meaning of a positive signal to be reversed by the sender's equilibrium choice. This result gives us an effective way to test whether receivers account for the strategic incentives and recognize the impact of the sender's decisions on the interpretation of the data, even when it leads to counterintuitive outcomes.

Finally, in both information environments, the introduction of a cost on the Easy DGP leads to the separating equilibrium in which θ_H chooses the Easy DGP and θ_L the Hard DGP. Given our parameters, this induces an identical outcome across information environments and the highest feasible amount of information transmission about θ and extreme reactions to information. This treatment variation allows us to compare senders' and receivers' choices in the two environments and attribute potential differences to the set of feasible DGPs.

3.2 Auxiliary Treatments

Our design includes auxiliary treatments to control for the effect of strategic DGP selection and unobservable DGP choice.

Non-Strategic Treatments. We conduct non-strategic sessions that recruit only receivers and use the same monetary incentives as in the main strategic treatment. Participants face the same information environments, varied *between subjects*. There is no sender. Instead, receivers are paired with computer "machines", which can be of high or low type (respectively, *Triangle* or *Circle*), and are informed of the DGP-selection rule used by these machines.

We have two sets of non-strategic treatments. In the *random* treatments, each DGP is chosen by each machine type with probability 50%, fixed throughout all rounds. In the *deterministic* treatments, subjects are explicitly told the DGP each machine type will choose (i.e., the selection rule), and we vary these rules *within subjects*. We consider three DGP-selection rules: both types choose the Easy DGP; high type chooses the Easy DGP and the low type chooses the Hard DGP; high type chooses the Hard DGP, and low type chooses the Easy DGP. The implemented rules reproduce the relevant theoretical predictions in each treatment, aiming to approximate the type of selection and the data distribution to which receivers are exposed in the strategic treatments.

Strategic Treatments with Observable DGP Choice. To control for the role of the unobservable DGP in information transmission, we conduct sessions that replicate the structure of the main strategic treatments, but make the sender's DGP choice in each round observable to the receiver. Specifically, the receiver observes whether the sender chooses the Easy or the Hard DGP but remains uninformed about the sender's type, on which the inference is formed. This change substantially affects the theoretical predictions of the game.[24] While multiple equilibria survive our refinement in this modified environment, all of them predict (weakly) greater information transmission about θ. The only exception arises in the unverifiable information environment with a cost on the Easy DGP, where less effective separation can be sustained in equilibrium.

3.3 Procedure and Additional Elicitations

Experimental Procedures. In all treatments, subjects complete 30 rounds. In treatments with within-subject variations, the different treatments are implemented in a block design. Each block—defined by a cost or a DGP selection rule—consists of 5 rounds and is repeated twice. In this way, before playing in the same treatment for the second time, subjects are exposed to all the cost/ selection rules combinations. We vary the block order across sessions to rule out any order effect.

At the end of each round, subjects get full feedback. In particular, they see a summary reporting the sender's or machine's true type, the selected DGP, the three signals drawn, the receiver's guesses, the realized score, and—only in strategic treatments—the sender's payoff. Points earned during the experiment are converted at \$0.01 per point, and participants receive an additional \$10 show-up fee.

To ensure consistency across sessions and treatments, the instructions are delivered through an AI-generated video. In the video, instructions are read out loud and paired with text and images to maximize clarity. After viewing it, subjects receive paper copies of the same material.[25] After the instruction period, subjects complete two practice rounds—one guided and one unguided—designed to illustrate the interface step by step (for both roles in the strategic treatments). The paid role is then revealed on-screen.

[24]The theoretical predictions in this auxiliary case follow the intuition in Perez-Richet (2014). In the absence of costs, this treatment constitutes a simple application of informed information design (see also Koessler and Skreta, 2023). In the presence of costs, this interaction resembles a model of signaling with grades (see Daley and Green, 2014).

[25]See Appendix J for a sample of the instructions in the main strategic treatment.

Table 3: Number of Subjects by Treatment

	Information Environment	
	Verifiable	**Unverifiable**
Strategic Unobservable (Main)	66	62
Random	30	34
Deterministic	32	34
Strategic Observable	52	58

Additional Elicitations. After the practice rounds and the 30 main rounds, subjects complete a series of additional elicitations. Specifically, in the strategic treatments, we elicit—in an incentivized manner—subjects' beliefs about the opponent's strategy. Also, we elicit receivers' preferences over different types of additional information to perform their task and senders' preferences over possible signal realizations.[26] Finally, we elicit risk aversion with two investment tasks following Gneezy and Potters (1997) and, in the strategic treatments, altruism with two dictator tasks.[27] More details about the additional elicitations are reported in Appendix H.

Data Collection. Sessions were conducted at the Columbia Experimental Laboratory for the Social Sciences (CELSS) and the Princeton Experimental Laboratory for the Social Sciences (PExL) in Spring/ Fall 2025. We recruited 368 subjects.[28] Table 3 reports the number of subjects per treatment. Average earnings were roughly $35 for sessions lasting 60–90 minutes. For each information environment, we ran 4 sessions for the main strategic treatment (with unobservable DGPs), 3 sessions for the deterministic treatments, 4 sessions for the observable DGP treatments. For the random treatments, we ran 2 sessions for the verifiable information environment and 3 sessions for the unverifiable information environment. In total, we ran 27 sessions. Strategic sessions had on average 16 subjects, with a minimum of 10 and a maximum of 18.

[26]These additional tasks did not reveal any systematic pattern beyond what is already evident from the main tasks and the standard belief elicitation. In the interest of brevity, they are omitted from the results.

[27]We find no evidence that risk aversion or altruism play a role in explaining the data.

[28]Subjects at Columbia University were recruited through the ORSEE software (Greiner, 2015). The experiment was coded in oTree (Chen et al., 2016).

4 Results

We report results in two parts: Section 4.1 analyzes senders and Section 4.2 analyzes receivers. To account for within-session learning, our main analyses use the second block of each treatment.[29] Focusing on this block ensures that subjects have been exposed to each cost parametrization/ DGP selection rule before playing again in the same treatment.[30]

All statistical analyses include either subject fixed effects or subject random effects, depending on whether we are comparing choices made by the same subjects within the session or choices made by different subjects across sessions. We cluster standard errors at the individual level when making within-subject comparisons (i.e., given an information environment) and at the session level when making between-subject comparisons (i.e., across information environments).[31]

For brevity, while discussing the results, we denote treatments by V for verifiable information treatments and U for unverifiable information ones. Costs are labeled as C_0 (no cost), C_{Hard} (cost on the Hard DGP), or C_{Easy} (cost on the Easy DGP). For example, $V - C_{Hard}$ refers to the verifiable information environment with a cost on the Hard DGP.

4.1 Senders' Behavior

4.1.1 Choice of Data-Generating Processes

We start by analyzing the senders' DGP choices. Our framework predicts that changes in the information environments and in the DGP costs lead to different types of DGP selection. This pattern modifies the distributions of data receivers are exposed to and, as a consequence, their meaning. Table 4 reports the share of Easy DGP choices for each sender type, and it compares them with our theoretical predictions (reported in square brackets). The predictions reported in the table coincide with the empirically optimal ones, given the receivers' guesses. In the $U - C_{Hard}$ treatment, where two monotone equilibria are possible, the receivers' behavior leads to the selection of the separating equilibrium in which θ_H chooses the Hard DGP.

First, we find that behavior qualitatively aligns with theory, as senders mostly select the pre-

[29] Figure 12 and Figure 13 show that modest learning occurs during the sessions, mainly among senders.

[30] We reproduce all the main analyses using all data and obtain the same qualitative patterns (Appendix G). In our sessions, we also vary the order in which costs/DGP selection rules are introduced, keeping no cost/ pooling first in the sequence, and then adding positive costs or separating rules. We find no evidence of order effects.

[31] However, the main qualitative patterns remain robust to different clustering levels.

Table 4: Senders' DGP choices: share of Easy DGP

	Verifiable Information		**Unverifiable Information**	
	θ_H	θ_L	θ_H	θ_L
C_0	0.96	0.91	0.78	0.83
	[1]	[1]	[1]	[1]
C_{Hard}	0.96	0.96	0.48	0.95
	[1]	[1]	[0 / 1]	[1]
C_{Easy}	1.00	0.15	0.85	0.11
	[1]	[0]	[1]	[0]

Notes: Theoretical predictions are reported in square brackets.

dicted DGP. This result shows that strategic DGP selection is a key driver of senders' choices. Specifically, senders internalize both endogenous incentives (receivers' reactions to data) and exogenous ones (selection costs), making the appropriate adjustments across treatments. Quantitatively, however, selection tends to be less extreme than predicted, mainly in the unverifiable information treatment with no selection costs, where senders face more uncertainty over the receivers' responses.

It is important to highlight that the aggregate senders' behavior is consistent with either a pooling or a separating equilibrium, depending on the specific predictions. This suggests that senders tend to appreciate when different DGP choices can serve as a signaling tool, and when, instead, that is not feasible—like, for instance, in the $V - C_{Hard}$ treatment, where θ_H senders cannot attempt to separate by choosing the Hard DGP. If, instead, we focus on the $U - C_{Hard}$ treatment, we can see that high-type senders choose the Easy DGP roughly half of the time. In this treatment, θ_H can signal its type by maximizing the expected number of negative signals, which are relatively more expensive to produce for θ_L. The significant drop in the share of the Easy DGP choices is driven by a large group of senders (roughly 40% of our subjects), who identify such a mechanism and try to exploit the signaling power of the Hard DGP—while the remaining group makes choices that are more noisy or in line with a pooling equilibrium.[32]

To confirm the relevance of the strategic forces in shaping the senders' behavior, we can

[32] Appendix E.3.1 confirms these patterns by studying the distribution of senders' individual choices.

measure how often the senders' elicited beliefs induce the empirically optimal choice.[33] We find that this is overwhelmingly true in the data. In particular, 88% of the subjects in the $V - C_0$ treatment hold beliefs that would lead to the empirically optimal action for both sender types, and the share grows to over 97% with the introduction of a selection cost. In the unverifiable information treatments, the same result holds for 94% of the subjects in the $U - C_0$ treatment and 90% in the $U - C_{Easy}$ one. Finally, 55% of the subjects in the $U - C_{Hard}$ treatment hold beliefs that would support a separating equilibrium. However, 86% of the remaining subjects hold beliefs that would, instead, justify pooling over the Easy DGP.

Sender - Result 1: *Senders' choices of data-generating processes align with theoretical predictions. Strategic DGP selection emerges as a key driver of behavior and shapes the distribution of data across environments.*

4.1.2 Implications of DGP Choice on Inference: Bayesian Guesses

The senders' DGP choices induce different data distributions across treatments. This modifies the type of inference that receivers should form from the data, making it optimal to react to the same data realizations in different ways. We can construct a benchmark for this empirically optimal behavior by computing, for each pair (θ, DGP), the guesses that a fully Bayesian receiver who best responds to the average senders' behavior would make—we will refer to these guesses as *Bayesian guesses*, and they will constitute our *Bayesian benchmark*. Figure 2 reports the Bayesian guesses for each realization of the data.

First, optimal guesses vary across information environments and cost parametrizations in ways that are consistent with the theory. In particular, in both information environments, we detect, on average, a significant decrease (p-value < 0.01) in the guess associated with θ_L and the costly DGP compared to the treatment without a cost. Consistent with the introduced signaling opportunity, we find, on average, a significant increase (p-value < 0.01) in the guess associated with the pair $(\theta_H, Hard)$ in the $U - C_{Hard}$ treatment, compared to the other cost parametrizations—especially for low $\bar{s}$, which are now optimally interpreted as good news. Finally, in both $V - C_{Easy}$ and $U - C_{Easy}$, the senders' behavior induces more extreme beliefs about θ than when costs are absent, and this gap is larger for the unverifiable information environment. As predicted by our theory—and as expected given the senders' choices—such posteriors are extremely close across information environments.

[33] Figure 16 and Figure 17 in Appendix E.3.2 report the senders' elicited beliefs and compare them with the average receivers' behavior.

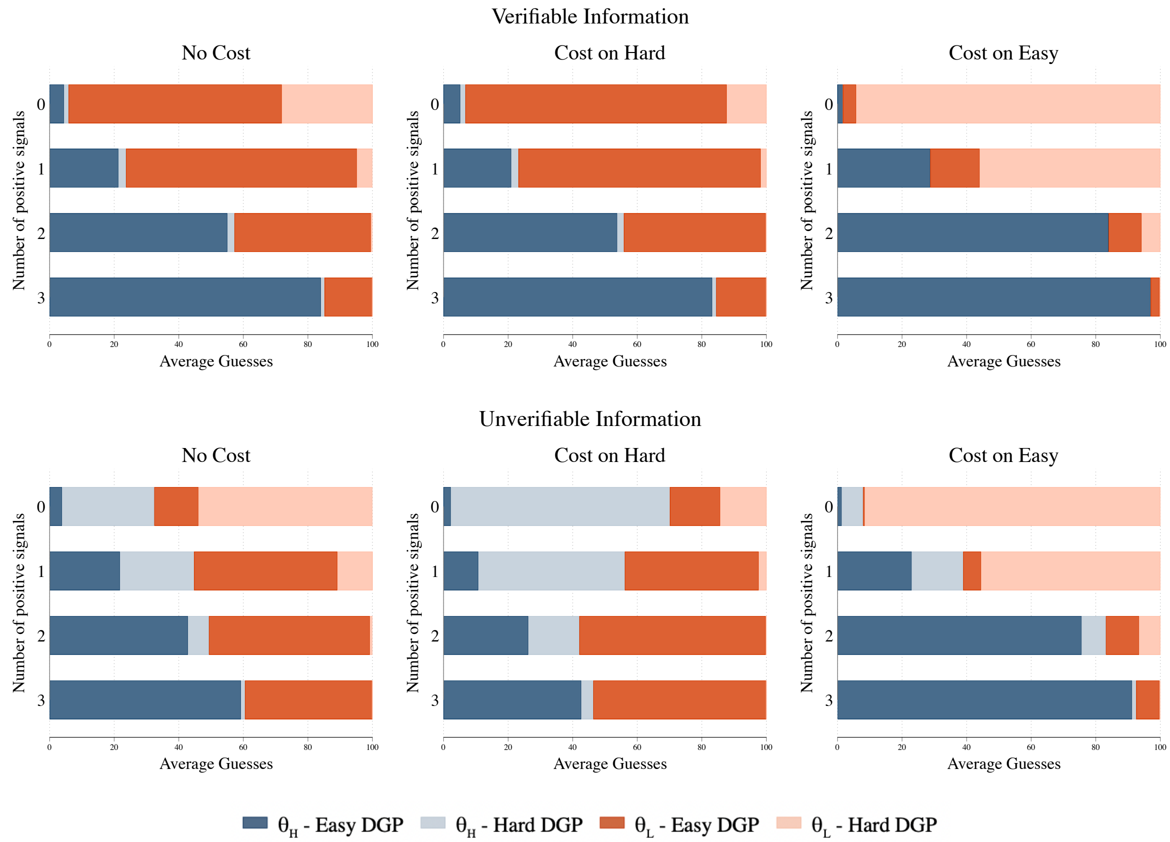


Figure 2: Bayesian guesses in the strategic treatments.

Sender - Result 2: *The sender's selection of data-generating processes affects the correct inference of the data, which varies across information environments and cost parametrizations.*

4.1.3 Summary and Implications of Sender's Behavior

Section 4.1.1 shows that the sender's behavior is broadly consistent with the theoretical predictions. We find that senders respond to both endogenous and exogenous incentives—respectively, the receivers' guesses and the exogenous cost. This pattern is qualitatively consistent with the standard communication literature, which usually documents senders' ability to adapt their choices to the receivers' behavior. However, our data shows that such a finding extends to a setting in which senders do not have direct control of the evidence, but need to reason in terms of distributions. This is a key observation, confirming that selection is likely to bias the data subjects are exposed to, even when shaping distributions is the only feasible manipulation.

Also, we do not detect any particular pattern other than the equilibrium forces, suggesting that DGP selection is not affected by other types of considerations, such as preferences towards honesty, which have been documented as an important force in the communication literature.[34] Choosing the most extreme DGPs would lead to the most informative outcome for the receiver, but, unless this is predicted by the theory, we do not find evidence of this pattern in the data.

Overall, the senders' behavior generates the type of variation in the data that we expected given the theoretical predictions. This provides us with the appropriate tool to study the receivers' inference in the different environments and their reaction to selected DGPs. Section 4.2 discusses our receivers' results. Combining the behavior of both players, Section 5 assesses the impact of the unobservable DGP selection on information transmission.

4.2 Receivers' Behavior

4.2.1 Receivers' Inference in the Exogenous Treatments

The receiver's task in the main experiment is demanding: it combines a strategic step—inferring the sender's DGP choice—and a probabilistic step—updating beliefs from the realized data. Even if receivers perfectly anticipate sender choices, they must still update along two dimensions, the state θ and the DGP, which can generate systematic mistakes.[35]

To assess how much probabilistic mistakes matter in our setting, we study receiver behavior in non-strategic treatments that remove the need to anticipate sender choices. In the *deterministic* treatments, the type-to-DGP mapping is known and mirrors the possible equilibrium combinations in our main strategic treatments. In the *random* treatments, each type selects each DGP with 50% probability. All subjects were carefully informed about these selection rules during the session.[36]

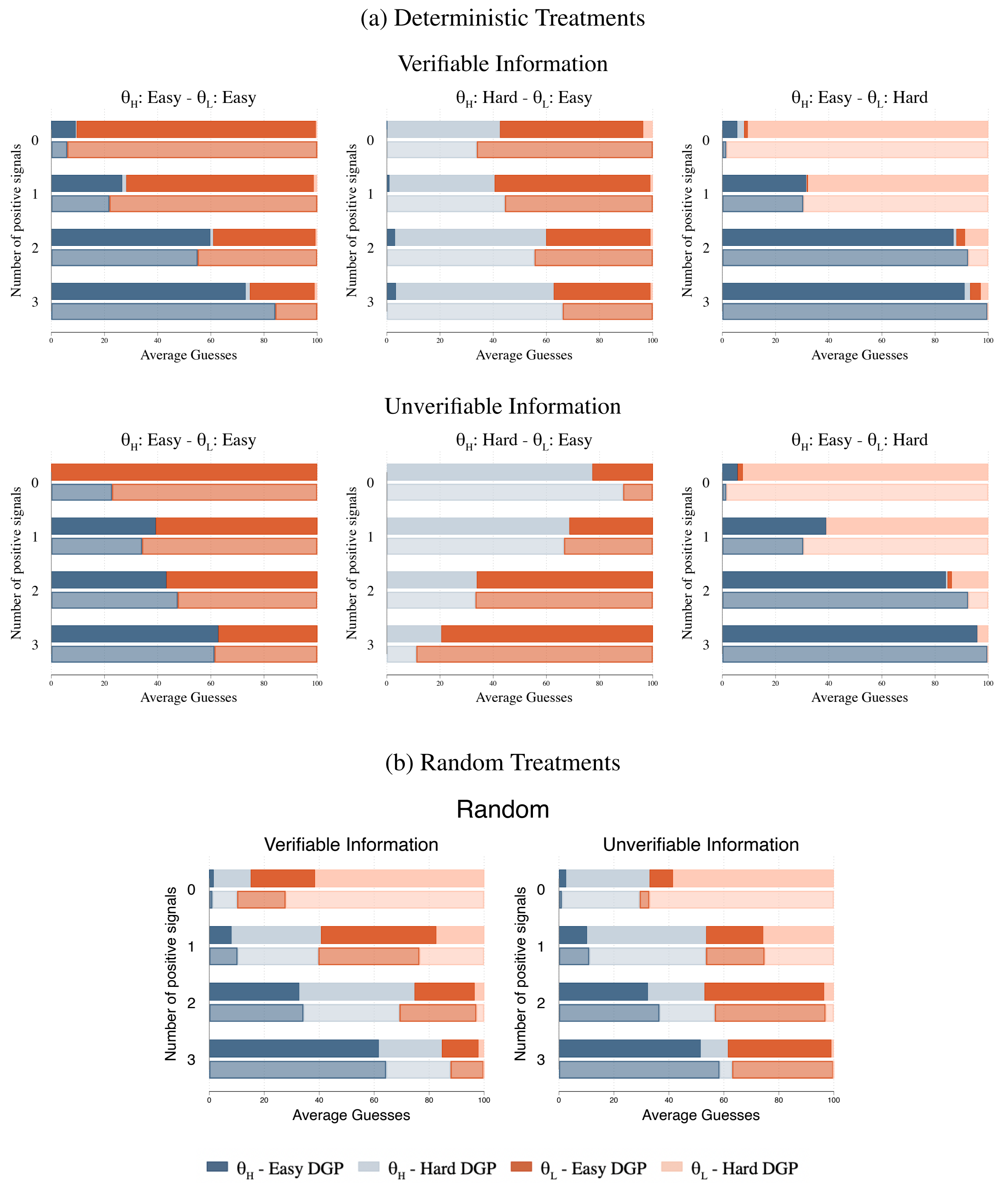
Figure 3 reports the average receivers' guesses in non-strategic treatments (top/darker bar for each $\bar{s}$) and compares them to the Bayesian guesses (bottom/lighter bar for each $\bar{s}$).[37] Subjects' aggregate inference is qualitatively consistent with the Bayesian benchmark. Said otherwise, in

[34] Abeler et al. (2019) conduct a meta-study of the existing experimental literature and document the prevalence of lying aversion in cheap-talk settings. Similarly, Farina et al. (2024) documents an aversion to deceive the receivers in an environment in which senders cannot lie, but only select among available evidence. See Sobel, 2020 for a theoretical distinction between lying and deceiving in games.

[35] See Woodford (2020) and Enke (2024) for reviews on complexity in belief formation. Complexity often leads to attenuation, though overreaction can also occur (Ba et al., 2024). Aina and Schneider (2025) and Liang (2025) document deviations from Bayesian updating when beliefs are formed over multiple DGPs.

[36] See Appendix I.2 for more details about the non-strategic interfaces.

[37] While, in each information environment, the number of total observations for each DGP selection rule is the same, the number of observations for each $\bar{s}$ varies given the different senders' choices.



Notes: For each number of positive signals, the top (darker) bar represents receivers' observed behavior, while the bottom (lighter) bar represents the Bayesian guesses implied by the DGP selection rule. Realizations of $\bar{s} = 0$ in the unverifiable information treatments with θ_H : Easy $- \theta_L$: Easy are extremely unlikely, and we have only one observation in the sample reported in this Figure.

Figure 3: Average receiver guesses in the non-strategic treatments.

the exogenous treatments—where the type to DGP mapping is known—we see no systematic mistakes: some data realizations show small deviations, but average posteriors over the type and DGP are well calibrated.

This result reveals two important insights. First, subjects know how to account for deterministic DGP selection, which well approximates the senders' choices in the experiment for most treatments, when the selection rule is known. This is in contrast with some of the results in the data selection-neglect literature, where neglect of unobserved data tends to be an issue also in the absence of endogenous or strategic selection (e.g., Enke, 2020). Second, when DGPs are randomly selected, subjects seem to properly account for the likelihood of the different DGPs, suggesting that there is no systematic tendency in overweighting the most likely DGP, as is possibly the case for conflicting processes (e.g., Aina and Schneider, 2025).

Receiver—Result 1: *In the exogenous treatments, the aggregate inference over types and DGPs reveals small deviations, with no systematic bias.*

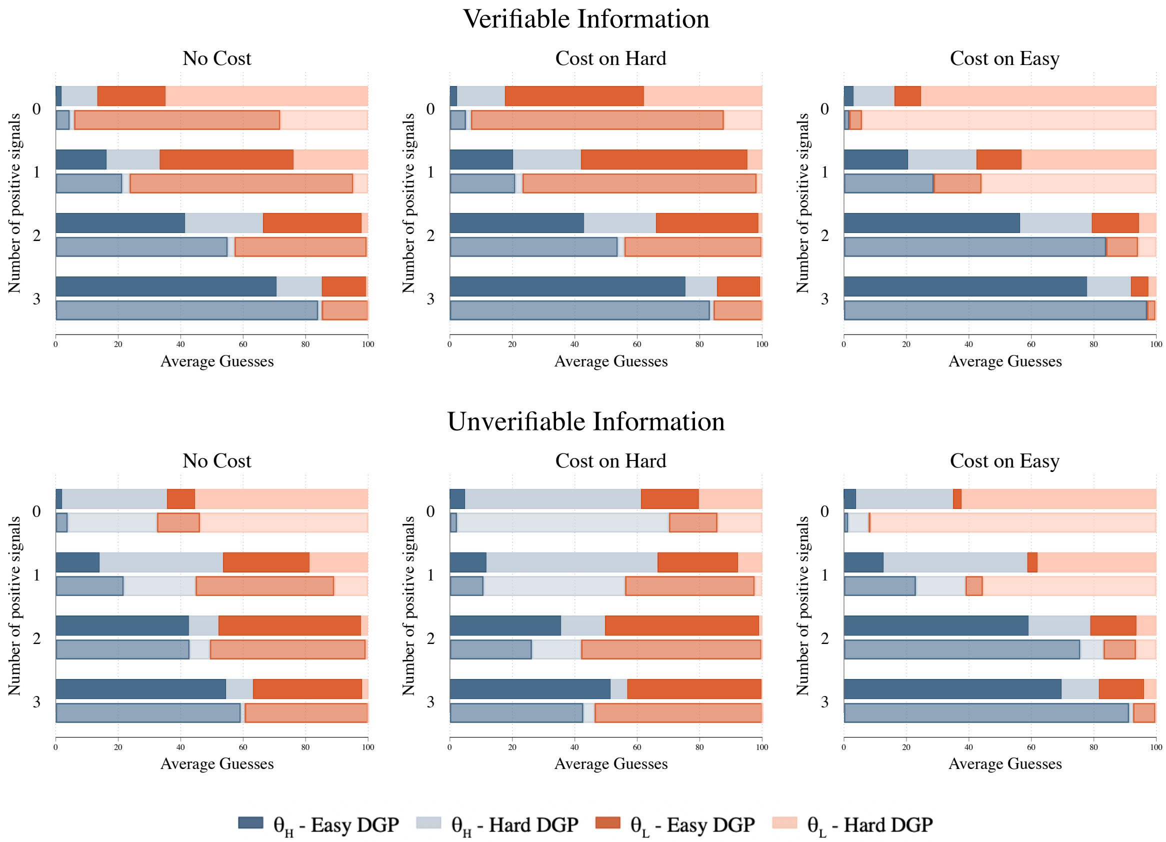
With this benchmark, we now turn to the strategic treatments, where we investigate how receivers' inference adapts to the senders' strategic DGP choices.

4.2.2 General Patterns in Receivers' Inference

Figure 4 reports, for each treatment, the average receiver's guess for each combination (θ, DGP) given the number of realized positive signals.[38] The top/darker bar summarizes our data, while the bottom/lighter bar reports the Bayesian guesses from Figure 2. The results reveal several patterns in the receivers' inference.

First, receivers adapt their guesses to the realized evidence and to the information environments in ways that are qualitatively consistent with the Bayesian benchmark. In particular, when a cost is imposed on a given DGP, receivers reduce, on average, the probability attached to the combination of that DGP and type θ_L (p-values < 0.01). This suggests a correct reaction to the introduction of exogenous incentives. Also, comparing $U - C_{Hard}$ with $U - C_0$, we find an increase in the likelihood attached to (θ_H, Hard), mainly for low realizations of $\bar{s}$. This implies that, after a low $\bar{s}$, receivers are more optimistic about θ in $U - C_{Hard}$ than in the other cost parametrizations (in the comparison with $U - C_0$, p-value < 0.01 for $\bar{s} = 0$ and p-value

[38]While, in each information environment, the number of total observations for each cost parametrizations is the same, the number of observations for each $\bar{s}$ varies given the different senders' choices. The patterns revealed by Figure 4 are consistent with the ones in Figure 33, which include all the data.



Notes: For each number of positive signals, the top (darker) bar represents receivers' observed behavior, while the bottom (lighter) bar represents the optimal guesses implied by the senders' behavior.

Figure 4: Average receiver guesses in the strategic treatments.

< 0.05 for $\bar{s} = 1$; in the comparison with $U - C_{Easy}$, p-value < 0.05 for $\bar{s} = 0$). This reveals a general understanding of the signaling opportunities available to type θ_H. Finally, in each information environment, when a cost on the Easy DGP is imposed, receivers tend to make significantly more extreme guesses about θ after high realizations of $\bar{s}$ compared to the other cost parametrizations (in the verifiable information treatment, p-value < 0.01 for $\bar{s} = 2$ and p-value < 0.1 for $\bar{s} = 3$; in the unverifiable information treatment, p-values < 0.01 for $\bar{s} = 2$ and $\bar{s} = 3$). This suggests a correct tendency to extract more information from more informative data.

Despite these positive patterns in the receivers' aggregate behavior, we also detect systematic deviations from the Bayesian benchmark. To comment on these deviations, we can consider the average gap between the guesses in our data and the Bayesian benchmark we constructed. This gap gives us an indication of whether subjects tend to systematically overweight or underweight the likelihood of a certain type-DGP combination. First, we find that receivers tend to overestimate the likelihood of (θ_H, Hard) and underestimate the one of (θ_H, Easy) ($p-$values < 0.01). This pattern is present in all treatments, except for $U - C_{Hard}$, where type θ_H chooses the Hard DGP more often: here we find that, on average, subjects tend to overweight (θ_H, Easy) ($p-$value < 0.01). Notably, the nature of these deviations is consistent across treatments. Overall, these results suggest that receivers struggle to incorporate θ_H's strategic incentives—which only need to account for the receivers' responses—into their inference, assessing the high type's DGP selection as less extreme than it actually is.

The same qualitative pattern applies to the inference about θ_L when selection costs are absent. Indeed, we find that, on aggregate, receivers tend to place too little weight on (θ_L, Easy) and too much on (θ_L, Hard) ($p-$values < 0.01), confirming that in the absence of costs, strategic selection is not fully internalized. It is worth mentioning that deviations from the Bayesian benchmark appear to be much smaller in $U - C_0$ than in $V - C_0$. This is partially due to the less extreme sender's behavior (see Table 4), which makes it optimal, from an empirical perspective, to account for both DGPs when assessing the data. However, it is still true that receivers underestimate the extent of the strategic selection, even if the misassessment is less serious.

As already mentioned, exogenous costs significantly change receivers' guesses for type θ_L. Indeed, while we still detect some systematic biases, their nature tends to change due to the correct reaction to costs. In $V - C_{Hard}$, for example, the gap from the Bayesian benchmark for (θ_L, Hard) narrows, and the same is true for (θ_L, Easy) in the treatments with a cost on the Easy DGP. Overall, this suggests that receivers struggle more in accounting for endogenous strategic incentives than for exogenous ones when making their inference.

Finally, the discussed biases imply that, in all our strategic treatments, subjects significantly overweight the probability of type θ_H and underweight the likelihood of the Easy DGP ($p-$values < 0.01). These results are in contrast with the ones discussed in Section 4.2.1, where we did not detect any systematic bias in the receivers' inference: this supports the insight that deviations are not the outcome of incorrect probabilistic inference, but of strategic mistakes. Comparing Figure 4 with Figure 3 reinforces this point: the observed deviations in our strategic treatment mirror the guesses made in the random (non-strategic) treatment—which are extremely close to the predicted one given the random DGP selection. This allows us to identify the Bayesian guesses following from random DGP selection as a second benchmark to interpret our data—and we will refer to that as the *Random DGP benchmark*.

Receiver - Result 2: *On average, receivers adapt their guesses to the different information environments and to the different selection costs. However, their inference presents systematic biases that are consistent with a partial neglect of DGP selection.*

4.2.3 Receivers' Heterogeneity

Section 4.2.2 shows how the aggregate behavior of the receivers deviates from the Bayesian benchmark in ways that resemble the Random DGP selection benchmark. However, our aggregate analysis does not reveal whether the documented deviations are consistently made by all receivers in our data, or whether there is some meaningful heterogeneity. Exploiting the within-subjects variations in the cost parameters, we can assess receivers' behavior across treatments. Doing that allows us to capture trends in the behaviors that go beyond a single task, but that adapt to the different strategic environments receivers are exposed to.

The ideal way to perform this analysis would be to assess the inference that each subject makes after observing a realization $\bar{s}$ in each cost parametrization. However, this approach is not feasible in our environment for several reasons. First, this would involve classifying subjects over 16 values (4 distinct guesses for each possible realization of the positive signals) per treatment, making the problem extremely high-dimensional. Second, the constraints on the number of rounds that we can implement and the randomness of the data realizations limit our ability to have a complete combination of data for each subject. Third, subjects have potentially seen different realizations of $\bar{s}$ throughout the session, preventing us from simply considering their average behavior in the session in the absence of any benchmark. To address these challenges, we define two measures that quantify the consistency of the subjects' inference to the Bayesian and to the Random DGP benchmarks, which emerged as meaningful reference points.

Measures of Consistency to the Bayesian and the Random DGP Benchmark. For each subject, in each treatment, we compute the following measures:

$$\text{Total Distance from Bayesian} = \frac{1}{2}\mathbb{E}\left[\sum_{\theta,i} |\bar{a}(\theta,i|\bar{s}) - a^*(\theta,i|\bar{s})|\right]$$

$$\text{Total Distance from Random DGP} = \frac{1}{2}\mathbb{E}\left[\sum_{\theta,i} \left|\bar{a}(\theta,i|\bar{s}) - a^R(\theta,i|\bar{s})\right|\right]$$

where, given a combination (θ, DGP) and a realization $\bar{s}$, $\bar{a}(\theta, \text{DGP}|\bar{s})$ represents the individual average guess, $a^*(\theta, \text{DGP}|\bar{s})$ the Bayesian guess and $a^R(\theta, \text{DGP}|\bar{s})$ the guess consistent with random DGP selection.[39] The two distances can be interpreted as the average number

[39] In the definition, we use the subjects' average guesses for a certain realization of $\bar{s}$ rather than the guesses taken in each round to mitigate the role of noise in the distances we compute. Our results are broadly consistent with an alternative definition using the guesses in each round.

of percentage points receivers misallocate in their inference compared to the benchmark under consideration. If a receiver is fully Bayesian and accounts for the sender's DGP selection, the *Total Distance from Bayesian* is 0, while the *Total Distance from Random DGP* is positive. Similarly, a subject behaving as if DGPs were randomly selected would present a positive *Total Distance from Bayesian* and null *Total Distance from Random DGP*. For brevity, in the rest of the Section, we will refer to the two distances simply as *Distance from Bayesian* and *Distance from Random DGP*, and we use the *Total* label when it is needed to avoid any confusion.

Table 9 in Appendix E.5 summarizes the value of these benchmarks for the strategic and the exogenous treatments. Two insights are worth mentioning. First, we find that the distance from the Bayesian benchmark is significantly higher in the strategic treatments than in the exogenous ones, and the magnitude of this gap is substantial in every information environment and cost parameterization. Second, we find that the distance from the random DGP benchmark is significantly higher in the exogenous treatments—where it is close to the values consistent with the Bayesian guesses—than in the strategic ones, and again, the magnitude of this gap is substantial. These observations are consistent with the analysis in Section 4.2.1 and Section 4.2.2, and convey similar messages as the aggregate receivers' behavior.

Classification of Receivers' Behavior. Given our two distance measures, we can summarize the behavior of each subject with 6 numbers, 2 for each treatment. Specifically, in our classification exercise, we use, for each subject, the average distance from the Bayesian benchmark and the gap between this measure and the average distance from the random DGP selection benchmark in each treatment. Using the difference between the two distances is preferable to using the second benchmark alone for several reasons. First, given the senders' behavior, the two benchmarks are similar in the unverifiable information environment without costs, and using this gap helps differentiate behavior. Second, while each distance should be close to zero in case subjects follow the predicted behavior—either the Bayesian or the random DGP one—their values could vary depending on the signal realizations and on alternative behavior subjects may follow. Again, using the gap as an input for the classification exercise, together with the Bayesian benchmark, would directly convey information on which benchmark is more descriptive of a subject's behavior.

For our classification exercise, we use a hierarchical clustering—Ward's linkage hierarchical clustering with Euclidean distance—and as inputs, the six values discussed above. This approach does not require us to specify a fixed number of groups, but it forms clusters by progressively merging individuals in a way that minimizes the increase in the input variance within

Table 5: Average total distance for each receiver cluster, both relative to the Bayesian benchmark and to the Random DGP benchmark.

	Verifiable Information					
	Equilibrium (15%)		Random DGP (33%)		Reactive (52%)	
	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP
C_0	8.89	38.02	40.58	15.78	32.33	26.39
C_{Hard}	7.26	40.30	40.33	16.22	24.15	30.79
C_{Easy}	13.70	37.35	43.81	20.58	22.86	29.61
	Unverifiable Information					
	Equilibrium (16%)		Random DGP (39%)		Reactive (45%)	
	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP
C_0	14.69	17.92	20.11	14.17	13.14	15.63
C_{Hard}	20.07	30.76	24.80	17.64	17.92	23.64
C_{Easy}	14.35	41.45	42.31	22.04	27.28	24.54

clusters. This guarantees that the algorithm produces cohesive groups that capture systematic patterns in the behavior.

For each information environment, we identify three clusters of receivers: an "*equilibrium*" cluster ($15-16\%$ of the subjects), i.e. a group of receivers who tend to make the correct inference in each cost parametrization; a "*random DGP*" cluster ($33-39\%$ of the subjects), i.e. a group of receivers who tend to behave as if the data generating processes were randomly selected; a "*reactive*" cluster ($45-52\%$ of the subjects), i.e. a group of receivers whose inference is not consistent with the Bayesian benchmark, but who can at least partially react to the strategic forces introduced by the costs for the low type.[40]

Table 5 reports the average distance from the Bayesian benchmark and from the random DGP benchmark for each receiver group. The behavior in the equilibrium group tends to induce

[40]We classify receivers in three groups since increasing the number would not induce meaningful changes in the interpretation.

lower distances from the Bayesian benchmark and higher distances from the random DGP benchmark than the behavior in the random DGP group. The reactive group, instead, presents intermediate values. We can more systematically compare the values of the total distances through OLS regressions in which we control for the distribution of the realized signals—that could potentially affect the magnitude of the deviations.

Focusing on the verifiable information case, we can see that the distance from the Bayesian benchmark for the equilibrium group is strictly below that of the other two groups (p-values < 0.01). Also, if we compare the distance from the random DGP benchmark, we find that it is strictly higher in the equilibrium group than in the others ($p-$values < 0.01), with the only exception being the reactive group in $V - C_{Easy}$. Comparing the reactive and the random DGP group, we find again a significant difference in the distance from the Bayesian benchmark, which is higher in the second one (p-values < 0.01 for the treatment with a cost and p-value < 0.05 for $V - C_0$). Consistently, we find that the distance from the random DGP benchmark is strictly smaller for the random DGP group (p-values < 0.01 for $V - C_0$ and $V - C_{Hard}$ and p-value < 0.05 for $V - C_{Easy}$).

We can perform the same type of analysis with the unverifiable information treatments. Given the behavior of the sender in this setting, we can expect the cleanest comparisons to apply for the $U - C_{Easy}$ treatment. This is, indeed, the case, with all the expected rankings being satisfied (p-values < 0.01, except for the random DGP distance between reactive and random DGP group).

For the other cost parametrizations, the comparisons, even if meaningful, are noisier. First, a weaker variation across groups is detected in the $U - C_0$ treatment, since the Bayesian and random DGP benchmark tends to be closer. However, we find that the distance from the Bayesian benchmark in the random DGP group is significantly higher than in the other groups (p-value < 0.1 for the equilibrium group and p-value < 0.05 for the reactive group), while we detect no significant difference for the distance from the random DGP benchmark. More complex is the analysis of the $U - C_{Hard}$ treatment. While the only statistically significant comparison for the distance from the Bayesian benchmark is between the random DGP and the reactive group (p-value < 0.05), the value for the equilibrium group needs to be more carefully interpreted. Two receivers assigned to the equilibrium group by our algorithm behave fully consistently with the separating equilibrium—and so with partially correct equilibrium forces—and this inflates the deviations from the Bayesian benchmark due to the senders' more intermediate behavior. Finally, we find that the distance from the random DGP benchmark in the random DGP group is significantly smaller than in the others (p-value < 0.1 for the equilibrium group and p-value

< 0.05 for the reactive group).[41]

Receiver - Result 3: *Across both information environments, receivers' behavior clusters into three consistent groups. The smallest group behaves in line with equilibrium predictions. A large group assesses data as if they were generated by randomly selected DGPs. Finally, introducing exogenous costs improves strategic inference for a substantial share of subjects.*

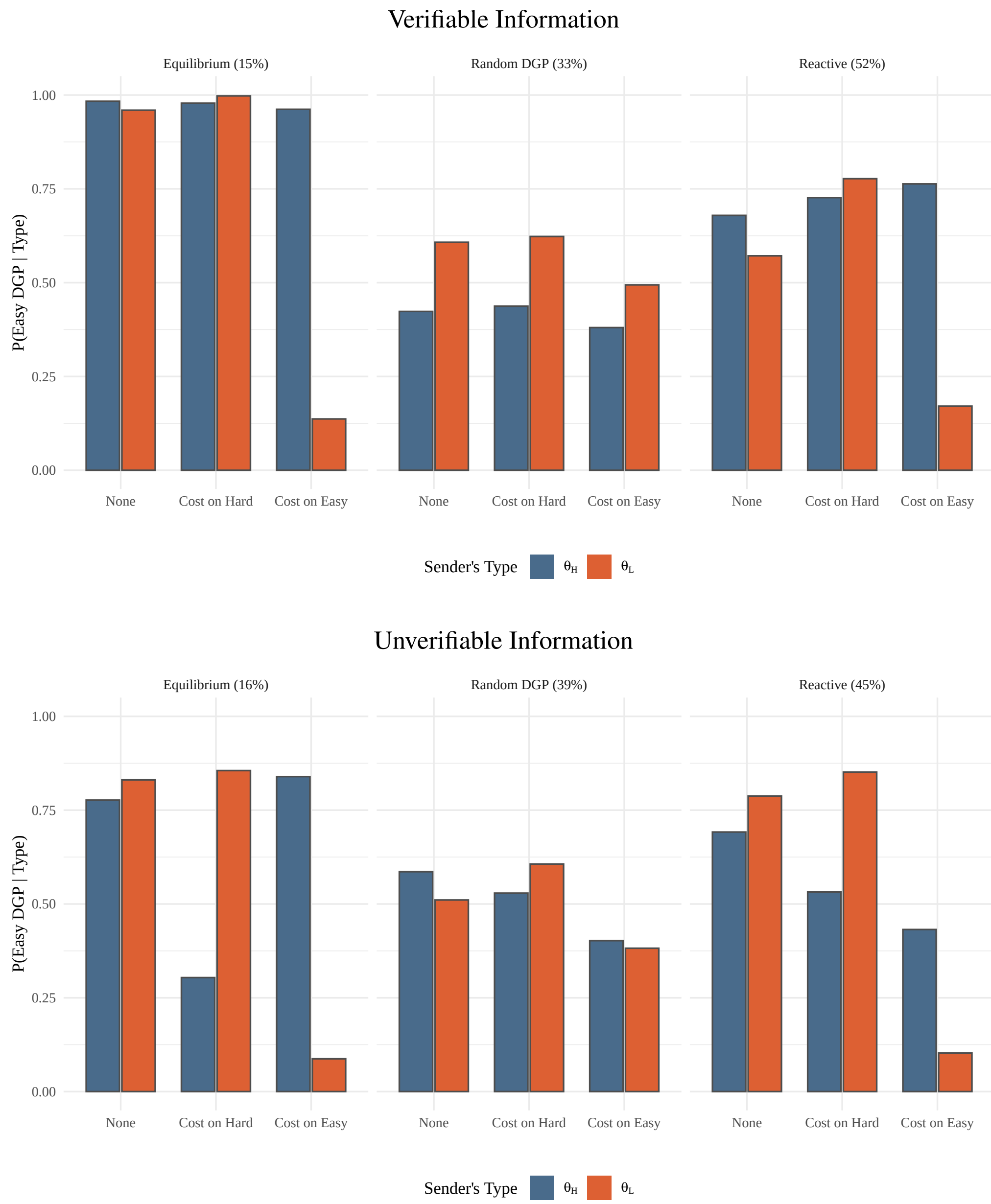
Heterogeneity in Receivers' Beliefs. We interpret the observed heterogeneity across receivers as reflecting differences in their beliefs about the relevant DGPs. To further validate this interpretation—and our classification—we examine receivers' beliefs about the sender's strategy. While our design includes an elicitation of these beliefs, we acknowledge that this measure may be noisy and may not perfectly capture what subjects thought while playing the game. First, we elicit beliefs only at the end of the experiment to avoid signaling to subjects the importance of strategic reasoning in their inference. Hence, asking the question may mechanically improve their responses. Second, receivers perform the belief-elicitation task only once, which increases the influence of noise in their answers.

To address these concerns, we construct a second measure of receivers' beliefs based on their guesses in the main task. Using a simple structural model, we estimate, within each treatment, the beliefs about senders' behavior that would rationalize those guesses. This approach not only yields an alternative belief measure but also offers a more compact summary of receivers' behavior to highlight heterogeneity across groups. The goal of our estimation is to retrieve, for each treatment, the perceived probability that each sender type chooses the Easy DGP. In each round, each subject makes four guesses, which can be described through Bayes' rule. The specification from Grether (1980) allows us to control for individual over-reaction and under-reaction to information. We cannot separately estimate the receivers' beliefs and the possible tendency to over-/under-weight the prior in the inference. For this reason, in our estimation, we set this parameter to 1. Given this theoretical benchmark, we use a Minimum Distance Estimator to retrieve the beliefs that best rationalize the observed guesses.[42]

Figure 5 summarizes the results of our estimation by reporting the average probability each receiver cluster attaches to the sender's choice of the Easy DGP. The pattern of our estimated beliefs captures systematic changes in the assessment of the sender's strategy, confirming our interpretation of the three clusters. Figure 26 in Appendix E.6.2 repeats the same analysis using

[41] Figure 20 and Figure 21 in Appendix E.6 present the average guesses by receiver group for each treatment.

[42] More details of our estimation are reported in Appendix D.



Notes: The top panel reports estimated beliefs in the verifiable information treatments, and the bottom panel the estimated beliefs in the unverifiable information treatments.

Figure 5: Average estimated beliefs over the Easy DGP by receiver cluster.

the elicited beliefs. Reassuringly, we find a positive correlation between our estimated beliefs and the elicited ones, and the same type of patterns across groups emerge.[43]

While this interpretation is consistent across information environments, it is important to

[43]Figure 24, Figure 25, Figure 27, and Figure 28 report the distribution of estimated and elicited beliefs for each treatment. Despite some heterogeneity across participants, we find that our interpretation remains robust.

highlight two main differences, which are consistent with Table 5. The first one concerns the accuracy of the reactive subjects in the treatments without costs. While in the verifiable information environment, the estimated beliefs of this group are closer to those of the random group, in the unverifiable information environment, the estimated beliefs tend to be closer to those of the equilibrium group. This fact is consistent with a less extreme sender's selection, which makes optimal beliefs more intermediate.

The second difference refers to the way in which subjects react to costs in the two environments. In the verifiable information treatment, receivers in the reactive group have more accurate estimated beliefs about both the high and the low types' behavior than those in the random group. In the unverifiable information environment, the improvement mostly concerns the low type, for which the cost is imposed—and this is particularly evident in the $U - C_A$ treatment. This pattern is likely the result of the larger residual uncertainty due to the unverifiability of the information, which can make inference about the high type more challenging.

Receiver - Result 4: *Receivers' systematic deviations in the strategic treatments can be attributed to biased beliefs about the sender's strategy.*

Robustness of the classification. Our receiver's classification survives multiple robustness checks. First, given the randomness of $\bar{s}$, which depends on the sender's behavior, one could be concerned about the effect of small sample biases on the computation of our benchmarks. To account for this possible issue, we repeat the clustering exercise on the whole sample, and/or we compute the average distances by reweighting the data realizations according to their frequency in each treatment. The clustering exercise yields very similar groups, to which the same interpretation discussed above applies.[44]

In addition, regarding the inputs of our algorithm, similar patterns would be detected if we only used the gap between our distance measures or the two measures separately. However, these two approaches would provide noisier classifications in the unverifiable information treatments, where the sender's behavior leads to more similarity between the Bayesian and random DGP benchmark, making the issues discussed above more likely to bind.

To highlight the validity of our clustering algorithm, it is worth noting that we obtain a very similar classification when subjects are assigned to groups manually, based on their distance from the Bayesian and random DGP benchmarks. Although this procedure is more arbitrary—

[44] Also, we do not detect any significant difference in the sample of evidence each cluster is presented, confirming that the data realizations are not affecting the results.

since it requires specifying a rule for group assignment[45]—it has the advantage of being directly applicable to the non-strategic treatment under the same criterion. Consistent with our earlier discussion, while the manual classification identifies three distinct groups in the strategic treatments, it classifies between 85% and 90% of subjects in the non-strategic treatments as equilibrium, confirming the absence of meaningful heterogeneity in the non-strategic environments discussed in Section 4.2.1.

Finally, since our estimated beliefs allow us to summarize subjects' behavior more compactly, we can use them to derive a final robustness check of our classification. We use the six individual beliefs derived from each information environment as the inputs of a k-medoids clustering algorithm from which we derive, again, three groups. The classification we get with this alternative method is broadly consistent with the ones derived through our distance measures.

4.2.4 Decomposing Receivers' Mistakes

Our analysis of receivers' heterogeneity indicates that a substantial share of the deviations in their inference can be attributed to biased beliefs about DGP selection. Consistent with the literature on Bayesian updating, we also detect individual biases in the inference that can be attributed to probabilistic biases.[46] Using our two measures of beliefs, we can try to quantify the impact of these different mistakes in our environments.

Specifically, we construct two counterfactuals for receivers' guesses, computing alternative Bayesian responses that use our estimated or elicited beliefs as priors over the sender's strategy.[47] For each treatment, we then calculate the implied total distance from the Bayesian benchmark and compare it with that observed in the data (Figure 6). In both counterfactuals, we find no substantial reduction in total distance, confirming that biased beliefs play a major role in explaining receivers' deviations.[48]

4.2.5 Why Do Receivers Form Wrong Beliefs?

There are multiple possible explanations for receivers' incorrect beliefs. Several behavioral models study the implications of some form of misspecification in subjects' mental models,

[45]We classify subjects as equilibrium if their average distance from the Bayesian benchmark is below 25% in all cost parametrizations, and as random DGP if their distance from the random DGP benchmark is below 25% in all cost parametrizations. When both conditions hold simultaneously, we compare the two measures to break ties.

[46]These biases have been shown to co-exist and be orthogonal to each other (Geng and Guan, 2025).

[47]Figure 29 summarizes the guesses implied by these counterfactuals.

[48]Appendix E.6.4 reports the distribution of changes in total distance, given a certain counterfactual, by receiver.

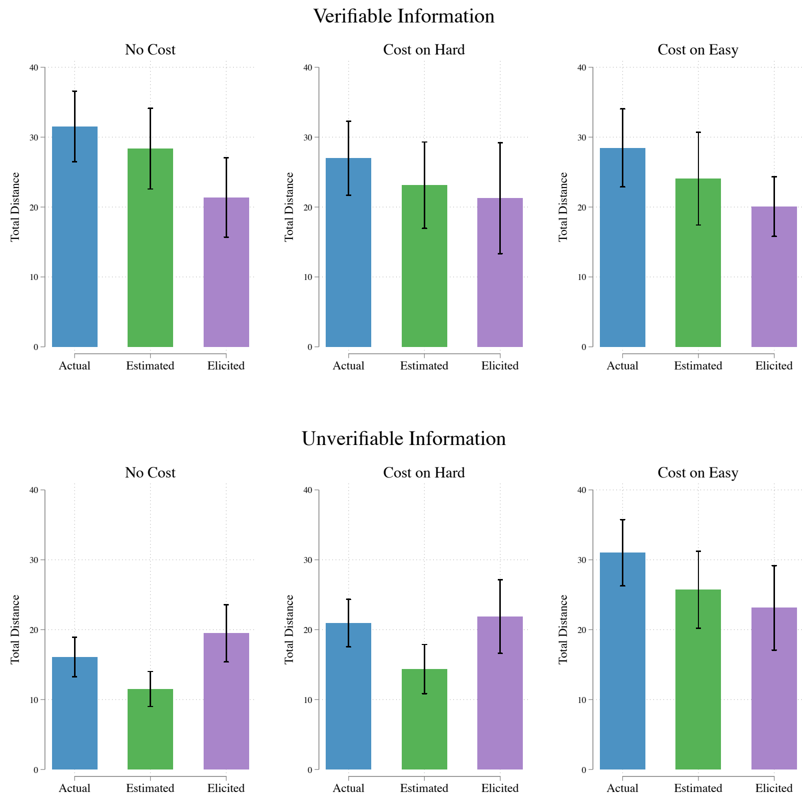


Figure 6: Counterfactuals of the average total distance from the Bayesian benchmark.

which can affect the outcome of strategic interactions. The underlying assumptions of such models differ, providing potentially different explanations for biases in the receivers' beliefs. Some models assume the inability of receivers to account for type-dependent selection, and, as a consequence, for the informational content of the others' action (see, for instance, Eyster and Rabin, 2005; Jehiel, 2005; Jehiel and Koessler, 2008; Esponda, 2008; Cohen and Li, 2022; Fong et al., 2025). While this is a possible explanation for the outcome in some of our treatments, it does not seem to be fully descriptive of the behavior in our data. Indeed, receivers' biases are detected also when the selection is type independent—for instance, the no-cost treatments—adding a different type of deviations.

Another possibility is that subjects best respond to lower levels of sophistication in their opponent, in a level-k fashion—either because they are unable to reason more strategically or because they hold incorrect beliefs about others' sophistication (see, for instance, Stahl II and Wilson, 1994; Nagel, 1995; Camerer et al., 2004; Crawford and Iriberri, 2007; Agranov et al., 2012; Alaoui and Penta, 2016; Alaoui et al., 2020). Under specific assumptions over the level-0 players, we could explain some of our patterns with these types of models—for instance, unsophisticated receivers may believe that senders only react to costs in their choices.

However, to explain the patter we see in the data, we would need to assume that receivers are not only strategically unsophisticated, but also probabilistically. This last assumption would be inconsistent with the accuracy we document in our non-strategic treatments. In addition, our analysis focuses on the second part of our session, after subjects have been exposed to 15 rounds of senders' choices—which reveal a high degree of sophistication—and have received detailed feedback about those in each cost parameterization. Hence, while we cannot reject this type of explanation, it seems reasonable that receivers would at least partially update their beliefs over senders' sophistication.

Finally, subjects could completely neglect the possibility of selection in DGPs and treat the data at *face-value*, as if they were the unbiased outcome of the exogenously given information environment. Related to this point, a growing experimental literature examines how sample selection distorts inference, finding that people struggle to adjust for data selection both in strategic and non-strategic settings (Esponda and Vespa, 2018; Brown and Fragiadakis, 2019; Enke, 2020; Araujo et al., 2021; Degan et al., 2023; Barron et al., 2024; Farina et al., 2024). In these cases, subjects tend to treat the evidence as if it were a random draw from the true DGP—overweighting the apparent representativeness of the sample they see.

This paper shifts the focus from selection of *data* to selection of *data-generating processes*, recreating, through different mechanisms, the feature that the meaning of the evidence depends on the specific selection rule. Ex-ante, it is unclear how biases from the selection-neglect literature translate to this setting, and how much weight the strategic uncertainty has on the receiver's inference. However, our data reveal similar patterns—i.e., the underestimation of the relevance of selection while interpreting the data—only when DGP selection is strategic. Our evidence complements the selection neglect literature by showing that misinterpretation of *selected processes* arises primarily when the selection rule itself is uncertain because the outcome of another agent's strategic choice—unlike classic sample selection, where the central difficulty is forming beliefs about the unobserved data.

To partially disentangle these potentially overlapping explanations, our design included an unincentivized post-experimental survey in which subjects were asked to briefly describe how they approached the game. The answers we recorded are quite heterogeneous, but some widespread points can be of help for our exercise. In particular, subjects classified as equilibrium often discuss the reasoning behind the sender's incentives. Those categorized as reactive almost always referred to the negative effect of costs on the likelihood of a given action, but less frequently mentioned incentives in the no-cost treatments. Finally, while some subjects in the random group noted the possible effect of costs on senders' choices, a substantial share did not mention the sender at all and instead focused on probabilistic inference in their guesses. This pattern is consistent with a tendency to ignore the strategic dimension of the problem and to interpret the data at *face value*, rather than conjecturing limited sophistication on the part of the opponent.

5 Information Transmission

The analysis shows that receivers rely on similar approaches when interpreting evidence across the two information environments. In both cases, many receivers appear to treat the data as if it resulted from random DGP selection, attributing differences primarily to the information environment rather than to the sender's strategic incentives.

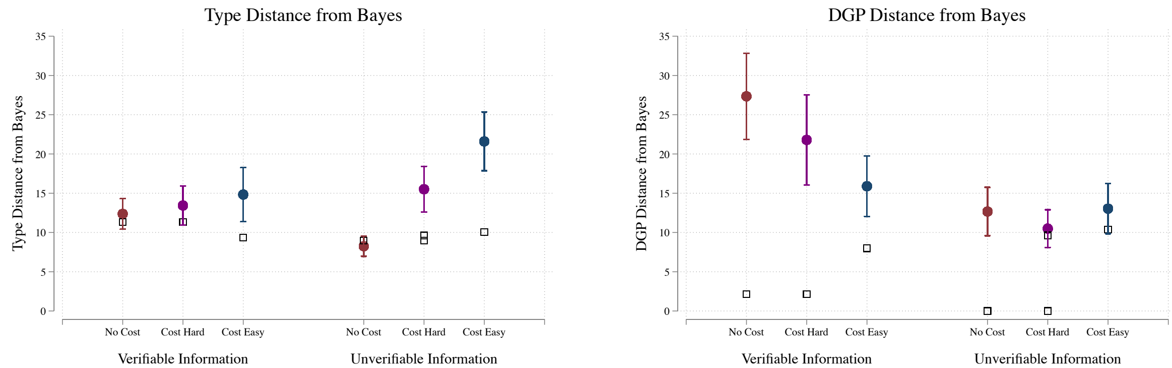
While our unified framework shows that the main drivers of behavior are consistent across levels of information verifiability, the implications for information transmission differ across environments. Indeed, in our treatments, the data tends to be more informative about different aspects of the sender's private information—either their type or the chosen DGP. To assess the implications of the receivers' biases on the information transmitted about these variables, we can introduce two new measures, which capture the inaccuracy of the receivers' inference about either the sender type or the chosen DGP, relative to the Bayesian benchmark. We define:

$$\text{Type Distance} = |\bar{a}(\theta_H \mid \bar{s}) - a^*(\theta_H \mid \bar{s})|$$

$$\text{DGP Distance} = |\bar{a}(A \mid \bar{s}) - a^*(A \mid \bar{s})|$$

where $\bar{a}(\cdot|\bar{s})$ represent the average individual guess about either θ_H or DGP A for a certain realization of $\bar{s}$ and $a^*(\cdot|\bar{s})$ the corresponding Bayesian benchmark. Similarly to before with the total distance, we can interpret these measures as the average percentage points that are misallocated in the subjects' inference about the type or the DGP.

While these measures do not capture the absolute amount of information transmitted, they are informative about the magnitude of the gap between maximal informativeness and the actual information conveyed in the game. Indeed, less accurate inference tends to generate larger gaps in information extraction. As with the total distance from the Bayesian benchmark introduced earlier, these measures share a theoretical benchmark of zero, making them particularly useful for cross-treatment comparisons. Figure 7 displays the average type distance (left panel) and the average DGP distance (right panel) for all treatments. These are then compared with the deterministic treatment that replicates the DGP-selection rule predicted by the theory (shown as hollow squares), which provides us with a benchmark for the probabilistic mistakes.



Notes: The hollow squares report the average type and DGP distance from the Bayesian benchmark in the deterministic treatments where the DGP selection rule replicates the theoretical predictions of the strategic game. For the $U - C_{Hard}$ treatment, we report the deterministic treatments that correspond to the two possible theoretical benchmarks.

Figure 7: Average type and DGP distances from the Bayesian benchmark.

The figure reveals a first clear pattern: in the verifiable information environment, most deviations from the Bayesian benchmark concern DGP inference, whereas in the unverifiable environment, inference about the type tends to be less accurate. This result is a direct implication of the biases documented in Section 4.2, since it is clear how receivers tend to do a better job in their inference when the importance of strategic considerations is limited—as is the case for the type when information is verifiable and for the DGP when information is not verifiable. Concerning the comparison with the deterministic treatments, we find that for the verifiable information case, the DGP distance is always significantly higher in the strategic treatments (p-values< 0.01), while only the type distance in $V - C_{Easy}$ significantly exceeds the one in the deterministic counterpart by a small magnitude (p-value< 0.01). For the unverifiable information treatments, we find that, when DGP selection is costly, the type distance is significantly higher than in the deterministic cases (p-values< 0.01). The inference about the DGP is, instead, equally accurate between the two types of treatments when the DGP selection rule is not degenerate (i.e., the Easy DGP is always chosen). This further reinforces our main result that most biases in inference are due to the uncertainty over the strategic DGP selection, which plays a larger role than standard probabilistic mistakes.

A second key insight from Figure 7 is the presence of clear comparative statics in our distance measures within each information environment.[49] In the verifiable information treatments,

[49] We compare the type or DGP distance in the different cost parametrizations controlling for the realized distribution of data, to account for the possibility that some values of $\bar{s}$ leads to systematically higher or lower accuracy. However, our comparisons are confirmed even in the absence of these controls.

we find that the inference about the DGP becomes more accurate when costs are positive (p-value < 0.05 for $V - C_0$ and $V - C_{Hard}$ and p-value < 0.01 for $V - C_0$ and $V - C_{Easy}$). This improvement is the result of two mechanisms: (i) the introduction of a cost improves the assessment of the DGP for the reactive group; (ii) given the change in sender's choices, the cost on the Easy DGP induces a more extreme distribution of data, which tends to be more informative about the chosen process (see Appedix E.7.1 for more details). Instead, no clear pattern exists in the type distance, where we see a slight increase in inaccuracy after costs are introduced, but such a change is both small in magnitude and not statistically significant.

The opposite reasoning applies to the verifiable information treatments. Indeed, our data reveal that subjects become more incorrect in their inference about the type when costs are introduced (p-values < 0.01), while the level of accuracy tends to be constant in the beliefs formed about the DGP.[50] While this result may seem to contradict the positive effect of cost on strategic inference through the reactive group, it is, instead, consistent with the results derived in Section 4.2. In $U - C_0$, sender types tend to use a (noisy) pooling behavior, making strategic considerations less important. This aspect, paired with the receivers' ability to make correct inferences about the DGP, limits the biases in information extraction. However, once costs are introduced, senders start making type-dependent choices—especially in the $U - C_{Easy}$ treatment, where a fully separating equilibrium is played. This substantially limits the amount of information that can be extracted from the data by the receivers who neglect these strategic choices, which are essential to interpret otherwise almost uninformative data (see Appendix E.7.1 for more details).

Finally, the last important comparison involves the two treatments with a cost on the Easy DGP. Given our parametrization and the senders' choices in the data, these two treatments lead to extremely similar outcomes across information environments. However, the existing biases in the receivers' behavior affect information extraction in different ways. As we could expect from the previous discussion, the inference about the type is more inaccurate when information is unverifiable ($p-$value < 0.01). Instead, while the inference about the DGP tends to be more inaccurate with unverifiable information, such a gap is not statistically significant.

Information Transmission. While type and DGP distance are useful for highlighting the main forces behind incomplete information extraction and for enabling direct comparisons across treatments, they do not provide insight into the overall level of informativeness in the

[50]We detect a significant decrease in the DGP distance for the $U - C_{Hard}$ treatment compared to the other cost parametrizations (p-values < 0.05), but the magnitude is limited.

Table 6: Information transmitted and processed (measured as expected receiver's payoff).

	Verifiable Information		**Unverifiable Information**	
	Info on Type	Info on DGP	Info on Type	Info on DGP
C_0	82.45	82.48	74.48	90.39
	[85.21]	[94.38]	[75.72]	[92.43]
C_{Hard}	77.65	88.61	70.85	84.14
	[83.04]	[96.76]	[75.40]	[85.98]
C_{Easy}	85.88	84.21	79.91	89.86
	[89.45]	[89.75]	[88.88]	[92.46]

Notes: The information transmitted, measured as the expected payoff of a Bayesian receiver, is reported in the square brackets. In the absence of information other than the prior, the benchmark payoff for the inference about the type is 75.

different environments. For each treatment, Table 6 reports the expected receiver payoff for the type and DGP inference (as in Section 2.1.1) and compares it with the expected payoff of a receiver following the Bayesian benchmark (reported in square brackets).[51]

First, we can assess changes in the amount of information transmitted. Given the senders' choices, the most informative outcome about the type is achieved when there is cost on the Easy DGP (respectively, p-value < 0.1 and p-value < 0.05 for the comparison with $V - C_0$ and in $V - C_{Hard}$ and p-value < 0.01 for the unverifiable information treatments). Instead, no difference is detected between the no-cost environment and the one with a cost on the Hard DGP. While this is consistent with our predictions in the verifiable information case, in the unverifiable information environment, this result is the outcome of an incomplete exploitation of the signaling opportunity for the high type.

Second, a large amount of information is transmitted about the DGP. In the verifiable information environment, the almost perfect pooling equilibrium played by the senders in $V - C_0$ and in $V - C_{Hard}$ makes the outcome extremely informative about the DGP, and leads to more information transmission than in $V - C_{Easy}$ (respectively, p-value < 0.1 and p-value < 0.01). Instead, in the unverifiable information treatments, the senders' choices lead to a similar amount of information transmission in $U - C_0$ and $U - C_{Easy}$, but strictly less in the

[51] Appendix E.7 reports additional tables summarizing the information predictions of our theoretical model and reporting the measure of informativeness derived by the joint inference over type and DGP.

$U - C_{Hard}$ (p-value < 0.01 for $U - C_0$ and p-value < 0.05 for $U - C_{Easy}$).

Finally, consistently with the theoretical predictions, we do not detect any difference in informativeness between $V - C_{Easy}$ and $U - C_{Easy}$.

Given the amount of information transmitted by the senders, we can now study the amount of information processed by the receivers. First, Table 6 reveals that the information extracted is generally below the information transmitted. For the verifiable information treatments, such a gap is always significant (p-value < 0.01), even if the magnitudes tend to be larger for the information about the DGP. Instead, for the unverifiable information treatments, we detect a significant gap in the information processed about the type when selection costs are positive (for $U - C_{Hard}$ p-value < 0.05 and for $U - C_{Easy}$ p-value < 0.01), while for the DGP, we only detect a significant gap in $U - C_{Easy}$ (p-value < 0.05), whose magnitude is contained. These results are consistent with the patterns discussed in Figure 7. It is important to highlight that the resulting expected payoff of the receiver in $U - C_{Hard}$ is strictly below the no information benchmark (p-value < 0.05): this is likely the consequence of the large uncertainty over the senders' behavior, which can lead to extremely different interpretations of the data.[52]

Despite these gaps, we generally find that receivers extract more information about the type when more is transmitted. In the verifiable information treatments, we find that more information about the type is processed in $V - C_{Easy}$ than in $V - C_{Hard}$ (p-value < 0.05), while the positive change between $V - C_0$ and $V - C_{Easy}$ is not statistically significant.[53] The same result holds in the unverifiable information treatments, where most information about the type is extracted in $U - C_{Easy}$ (p-value < 0.05 for the comparisons with $U - C_0$ and p-value < 0.01 for the comparison with $U - C_{Hard}$). While there is a decrease in type informativeness between $U - C_0$ and $U - C_{Hard}$, such a change is not significant in our data (p-value $= 0.12$).

Moving to the information extracted about the DGP, in the verifiable information cases, we find that the lowest amount of information is extracted in $V - C_0$, even if the only significant difference is detected with $V - C_{Hard}$ (p-value < 0.1). In the unverifiable information case, instead, the least informative outcome is induced in $U - C_{Hard}$ (p-values < 0.05)—and this is again consistent with the increased uncertainty over the interpretation of the data.

Finally, we can compare the information processed in the $V - C_{Easy}$ and $U - C_{Easy}$ treat-

[52]See Table 13 in Appendix E.7.1 for more details about how the behavior of the different receiver clusters affects information extraction.

[53]The introduction of a cost on the Hard DGP leads to a decrease compared to the $V - C_0$ treatment (p-value < 0.1). This decrease is mainly the result of more inaccurate guesses of the random DGP group, due to more extreme choices of the sender. For more details, see Table 4 and Table 13.

ments, which are theoretically equivalent. We find that the information processed about the type is higher in $V - C_{Easy}$ (p-value < 0.1), while the information processed about the DGP is higher in $U - C_{Easy}$ (p-value < 0.05). It is worth reiterating that senders' choices are extremely similar in these two treatments, inducing the same information extraction opportunities. Any difference is the outcome of how receivers react to the nature of the feasible DGPs, and so the characteristics of the exogenously given information environment.

Implications of the Results. Overall, this section's analysis reveals two main insights. First, receivers' inference tends to be more accurate for the variable that the data are intrinsically more informative about. Consequently, the effect of introducing exogenous costs varies with the environment and the equilibrium predictions, which determine the relative importance of strategic considerations in the different settings. Second, despite these biases, receivers can still extract more information when more is transmitted. However, particularly in unverifiable environments, achieving high levels of informativeness requires type-specific selection, and failing to account for DGP choices attenuates such gains.

These insights lead to two broader implications relevant for applications and policy design. First, restricting the senders' ability to select among DGPs may be suboptimal, especially when these choices help shape more informative data distributions. When feasible, introducing frictions that promote specific types of selection may enhance information transmission.[54] Second, verifiable information environments generally facilitate more effective transmission about the senders' private type, regardless of the equilibrium outcome.

Information Transmission - Result 1: *Neglecting DGP selection affects information transmission differently across environments and cost parameterizations. Nonetheless, subjects tend to extract more information when more is transmitted by the sender.*

5.1 Transparency in DGP Choice

Recent regulatory efforts in science and accounting emphasize transparency in the processes that generate the data underlying inference. In research, pre-registration policies make analysis and data-collection plans observable in advance, limiting ex-post flexibility in how data are

[54]Previous studies have documented the positive effects of data selection (e.g., Di Tillio et al., 2021; Farina et al., 2024). Failing to account for selected samples weakens these benefits for inference, which, in our setting, may instead be preserved through the generation of more extreme and informative data distributions.

produced or results derived.[55] In accounting, the *Securities and Exchange Commission (SEC)*'s guidance on key performance indicators disclosures similarly requires firms to explain how metrics are computed and to reconcile them with standard measures.[56] Both initiatives aim to make the data-generation process itself observable and to enhance the transparency of the source of information. This mechanism corresponds to our strategic treatment with observable DGP choices. Importantly, in this modified setting, the sender remains informed about her type at the time of this choice, preserving the potential strategic effect of the decision on the meaning of the data. If receivers neglect these choices, information extraction may still be impaired, despite the increase in transparency.

We can then ask the question of how information transmission about the type is affected by the observability of the DGP. From a theoretical perspective, the answer to this question is ambiguous.[57] Indeed, the observability of the DGP increases the number of equilibria satisfying our monotonicity refinement and, when costs are imposed, senders' choices are predicted to be less extreme and, as a consequence, potentially less informative.[58] Appendix F summarizes our main results about senders' and receivers' choices. Overall, we find that the observability of the action makes the senders' choices of DGP less extreme, reducing the extent of selection. In the treatments with a positive DGP selection cost, this also implies that low-type senders pay the selection cost more often. While full separation cannot be an equilibrium outcome in this setting, in our data, given the receivers' behavior, it is optimal for the low type to not pay the cost—and for the high type to choose the alternative action—except for the $U - C_{Easy}$ treatment. The tendency to choose the costly DGP too often leads to inefficiencies in terms of information transmission—preventing informative separation—and suboptimally reduces the effective payoff of low type senders.[59] Concerning receivers, we find that they tend to respond to the different environments and strategic incentives, even if imperfectly.

Given these general patterns, we can assess whether the effectiveness of communication about the type is improved by the increase in DGP transparency. First, we can evaluate the accuracy of the inference by replicating our type distance measure in these auxiliary treatments.

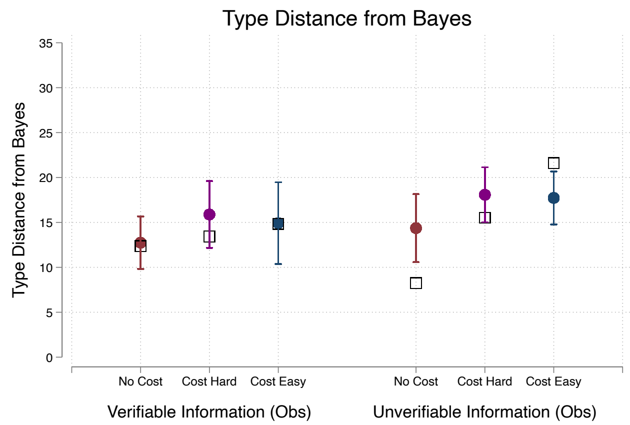
[55] See, for instance, Nosek et al. (2018), Adda et al. (2020), and Decker and Ottaviani (2023) for discussions about the effects of scientific pre-registration.

[56] See *SEC* (2020), Commission Guidance on Management's Discussion and Analysis of Financial Condition and Results of Operations, Release Nos. 33-10751; 34-88228.

[57] See Perez-Richet (2014) for an intuition about the theoretical argument.

[58] Downing et al. (2024) and Cohen and Nguyen (2024) show that firms strategically choose benchmarks in response to scrutiny or strategically adjust their targets across periods.

[59] We detect smaller changes in the high type payoff, with the only significant increase in $U - C_B$.



Notes: The hollow squares report the average type distance from the Bayesian benchmark for the corresponding strategic treatment with unobservable DGP.

Figure 8: Average type distances from Bayesian benchmark for the observable DGP treatments.

The main difference is that now the Bayesian benchmark also varies across DGP choices. Figure 8 reports the average type distance for our observable DGP treatments and compares it with our corresponding measure in the main strategic treatments (the hollow squares). The figure gives us two main insights. First, we do not detect a substantial improvement in the overall accuracy of receivers' guesses when the chosen DGP is observable. Second, differently from before, we find no meaningful comparative statics, suggesting that receivers face common challenges across cost parametrizations.

Finally, we can compare the informativeness of communication in the two strategic games. Table 7 reports the amount of information about the type that is transmitted and processed in our set of treatments. First, we can compare the information transmitted by the sender in the two observability regimes. In the verifiable information treatments, we find no significant changes, even when costs are introduced: the less extreme sender's behavior counterbalances the positive impact of the DGP observability, leading to comparable outcomes. Instead, in the unverifiable information environment, the effect depends on the specific cost parametrization. When the DGP becomes observable, more information is transmitted in $U - C_0$ (p-value < 0.01) and in $U - C_{Hard}$ (p-value < 0.01), while less in $U - C_{Easy}$ (p-value < 0.01). This is the result of three main mechanisms. First, in the absence of costs, senders choose the Hard DGP more often, and its induced distribution is, by design, more informative about the type. Second, in the $U - C_{Hard}$ treatment, the signaling opportunity of the cost is exploited more effectively, reducing uncertainty over the interpretation of the data. Third, type θ_L pays the DGP selection cost more often in $U - C_{Easy}$, reducing the informative value of the DGP selection.

Table 7: Information transmitted and processed about the type (measured as the expected receiver's payoff).

	Verifiable Information		**Unverifiable Information**	
	Unobservable DGP	Observable DGP	Unobservable DGP	Observable DGP
C_0	82.45	79.59	74.48	76.86
	[85.21]	[83.63]	[75.72]	[81.23]
C_{Hard}	77.65	82.36	70.85	78.39
	[83.04]	[85.75]	[75.40]	[84.43]
C_{Easy}	85.88	84.08	79.91	78.66
	[89.45]	[88.05]	[88.88]	[83.89]

Notes: The information transmitted, measured as the expected payoff of a Bayesian receiver, is reported in the square brackets.

Finally, we compare the levels of information processed. In the verifiable information treatments, we find no significant change in the amount of information extracted by receivers.[60] The same holds for the $U - C_0$ and $U - C_{Easy}$ treatments. In contrast, for $U - C_{Hard}$ we detect a substantial increase in the amount of information processed ($p-$value < 0.05).[61]

Overall, these results suggest that making the DGP observable does not necessarily enhance information transmission when it remains the outcome of an informed sender's strategic choice. While the neglect of strategic selection may differ across settings, greater transparency over the data-generating process alone is unlikely to offer a universally effective regulatory remedy to ex-ante data manipulation.

Information Transmission - Result 2: *Increasing transparency over the choice of the DGP does not generally improve information transmission when the selection remains strategically determined by an informed sender.*

[60] Even in $V - C_{Hard}$ where the gap is the largest, the $p-$value is only 0.143.

[61] We confirm these results in the full sample (see Appendix G.3).

6 Conclusion

This paper develops an experimental framework to study information transmission when data are generated through an unobservable and strategically chosen process. The design varies both the sender's feasible data-generating processes and the costs of their selection, thereby changing the degree of information verifiability and the incentives for strategic choices. These variations make it possible to isolate the mechanisms behind DGP selection and inference and to assess how the relative importance of strategic and probabilistic considerations affects communication outcomes.

Across treatments, several patterns emerge. Senders' choices broadly align with the theoretical predictions. They respond systematically to both exogenous incentives, induced by selection costs, and endogenous incentives, arising from receivers' inference. This reveals that strategic DGP selection is descriptive of senders' behavior and must be accounted for by receivers. Receivers, instead, only partially account for the strategic selection: many interpret data as if they were produced by randomly selected DGPs. Introducing asymmetric selection costs improves strategic reasoning but does not eliminate these biases.

The implications for information transmission differ across environments. When information is verifiable, inference is mostly inaccurate about the DGP, while inference about the sender's type does not exhibit systematic biases. When information is unverifiable, inference about the type deteriorates, especially when costs induce separation. Information transmission is thus most impaired when the realized data are highly sensitive to the sender's DGP choice.

Finally, we examine whether greater transparency over DGP choices improves communication. Making senders' choices observable does not generally enhance information transmission. Although transparency removes one dimension of uncertainty, it also alters senders' incentives and can reduce the informativeness of the data. Overall, the results suggest that policies targeting the production stage of information may be ineffective if they fail to account for the context-specific strategic nature of data generation itself.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for truth-telling," *Econometrica*, 87, 1115–1153.

Adda, J., C. Decker, and M. Ottaviani (2020): "P-hacking in clinical trials and how incentives shape the distribution of results across phases," *Proceedings of the National Academy of Sciences*, 117, 13386–13392.

Agranov, M., E. Potamites, A. Schotter, and C. Tergiman (2012): "Beliefs and endogenous cognitive levels: An experimental study," *Games and Economic Behavior*, 75, 449–463.

Aina, C. and F. Schneider (2025): "Weighting competing models," .

Alaoui, L., K. A. Janezic, and A. Penta (2020): "Reasoning about others' reasoning," *Journal of Economic Theory*, 189, 105091.

Alaoui, L. and A. Penta (2016): "Endogenous depth of reasoning," *The Review of Economic Studies*, 83, 1297–1333.

Ali, S. N., M. Mihm, L. Siga, and C. Tergiman (2021): "Adverse and advantageous selection in the laboratory," *American Economic Review*, 111, 2152–2178.

Allen, F. (1987): "Discovering personal probabilities when utility functions are unknown," *Management Science*, 33, 542–544.

Ambuehl, S. and H. C. Thysen (2024): "Choosing between causal interpretations: An experimental study," *NHH Dept. of Economics Discussion Paper*.

Araujo, F. A., S. W. Wang, and A. J. Wilson (2021): "The times they are A-changing: Experimenting with dynamic adverse selection," *American Economic Journal: Microeconomics*, 13, 1–22.

Ba, C., J. A. Bohren, and A. Imas (2024): "Over-and underreaction to information," *Available at SSRN 4274617*.

Barron, K., S. Huck, and P. Jehiel (2024): "Everyday econometricians: Selection neglect and overoptimism when learning from others," *American Economic Journal: Microeconomics*, 16, 162–198.

Benjamin, D. J. (2019): "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2, 69–186.

Benndorf, V., D. Kübler, and H.-T. Normann (2015): "Privacy concerns, voluntary disclosure of information, and unraveling: An experiment," *European Economic Review*, 75, 43–59.

Bergemann, D. and S. Morris (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

Bertomeu, J., P. Ma, and I. Marinovic (2020): "How Often Do Managers Withhold Information?" *The Accounting Review*, 95, 73–102.

Blume, A., D. V. DeJong, Y.-G. Kim, and G. B. Sprinkle (1998): "Experimental evidence on the evolution of meaning of messages in sender-receiver games," *The American Economic Review*, 88, 1323–1340.

Blume, A., E. K. Lai, and W. Lim (2020): "Strategic information transmission: A survey of experiments and theoretical foundations," *Handbook of experimental game theory*, 311–347.

Brandts, J. and C. A. Holt (1992): "An experimental test of equilibrium dominance in signaling games," *The American Economic Review*, 82, 1350–1365.

——— (1993): "Adjustment patterns and equilibrium selection in experimental signaling games," *International Journal of Game Theory*, 22, 279–302.

Brown, A. L., C. F. Camerer, and D. Lovallo (2012): "To review or not to review? Limited strategic thinking at the movie box office," *American Economic Journal: Microeconomics*, 4, 1–26.

Brown, A. L. and D. E. Fragiadakis (2019): "Benign vs. Self-Serving Information Reduction: Do Individuals Understand the Difference?" *Self-Serving Information Reduction: Do Individuals Understand the Difference*.

Burdea, V., M. Montero, and M. Sefton (2023): "Communication with partially verifiable information: an experiment," *Games and Economic Behavior*, 142, 113–149.

Cai, H. and J. T.-Y. Wang (2006): "Overcommunication in strategic information transmission games," *Games and Economic Behavior*, 56, 7–36.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A cognitive hierarchy model of games," *The quarterly journal of economics*, 119, 861–898.

Carlsson, H. and S. Dasgupta (1997): "Noise-proof equilibria in two-action signaling games," *Journal of Economic Theory*, 77, 432–460.

Charles, C. and C. Kendall (2024): "Causal narratives," *Available at SSRN 4669371*.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Cohen, L. and Q. Nguyen (2024): "Moving Targets," *Available at SSRN 4736129*.

Cohen, S. and S. Li (2022): "Sequential cursed equilibrium," *arXiv preprint arXiv:2212.06025*.

Crawford, V. P. and N. Iriberri (2007): "Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?" *Econometrica*, 75, 1721–1770.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1431–1451.

Daley, B. and B. Green (2014): "Market signaling with grades," *Journal of Economic Theory*, 151, 114–145.

Danz, D., L. Vesterlund, and A. J. Wilson (2022): "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 112, 2851–2883.

de Clippel, G. and K. Rozen (2024): "Communication, perception, and strategic obfuscation," Tech. rep., Working Paper.

de Haan, T., T. Offerman, and R. Sloof (2011): "Noisy signaling: theory and experiment," *Games and Economic Behavior*, 73, 402–428.

Decker, C. and M. Ottaviani (2023): "Preregistration and Credibility of Clinical Trials," *medRxiv*, 2023–05.

Degan, A., M. Li, and H. Xie (2023): "An experimental investigation of persuasion through selective disclosure of evidence," *Canadian Journal of Economics/Revue canadienne d'économique*, 56, 1490–1516.

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2021): "Strategic sample selection," *Econometrica*, 89, 911–953.

Dickhaut, J., M. Ledyard, A. Mukherji, and H. Sapra (2003): "Information management and valuation: an experimental investigation," *Games and Economic Behavior*, 44, 26–53.

Dickhaut, J. W., K. A. McCabe, and A. Mukherji (1995): "An experimental study of strategic information transmission," *Economic Theory*, 6, 389–403.

Downing, C., B. Levy, M. Phillips, and E. C. So (2024): "Odd Lots & Optics: Manipulation in Response to Scrutiny," .

Dye, R. A. (1985): "Disclosure of nonproprietary information," *Journal of accounting research*, 123–145.

Enke, B. (2020): "What you see is all there is," *The Quarterly Journal of Economics*, 135, 1363–1398.

——— (2024): "The cognitive turn in behavioral economics," *Unpublished manuscript. https://benjamin-enke. com*.

Esponda, I. (2008): "Behavioral equilibrium in economies with adverse selection," *American Economic Review*, 98, 1269–1291.

Esponda, I. and E. Vespa (2018): "Endogenous sample selection: A laboratory study," *Quantitative Economics*, 9, 183–216.

Eyster, E. and M. Rabin (2005): "Cursed equilibrium," *Econometrica*, 73, 1623–1672.

Farina, A., G. Fréchette, A. Ispano, A. Lizzeri, and J. Perego (2024): "The Selective Disclosure of Evidence: An Experiment," Tech. rep., National Bureau of Economic Research.

Farina, A. and M. Leccese (2024): "Hiding a Flaw: A Lab Experiment on Multi-Dimensional Information Disclosure," *Available at SSRN 4785923*.

Fields, T. D., T. Z. Lys, and L. Vincent (2001): "Empirical research on accounting choice," *Journal of accounting and economics*, 31, 255–307.

Fong, M.-J., P.-H. Lin, and T. R. Palfrey (2025): "Cursed sequential equilibrium," *American Economic Review*, 115, 2616–2658.

Forsythe, R., R. M. Isaac, and T. R. Palfrey (1989): "Theories and tests of" blind bidding" in sealed-bid auctions," *The Rand Journal of Economics*, 214–238.

Forsythe, R., R. Lundholm, and T. Rietz (1999): "Cheap talk, fraud, and adverse selection in financial markets: Some experimental evidence," *The Review of Financial Studies*, 12, 481–518.

FRÉCHETTE, G. R., A. LIZZERI, AND J. PEREGO (2022): "Rules and commitment in communication: An experimental analysis," *Econometrica*, 90, 2283–2318.

FRÉCHETTE, G. R., E. VESPA, AND S. YUKSEL (2024): "Extracting Models From Data Sets: An Experiment," *Available at SSRN 5026751*.

GENG, S. AND M. GUAN (2025): "Disentangling Probabilistic and Strategic Errors," *Available at SSRN 5224282*.

GNEEZY, U. AND J. POTTERS (1997): "An experiment on risk taking and evaluation periods," *The quarterly journal of economics*, 112, 631–645.

GREEN, J. R. AND N. L. STOKEY (2007): "A two-person game of information transmission," *Journal of economic theory*, 135, 90–104.

GREINER, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125.

GRETHER, D. M. (1980): "Bayes rule as a descriptive model: The representativeness heuristic," *The Quarterly journal of economics*, 95, 537–557.

GROSSMAN, S. J. (1981): "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 24, 461–483.

GÜTH, W. AND F. WINTER (2013): "Sorting via screening versus signaling: a theoretic and experimental comparison," Tech. rep., Jena Economic Research Papers.

HAGENBACH, J. AND E. PEREZ-RICHET (2018): "Communication with Evidence in the Lab," *Games and Economic Behavior*, 112, 139–165.

HAGENBACH, J. AND C. SAUCET (2024): "Motivated Skepticism," *Review of Economic Studies*, rdae058.

HOSSAIN, T. AND R. OKUI (2013): "The binarized scoring rule," *Review of Economic Studies*, 80, 984–1001.

IOANNIDIS, J. P. (2005): "Why most published research findings are false," *PLoS medicine*, 2, e124.

JEHIEL, P. (2005): "Analogy-based expectation equilibrium," *Journal of Economic theory*, 123, 81–104.

Jehiel, P. and F. Koessler (2008): "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 62, 533–557.

Jeitschko, T. D. and H.-T. Normann (2012): "Signaling in deterministic and stochastic settings," *Journal of Economic Behavior & Organization*, 82, 39–55.

Jin, G. Z. and P. Leslie (2003): "The effect of information on product quality: Evidence from restaurant hygiene grade cards," *The Quarterly Journal of Economics*, 118, 409–451.

Jin, G. Z., M. Luca, and D. Martin (2021): "Is no news (perceived as) bad news? An experimental investigation of information disclosure," *American Economic Journal: Microeconomics*, 13, 141–173.

——— (2022): "Complex disclosure," *Management Science*, 68, 3236–3261.

Jovanovic, B. (1982): "Truthful disclosure of information," *The Bell Journal of Economics*, 36–44.

Jung, W.-O. and Y. K. Kwon (1988): "Disclosure when the market is unsure of information endowment of managers," *Journal of Accounting research*, 146–153.

Kagel, J. H. and D. Levin (2016): "Auctions: A survey of experimental research," *The handbook of experimental economics*, 2, 563–637.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

Kendall, C. and R. Oprea (2024): "On the complexity of forming mental models," *Quantitative Economics*, 15, 175–211.

Kihlstrom, R. E. and M. H. Riordan (1984): "Advertising as a Signal," *Journal of political economy*, 92, 427–450.

King, R. R. and D. E. Wallin (1991): "Market-induced information disclosures: An experimental markets investigation," *Contemporary Accounting Research*, 8, 170–197.

Koessler, F. and V. Skreta (2023): "Informed information design," *Journal of Political Economy*, 131, 3186–3232.

Kübler, D., W. Müller, and H.-T. Normann (2008): "Job-market signaling and screening: An experimental comparison," *Games and Economic Behavior*, 64, 219–236.

Leland, H. E. and D. H. Pyle (1977): "Informational asymmetries, financial structure, and financial intermediation," *The journal of Finance*, 32, 371–387.

Liang, Y. (2025): "Learning from unknown information sources," *Management Science*, 71, 3873–3890.

Mathios, A. D. (2000): "The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market," *The Journal of Law and Economics*, 43, 651–678.

Matthews, S. A. and L. J. Mirman (1983): "Equilibrium limit pricing: The effects of private information and stochastic demand," *Econometrica: Journal of the Econometric Society*, 981–996.

McKelvey, R. D. and T. Page (1990): "Public and private information: An experimental study of information pooling," *Econometrica: Journal of the Econometric Society*, 1321–1339.

Milgrom, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391.

Musolff, R. and F. Zimmermann (2025): "Model uncertainty," *Working Paper*.

Nagel, R. (1995): "Unraveling in guessing games: An experimental study," *The American economic review*, 85, 1313–1326.

Niederle, M. and E. Vespa (2023): "Cognitive limitations: Failures of contingent thinking," *Annual Review of Economics*, 15, 307–328.

Nosek, B. A., C. R. Ebersole, A. C. DeHaven, and D. T. Mellor (2018): "The preregistration revolution," *Proceedings of the National Academy of Sciences*, 115, 2600–2606.

Ognjanovic, I., D. Gasevic, and S. Dawson (2016): "Using institutional data to predict student course selections in higher education," *The Internet and Higher Education*, 29, 49–62.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990): "Strategic information revelation," *The Review of Economic Studies*, 57, 25–47.

Ortoleva, P. (2012): "Modeling the change of paradigm: Non-Bayesian reactions to unexpected news," *American Economic Review*, 102, 2410–2436.

PENCZYNSKI, S., C. KOCH, AND S. ZHANG (2023): "Disclosure of verifiable information under competition: An experimental study," *SSRN Working Paper*, 4130449.

PEREZ-RICHET, E. (2014): "Interim bayesian persuasion: First steps," *American Economic Review*, 104, 469–474.

RENTSCHLER, L. AND Z. SAMAD (2025): "Bayesian persuasion with fact-checking: An experimental investigation," *European Economic Review*, 105114.

RILEY, J. G. (2001): "Silver signals: Twenty-five years of screening and signaling," .

SÁNCHEZ-PAGÉS, S. AND M. VORSATZ (2007): "An experimental study of truth-telling in a sender–receiver game," *Games and Economic Behavior*, 61, 86–112.

SCHLAG, K. H. AND J. J. VAN DER WEELE (2013): "Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality," *Theoretical Economics Letters*, 3, 38–42.

SOBEL, J. (2020): "Lying and deception in games," *Journal of Political Economy*, 128, 907–947.

SPENCE, M. (1978): "Job market signaling," in *Uncertainty in economics*, Elsevier, 281–306.

STAHL II, D. O. AND P. W. WILSON (1994): "Experimental evidence on players' models of other players," *Journal of economic behavior & organization*, 25, 309–327.

TING, D. H. AND C. K. C. LEE (2012): "Understanding students' choice of electives and its implications," *Studies in Higher Education*, 37, 309–325.

VERRECCHIA, R. E. (1983): "Discretionary disclosure," *Journal of accounting and economics*, 5, 179–194.

WANG, J. T.-Y., M. SPEZIO, AND C. F. CAMERER (2010): "Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games," *American economic review*, 100, 984–1007.

WILSON, A. J. AND E. VESPA (2020): "Information transmission under the shadow of the future: An experiment," *American Economic Journal: Microeconomics*, 12, 75–98.

WOODFORD, M. (2020): "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics*, 12, 579–601.

ZHOU, J. (2023): "Costly verification and commitment in persuasion," *Journal of Economic Behavior & Organization*, 212, 1100–1142.

Appendix

A Equilibrium Concept and Refinement

This section formally defines what constitutes an equilibrium in our setting and introduces our refinement. First, let us provide a formal definition of PBE for our framework. To this end, note that the receiver's equilibrium strategy is pinned down by her belief on the type θ and on the DGP i after observing any realization of $\bar{s}$. We denote such belief with $\tilde{\mu}(\cdot,\cdot|\bar{s}) \in \Delta(\Theta \times \{A, B\})$. Indeed, given $\bar{s} \leq N$, the receiver maximizes her expected payoff by solving

$$\max_{a\in\mathcal{A}} \quad 1 - \frac{1}{2} \sum_{\substack{\hat{i}\in\{A,B\} \\ \hat{\theta}\in\{\theta_H,\theta_L\}}} \tilde{\mu}(\hat{\theta},\hat{i}|\bar{s}) \sum_{\substack{i\in\{A,B\} \\ \theta\in\{\theta_H,\theta_L\}}} \left(\mathbb{1}\{i=\hat{i},\theta=\hat{\theta}\} - a(\theta,i|\bar{s})\right)^2$$

subject to the constraints that $\sum a(i,\theta,\bar{s}) = 1$ for each $\bar{s} \leq N$. It is easy to verify that for each realization of $\bar{s}$, the receiver optimally chooses a such that $a(\theta,i|\bar{s}) = \tilde{\mu}(\theta,i|\bar{s})$.

The sender, instead, maximizes his payoff by choosing a strategy $q : \{\theta_H, \theta_L\} \to [0,1]$ that solves

$$\max_{q} \quad q(\theta)\mathbb{E}_{\bar{s}}\left[u_S(a,\theta,A|\bar{s})\right] + (1-q(\theta))\,\mathbb{E}_{\bar{s}}\left[u_S(a,\theta,B|\bar{s})\right]$$

where $\mathbb{E}_{\bar{s}}\left[u_S(a,\theta,i|\bar{s})\right] = \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta,i)\left[a(\theta_H,A|\bar{s}) + a(\theta_H,B|\bar{s})\right] - c(\theta,i)$. In what follows, we can simplify the notation by defining $a(\bar{s}) = a(\theta_H,A|\bar{s}) + a(\theta_H,B|\bar{s})$ to be the total probability that the receiver assigns to the type being θ_H.

To simplify the equilibrium definition, we can further define $\mu(\theta_H|\bar{s}) = \tilde{\mu}(\theta_H,A|\bar{s}) + \tilde{\mu}(\theta_H,B|\bar{s})$, which represents the receiver's belief that the type is θ_H given $\bar{s}$.

Definition A.1. (PBE) *Fix $p_A(\theta_H), p_A(\theta_L), p_B(\theta_H), p_B(\theta_L)$ and μ_0. The triple $(q^*, \tilde{\mu}^*, \mu^*)$ is a Perfect Bayesian Equilibrium if*

1. *Given $\mu^*(\cdot|\bar{s}) \in \Delta(\Theta)$, $q^*(\theta) \in [0,1]$ maximizes*

$$q^*(\theta)\left(\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta,A)\mu^*(\theta_H|\bar{s}) - c(\theta,A)\right) + (1-q^*(\theta))\left(\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta,B)\mu^*(\theta_H|\bar{s}) - c(\theta,B)\right)$$

   for each $\theta \in \{\theta_H, \theta_L\}$.

2. *Given $q^* \in [0,1]^{\Theta}$, $\tilde{\mu}^*(\cdot,\cdot|\bar{s}) \in \Delta(\Theta \times \{A,B\})$ is defined as*

$$\tilde{\mu}^*(\theta,i|\bar{s}) = Prob(\theta,i|\bar{s},q^*) = Prob(i|\bar{s},q^*)Prob(\theta|i,\bar{s},q^*),$$

where

$$Prob(A|\bar{s},q^*) = \frac{\mu_0 q^*(\theta_H)Prob(\bar{s}|\theta_H,A) + (1-\mu_0)q^*(\theta_L)Prob(\bar{s}|\theta_L,A)}{Prob(\bar{s}|q^*)},$$

$$Prob(\bar{s}|q^*) = \mu_0\left(q^*(\theta_H)Prob(\bar{s}|\theta_H,A) + (1-q^*(\theta_H))\,Prob(\bar{s}|\theta_H,B)\right)$$
$$+(1-\mu_0)\left(q^*(\theta_L)Prob(\bar{s}|\theta_L,A) + (1-q^*(\theta_L))\,Prob(\bar{s}|\theta_L,B)\right),$$

$$Prob(\theta_H|A,\bar{s},q^*) = \frac{\mu_0 q^*(\theta_H)Prob(\bar{s}|\theta_H,A)}{\mu_0 q^*(\theta_H)Prob(\bar{s}|\theta_H,A) + (1-\mu_0)q^*(\theta_L)Prob(\bar{s}|\theta_L,A)}$$

and

$$Prob(\theta_H|B,\bar{s},q^*) = \frac{\mu_0\,(1-q^*(\theta_H))\,Prob(\bar{s}|\theta_H,B)}{\mu_0\,(1-q^*(\theta_H))\,Prob(\bar{s}|\theta_H,B) + (1-\mu_0)\,(1-q^*(\theta_L))\,Prob(\bar{s}|\theta_L,B)}.$$

3. *Given* $\tilde{\mu}^*(\cdot,\cdot|\bar{s}) \in \Delta(\Theta \times \{A,B\})$, $\mu^*(\cdot|\bar{s}) \in \Delta(\Theta)$ *is defined as*

$$\mu^*(\theta_H|\bar{s}) = \tilde{\mu}^*(\theta_H,A|\bar{s}) + \tilde{\mu}^*(\theta_H,B|\bar{s}).$$

It is important to highlight that, even in a setting in which the receiver is not explicitly required to make any inference about the DGP i, but only about the state θ, the receiver would still need to make this first step to compute the final posterior over θ_H.

As in many communication and signaling games, this game admits multiple PBEs. We aim to refine the set of equilibria on which we focus attention, and we do so by assuming that, in equilibrium, the receiver holds beliefs about θ_H which are strictly monotone in $\bar{s}$.

Definition A.2. (Monotone PBE) *Fix any PBE* $(q^*,\tilde{\mu}^*,\mu^*)$. *The PBE is monotone if* $\mu^*(\theta_H|\bar{s})$ *is strictly monotone in* $\bar{s}$.

With this refinement, we aim to restrict attention to the class of PBE in which the receiver treats more successes as either good or bad news. When making this restriction, we implicitly make the following assumptions:

1. The players believe that some information is contained in the signals, and so that beliefs will change compared to the prior after new information is revealed;

2. Successes are either good or bad news, and they cannot follow more complex patterns.

B Proofs

Before proving the main results, we state some general results that will be useful in the equilibrium derivation later on. We start with the proof of a simple lemma.

Lemma B.1. *Fix any binomial distribution with success parameter $p \in (0,1)$ and domain $\{0,...,N\}$. Consider an alternative binomial distribution with the same domain but success parameter $p' \in (0,1)$, $p' > p$. There will exists $\hat{s}(p,p') \in \{1,...,N-1\}$ such that $Prob(\bar{s}|p') \geq Prob(\bar{s}|p)$ for all $\bar{s} \geq \hat{s}(p,p')$, with at least one strict inequality, and $Prob(\bar{s}|p') < Prob(\bar{s}|p)$ for all $\bar{s} < \hat{s}(p,p')$.*

Proof. Consider any pair of binomial parameters $p' > p$ that satisfy the stated assumptions. To have that $Prob(\bar{s}|p') > Prob(\bar{s}|p)$ we need to satisfy the following inequality:

$$\frac{N!}{\bar{s}!(N-\bar{s}!)} p'^{\bar{s}} \left(1-p'\right)^{N-\bar{s}} > \frac{N!}{\bar{s}!(N-\bar{s}!)} p^{\bar{s}} \left(1-p\right)^{N-\bar{s}}.$$

For the above relationship to be true, we need that $\left(\frac{p'}{p}\right)^{\bar{s}} > \left(\frac{1-p}{1-p'}\right)^{N-\bar{s}}$. Under the assumption that $p' > p$, both sides of the inequality are weakly above 1. In particular, the left-hand side is equal to 1 when $\bar{s} = 0$ and monotonically grows when $\bar{s}$ becomes larger, up to $\left(\frac{p'}{p}\right)^{N}$. On the other hand, the right-hand side is equal to 1 when $\bar{s} = N$, and it is easy to see that monotonically decreases in $\bar{s}$, with maximal value attained at $\bar{s} = 0$ and equal to $\left(\frac{1-p}{1-p'}\right)^{N}$. This argument implies that the two functions will cross for some value of $\tilde{s} \in (0,N)$. Hence, $\left(\frac{p'}{p}\right)^{\bar{s}} \geq \left(\frac{1-p}{1-p'}\right)^{N-\bar{s}}$ for every $s \geq \tilde{s}$ and $\left(\frac{p'}{p}\right)^{\bar{s}} < \left(\frac{1-p}{1-p'}\right)^{N-\bar{s}}$ for every $s < \tilde{s}$.

Notice also that $\sum_{\bar{s}=0}^{N} \frac{N!}{\bar{s}!(N-\bar{s}!)} p'^{\bar{s}} \left(1-p'\right)^{N-\bar{s}} = \sum_{\bar{s}=0}^{N} \frac{N!}{\bar{s}!(N-\bar{s}!)} p^{\bar{s}} \left(1-p\right)^{N-\bar{s}} = 1$, so it must be that $\tilde{s} \geq 1$ and $\tilde{s} < N$, i.e. there is at least one value of $\bar{s}$ for which the likelihood increases and one for which the likelihood decreases when we change the success parameter from p to p', otherwise the total probability of $s \in \{0,..,N\}$ would be different from 1.

We can then conclude that, for each pair $p' > p$, there exists $\hat{s}(p,p') \in \{1,...,N-1\}$, with $\hat{s}(p,p') \geq \tilde{s}$[62] such that $Prob(\bar{s}|p') \geq Prob(\bar{s}|p)$ for all $\bar{s} \geq \hat{s}(p,p')$, with at least one strict inequality, and $Prob(\bar{s}|p') < Prob(\bar{s}|p)$ for all $\bar{s} < \hat{s}(p,p')$. □

We can now use Lemma B.1 to prove a result that will be helpful in the equilibrium analysis.

Proposition 4. *Fix any $\mu(\theta_H|\bar{s})$ that is strictly monotone in $\bar{s}$. If $\mu(\theta_H|\bar{s})$ is increasing in $\bar{s}$, then $\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta,A)\mu(\theta_H|\bar{s}) > \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta,B)\mu(\theta_H|\bar{s})$ for each $\theta \in \{\theta_H,\theta_L\}$. If*

[62] More precisely, $\hat{s}(p,p')$ is the first integer value weakly higher than $\tilde{s}$.

$\mu(\cdot|\bar{s})$ is decreasing in $\bar{s}$, then $\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, A)\mu(\theta_H|\bar{s}) < \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, B)\mu(\theta_H|\bar{s})$ for each $\theta \in \{\theta_H, \theta_L\}$.

Proof. Fix any $\theta \in \{\theta_H, \theta_L\}$, and remember our assumption $p_A(\theta) > p_B(\theta)$. Given this observation, we can use Lemma B.1 to directly prove our result.

Indeed, there exists $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$ such that $Prob\left(\bar{s}|\theta, A\right) \geq Prob\left(\bar{s}|\theta, B\right)$ for all $\bar{s} \geq \hat{s}\left(p_A(\theta), p_B(\theta)\right)$ (with at least one strict inequality), and $Prob\left(\bar{s}|\theta, A\right) < Prob\left(\bar{s}|\theta, B\right)$ for all $\bar{s} < \hat{s}\left(p_A(\theta), p_B(\theta)\right)$. This implies that, when $\mu(\cdot|\bar{s})$ is increasing in $\bar{s}$,

$$\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, A)\mu(\theta_H|\bar{s}) > \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, B)\mu(\theta_H|\bar{s})$$

since the binomial distribution under DGP A allocates weakly more weight on each value of $\bar{s} \geq \hat{s}\left(p_A(\theta), p_B(\theta)\right)$ (with at least one strict ranking) and less weight on on each value of $\bar{s} < \hat{s}\left(p_A(\theta), p_B(\theta)\right)$. It does not matter how the weight is reallocated above and below $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$ for two reasons. First, $\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, A) = \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, B) = 1$ and so any loss in probability below $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$ has to be offset with the same increase in probability above $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$. Second, $\mu(\cdot|\bar{s})$ is monotone increasing and so any value above $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$ dominates any value below the threshold.

Using the same argument, we can easily see that, if $\mu(\cdot|\bar{s})$ is decreasing in $\bar{s}$, then the opposite relation holds, i.e.

$$\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, A)\mu(\theta_H|\bar{s}) < \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta, B)\mu(\theta_H|\bar{s}).$$

In this second case, experiment B would shift likelihood from above to below $\hat{s}\left(p_A(\theta), p_B(\theta)\right)$, increasing the overall value of the expectation when $\mu(\cdot, q)$ is decreasing. □

Finally, we can prove an additional result that highlights under what conditions the assumption we make under our monotone belief refinement restricts the set of equilibria.

Lemma B.2. *Assume that $p_A(\theta_H) > p_B(\theta_H) > p_A(\theta_L) > p_B(\theta_L)$. For any possible sender's strategy $q : \{\theta_H, \theta_L\} \to [0, 1]$, a belief $\mu(\theta_H|\bar{s})$ computed consistently with q is (strictly) increasing in $\bar{s}$.*

Proof. Fix any $q(\cdot)$ and consider the resulting belief

$$\mu(\theta_H|\bar{s}) = Prob(\theta_H|\bar{s}, q) = \frac{\mu_0 Prob(\bar{s}|\theta_H, q)}{\mu_0 Prob(\bar{s}|\theta_H, q) + (1 - \mu_0) Prob(\bar{s}|\theta_L, q)}.$$

We need to show that for any $\bar{s} > \bar{s}'$, $\mu(\theta_H|\bar{s}) > \mu(\theta_H|\bar{s}')$. The statement can be rearranged as

$$Prob(\bar{s}|\theta_H, q)Prob(\bar{s}'|\theta_L, q) > Prob(\bar{s}'|\theta_H, q)Prob(\bar{s}|\theta_L, q)$$

where

$$Prob(\bar{s}|\theta, q) = q(\theta)\frac{N!}{\bar{s}!(N-\bar{s})!}p_A(\theta)^{\bar{s}}\left(1-p_A(\theta)\right)^{N-\bar{s}} + (1-q(\theta))\frac{N!}{\bar{s}!(N-\bar{s})!}p_B(\theta)^{\bar{s}}\left(1-p_B(\theta)\right)^{N-\bar{s}}$$

This implies that we need to show

$$\left(q(\theta_H)p_A(\theta_H)^{\bar{s}}\left(1-p_A(\theta_H)\right)^{N-\bar{s}} + (1-q(\theta_H))\,p_B(\theta_H)^{\bar{s}}\left(1-p_B(\theta_H)\right)^{N-\bar{s}}\right)\cdot$$
$$\left(q(\theta_L)p_A(\theta_L)^{\bar{s}'}\left(1-p_A(\theta_L)\right)^{N-\bar{s}'} + (1-q(\theta_L))\,p_B(\theta_L)^{\bar{s}'}\left(1-p_B(\theta_L)\right)^{N-\bar{s}'}\right) \geq$$
$$\left(q(\theta_H)p_A(\theta_H)^{\bar{s}'}\left(1-p_A(\theta_H)\right)^{N-\bar{s}'} + (1-q(\theta_H))\,p_B(\theta_H)^{\bar{s}'}\left(1-p_B(\theta_H)\right)^{N-\bar{s}'}\right)\cdot$$
$$\left(q(\theta_L)p_A(\theta_L)^{\bar{s}}\left(1-p_A(\theta_L)\right)^{N-\bar{s}} + (1-q(\theta_L))\,p_B(\theta_L)^{\bar{s}}\left(1-p_B(\theta_L)\right)^{N-\bar{s}}\right)$$

Given that $\bar{s} - \bar{s}' \geq 1$ can rearrange the inequality as

$$q(\theta_H)q(\theta_L)p_A(\theta_H)^{\bar{s}'}p_A(\theta_L)^{\bar{s}'}\left(1-p_A(\theta_H)\right)^{N-\bar{s}}\left(1-p_A(\theta_L)\right)^{N-\bar{s}}\cdot$$
$$\left(p_A(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_A(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_H)\right)^{\bar{s}-\bar{s}'}\right)+$$
$$q(\theta_H)\left(1-q(\theta_L)\right)p_A(\theta_H)^{\bar{s}'}p_B(\theta_L)^{\bar{s}'}\left(1-p_A(\theta_H)\right)^{N-\bar{s}}\left(1-p_B(\theta_L)\right)^{N-\bar{s}}\cdot$$
$$\left(p_A(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_B(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_H)\right)^{\bar{s}-\bar{s}'}\right)+$$
$$\left(1-q(\theta_H)\right)q(\theta_L)p_B(\theta_H)^{\bar{s}'}p_A(\theta_L)^{\bar{s}'}\left(1-p_B(\theta_H)\right)^{N-\bar{s}}\left(1-p_A(\theta_L)\right)^{N-\bar{s}}\cdot$$
$$\left(p_B(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_A(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_H)\right)^{\bar{s}-\bar{s}'}\right)+$$
$$\left(1-q(\theta_H)\right)\left(1-q(\theta_L)\right)p_B(\theta_H)^{\bar{s}'}p_B(\theta_L)^{\bar{s}'}\left(1-p_B(\theta_H)\right)^{N-\bar{s}}\left(1-p_B(\theta_L)\right)^{N-\bar{s}}\cdot$$
$$\left(p_B(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_B(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_H)\right)^{\bar{s}-\bar{s}'}\right) > 0.$$

At this point, we can notice that:

- $p_A(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_A(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_H)\right)^{\bar{s}-\bar{s}'} > 0$ since $p_A(\theta_H) > p_A(\theta_L)$;
- $p_A(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_B(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_H)\right)^{\bar{s}-\bar{s}'} > 0$ since $p_A(\theta_H) > p_B(\theta_L)$;
- $p_B(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_A(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_A(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_H)\right)^{\bar{s}-\bar{s}'} > 0$ since $p_B(\theta_H) > p_A(\theta_L)$;

- $p_B(\theta_H)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_B(\theta_L)^{\bar{s}-\bar{s}'}\left(1-p_B(\theta_H)\right)^{\bar{s}-\bar{s}'} > 0$ since $p_B(\theta_H) > p_B(\theta_L)$.

Notice that the third condition is the one that would fail if the assumption $p_B(\theta_H) > p_A(\theta_L)$ is not satisfied and the game admits a cheap talk/ signaling flavor in the overall dynamics. Such conditions complete the proof that, no matter $q(\cdot)$, our target inequality is satisfied. This reasoning allows us to conclude that any receiver's beliefs computed according to Bayes' rule for some sender's behavior needs to be monotonically increasing in $\bar{s}$ when $p_B(\theta_H) > p_A(\theta_L)$. □

Lemma B.2 shows that, if the pair of DGPs is such that, no matter the sender's strategy, the likelihood of producing a success is always higher for θ_H, then there cannot be any equilibrium in which a higher $\bar{s}$ is not good news. Indeed, the only structure that is imposed on the beliefs in the statement of the Lemma is that it is computed consistently with what the receiver believes is the strategy of the sender. No matter what this belief is, the resulting belief over θ_H would be monotonically increasing in $\bar{s}$, guaranteeing that, in each equilibrium, such a pattern needs to be satisfied. Given this, Lemma B.2 also shows that our monotone refinement only effectively restricts the set PBE when $p_A(\theta_L) > p_B(\theta_H)$.

Proposition (1)**.** *Fix any pair of DGPs $\{p_A(\theta), p_B(\theta)\}_{\theta\in\{\theta_H,\theta_L\}}$ that satisfy our main assumptions and let $c(\theta, i) = 0$ for all $\theta \in \{\theta_H, \theta_L\}$ and $i \in \{A, B\}$. There is a unique monotone PBE with $q^*(\theta_H) = q^*(\theta_L) = 1$.*

Proof of Proposition 1. Fix any posterior belief $\mu(\theta_H|\bar{s})$ that the receiver might have. Under the restrictions to monotone beliefs, in equilibrium $\mu(\theta_H|\bar{s})$ will be either increasing or decreasing in $\bar{s}$. We can then restrict attention to the sender's best reply to beliefs with this structure, given that, in equilibrium, beliefs must be correct and the sender must best respond to them.

Assume that $\mu(\theta_H|\bar{s})$ is increasing. By Lemma B.1, any type of sender will find it profitable to choose $i = A$. Since beliefs need to be correct in equilibrium, we only need to check whether an increasing $\mu(\theta_H|\bar{s})$ is consistent with $q(\theta_H) = q(\theta_L) = 1$. This result is trivially true given that $p_A(\theta_H) > p_A(\theta_L)$ and $\mu(\theta_H|\bar{s}) = \tilde{\mu}(\theta_H, A|\bar{s})$ for every $\bar{s} \in \{0, ..., N\}$. This allows us to conclude that $q^*(\theta_H) = q^*(\theta_L) = 1$ constitutes a monotone PBE of our communication game, no matter the DGP parameters.

Assume now that $\mu(\theta_H|\bar{s})$ is decreasing. By Lemma B.1, any type of sender will find it profitable to choose $i = B$. Differently from before, a decreasing belief cannot be consistent with the choice $q(\theta_H) = q(\theta_L) = 0$. It is easy to see that, since $p_B(\theta_H) > p_B(\theta_L)$, the receiver's beliefs will need to be increasing in $\bar{s}$, contradicting the initial assumption that $\mu(\cdot|\bar{s})$ is decreasing.

Finally, it is easy to see that we cannot have any mixed PBE that satisfies our monotonicity assumption. Given the absence of any cost and the result in Lemma B.1, for the sender to be indifferent between $i = A$ and $i = B$, the receiver's posterior cannot be strictly monotone, and so they would be inconsistent with our refinement. Given the structure of the game, there is no out-of-equilibrium $\bar{s}$ to account for. For this reason, we can conclude that there exists a unique monotone PBE in which $q^*(\theta_H) = q^*(\theta_L) = 1$. $\square$

Proposition (2)**.** *Fix any pair of DGPs $\{p_A(\theta), p_B(\theta)\}_{\theta\in\{\theta_H,\theta_L\}}$ that satisfy our main assumptions and let the exogenous costs be $c(\theta_L, A) = 0$ and $c(\theta_L, B) = c$. There exists a threshold $\bar{c}$, which depends on the DGP parameters, such that the following statements are true.*

1. *If $p_B(\theta_H) > p_A(\theta_L)$, there is a unique monotone PBE with $q^*(\theta_H) = q^*(\theta_L) = 1$.*
2. *If $p_A(\theta_L) > p_B(\theta_H)$, there is a monotone PBE in which $q^*(\theta_H) = q^*(\theta_L) = 1$. If $c \geq \bar{c}$, there exists a monotone PBE in which $q^*(\theta_H) = 0$ and $q^*(\theta_L) = 1$ (high cost case). If $c < \bar{c}$, there may exist a monotone PBE in which $q^*(\theta_H) = 0$ and $q^*(\theta_L) \in (0, 1)$ (low/intermediate cost case). However, whether this is the case depends on the specific DGP parameters.*

Proof Proposition 2. Let's start by assuming that $p_B(\theta_H) > p_A(\theta_L)$. By Lemma B.2, we know that, for any possible $q(\cdot)$, a belief $\mu(\theta_H|\bar{s})$ computed consistently with $q(\cdot)$ needs to be strictly increasing in $\bar{s}$. This implies that, without any cost, the best response of the sender to any receiver's belief is to choose DGP A. Given that $c(\theta_H, A) = c(\theta_L, A) = 0$, this choice is still available at no cost for both types, and the unique monotone PBE predicts $q^*(\theta_H) = q^*(\theta_L) = 1$.

Let's now study the case in which $p_A(\theta_L) > p_B(\theta_H)$. It is easy to see that the absence of any cost to choose DGP A makes $q^*(\theta_H) = q^*(\theta_L) = 1$ still a monotone PBE. Indeed, if the receiver believes that DGP A is used by both types, the resulting posterior belief will be strictly increasing, and by Lemma B.1, the unique best response for the sender is $q^*(\theta_H) = q^*(\theta_L) = 1$. Lemma B.1 also suggests that this is the only monotone PBE that is consistent with increasing beliefs, since any increasing pattern of $\mu(\cdot|\bar{s})$ would induce both types to choose DGP A.

Hence, to check the existence of alternative monotone equilibria, we need to check whether there exists one supported by decreasing beliefs.

To do so, let's assume that the receiver's beliefs are such that the posterior $\mu(\theta_H|\bar{s})$ is decreasing in $\bar{s}$. By Lemma B.1, θ_H will optimally choose $q^*(\theta_H) = 0$, to minimize the expected number of successes. What is, instead, the best response of θ_L depends on the cost of choosing DGP B. In particular, θ_L will also choose DGP B if the cost is not too high given the benefits, i.e. if

$$\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta_L, B)\mu(\theta_H|\bar{s}) - c > \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta_L, A)\mu(\theta_H|\bar{s}).$$

For the above relationship to be true, it must be that the cost of choosing DGP B is below the increase in expected gain resulting from the choice:

$$c < \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$$

and this relationship needs to be satisfied under the equilibrium beliefs to be consistent with an equilibrium. At this point, we can notice that, if we compute $\mu(\theta_H|\bar{s})$ in a way that is consistent with q and we fix $q^*(\theta_H) = 0$, the right-hand side of the above inequality is increasing in $q(\theta_L)$.

To see that, we can plug in it the definition of $\mu(\theta_H|\bar{s})$ given $q^*(\theta_H) = 0$:

$$\mu(\theta_H|\bar{s}) = \frac{\mu_0 Prob(\bar{s}|\theta_H, B)}{\mu_0 Prob(\bar{s}|\theta_H, B) + (1-\mu_0)\left(q(\theta_L)Prob(\bar{s}|\theta_L, A) + (1-q(\theta_L))Prob(\bar{s}|\theta_L, B)\right)}.$$

At this point we can compute the derivative of $\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$ with respect to $q(\theta_L)$ and we get

$$\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right)^2 \mu(\theta_H|\bar{s}) \frac{(1-\mu_0)}{Prob(\bar{s}|q)} > 0$$

where

$$Prob(\bar{s}|q) = \mu_0 Prob(\bar{s}|\theta_H, B) + (1-\mu_0)\left(q(\theta_L)Prob(\bar{s}|\theta_L, A) + (1-q(\theta_L))Prob(\bar{s}|\theta_L, B)\right).$$

Given the above argument, $\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$ is increasing in $q(\theta_L)$. This is not surprising: the value of choosing DGP B for type θ_L is higher when the receiver assigns a greater probability to the choice of DGP A (i.e. when $q(\theta_L)$ is higher). This is because in these cases a lower $\bar{s}$ will be stronger evidence that the type is θ_H, making the deviation more profitable. The right-hand side will reach the lowest possible value when $q(\theta_L) = 0$ and the highest possible one when $q(\theta_L) = 1$.

Then, if under the assumption that $q(\theta_H) = q(\theta_L) = 0$ the relationship

$$c < \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$$

is satisfied, we can conclude that the cost is low enough, no matter the receiver's beliefs, and also type θ_L would best respond to the decreasing belief by choosing DGP B. However, it is easy to see that this scenario is not possible, since $q(\theta_H) = q(\theta_L) = 0$ would imply increasing receivers' beliefs and so the value of choosing DGP A would be higher even in the absence of a cost, leading to negative gains. This pattern would never be consistent with an equilibrium outcome, since it would also contradict the fact that $q^*(\theta_H) = 0$.

However, given the possible change in the direction of update, we know that the receiver's gain becomes positive at some point, potentially leading to the choice of DGP B if the cost allows.

Given the same argument, $\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$ reaches the highest possible value when $q(\theta_L) = 1$. This implies that if, under the assumption $q(\theta_H) = 0$ and $q(\theta_L) = 1$

$$c > \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s}) = \overline{c}$$

the cost is too high, no matter the receiver's actions, and it can never be optimal for θ_L to choose DGP B: it must be that $q^*(\theta_L) = 1$. To verify whether this separating equilibrium can be supported, we need to understand if $q^*(\theta_H) = 0$ and $q^*(\theta_L) = 1$ can be consistent with decreasing beliefs. Notice that we are considering the case in which $p_A(\theta_L) > p_B(\theta_H)$, so it is easy to see how a higher number of successes must constitute bad news and $\mu^*(\theta_H|\bar{s})$ will need to be decreasing in $\bar{s}$. This implies that the identified combination would constitute an equilibrium, in which type θ_H chooses the DGP that minimizes the expected number of successes and separates itself from θ_L.

Finally, there are cases in which the value of the cost is below $\overline{c}$. When this is the case, we can check whether there exists a value of $\bar{q}$ such that, under the assumption $q(\theta_H) = 0$ and $q(\theta_L) = \bar{q}$ leads to

$$c = \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, B) - Prob(\bar{s}|\theta_L, A)\right) \mu(\theta_H|\bar{s})$$

where $\mu(\cdot|\bar{s})$ computed according to the considered sender's strategy. The imposed cost needs to be above the right-hand side for $q(\theta_L) < \bar{q}$ and below the right-hand side for $q(\theta_L) > \bar{q}$. In this case, for the additional equilibrium to be supported, we need that type θ_L mixes, while, instead, type θ_H always chooses DGP B. For this behavior to be consistent with a monotone PBE, we need the sender's strategy to induce a decreasing $\mu^*(\theta_H, q^*)$. This is not necessarily the case. In particular, looking back at the argument in Lemma B.2, we would need that, for each $\bar{s} > \bar{s}'$

$$\bar{q} p_B(\theta_H)^{\bar{s}'} p_A(\theta_L)^{\bar{s}'} \left(1 - p_B(\theta_H)\right)^{N-\bar{s}} \left(1 - p_A(\theta_L)\right)^{N-\bar{s}} \cdot$$
$$\left(p_B(\theta_H)^{\bar{s}-\bar{s}'} \left(1 - p_A(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_A(\theta_L)^{\bar{s}-\bar{s}'} \left(1 - p_B(\theta_H)\right)^{\bar{s}-\bar{s}'} \right) +$$
$$\left(1 - \bar{q}\right) p_B(\theta_H)^{\bar{s}'} p_B(\theta_L)^{\bar{s}'} \left(1 - p_B(\theta_H)\right)^{N-\bar{s}} \left(1 - p_B(\theta_L)\right)^{N-\bar{s}} \cdot$$
$$\left(p_B(\theta_H)^{\bar{s}-\bar{s}'} \left(1 - p_B(\theta_L)\right)^{\bar{s}-\bar{s}'} - p_B(\theta_L)^{\bar{s}-\bar{s}'} \left(1 - p_B(\theta_H)\right)^{\bar{s}-\bar{s}'} \right) < 0.$$

Whether such a relationship is satisfied depends on the structure of the two DGPs A and B and on the value of $\bar{q}$. So while this PBE generally exists under the appropriate assumption on c, it may be non-monotone. □

Proposition (3). *Fix any pair of DGPs $\{p_A(\theta), p_B(\theta)\}_{\theta \in \{\theta_H, \theta_L\}}$ that satisfy our main assumptions and let the exogenous costs to be $c(\theta_L, A) = c$ and $c(\theta_L, B) = 0$. There exist two thresholds, $\underline{c} < \bar{c}$, which depend on the DGP parameters, such that the following statements are true.*

If $c \leq \underline{c}$, there is a unique monotone PBE in which $q^*(\theta_H) = q^*(\theta_L) = 1$ (low cost case). If, instead, $c > \underline{c}$ but $c < \bar{c}$, there exists a unique monotone PBE in which $q^*(\theta_H) = 1$ and $q^*(\theta_L) \in (0, 1)$ (intermediate cost case). Finally, if $c \geq \bar{c}$, there exists a unique monotone PBE in which $q^*(\theta_H) = 1$ and $q^*(\theta_L) = 0$ (high cost case).

Proof of Proposition 3. In a monotone PBE, the receiver's beliefs can only be increasing or decreasing in $\bar{s}$. For this reason, we can focus on these two cases, study the best responses of the sender, and assess which ones can be consistent with an equilibrium.

Let's start by assuming that given the receiver's beliefs, $\mu(\theta_H|\bar{s})$ is decreasing in $\bar{s}$. It is easy to see both θ_H and θ_L would choose $q^*(\theta_H) = q^*(\theta_L) = 0$, for which no cost is imposed. However, as already discussed in Proposition 1, this cannot be part of an equilibrium because the induced beliefs would be increasing, contradicting our initial assumption.

We can now assume that $\mu(\theta_H|\bar{s})$ is increasing in $\bar{s}$. Under this assumption, by Lemma B.1, it must be the case that $q^*(\theta_H) = 1$, since the high type pays no cost to choose the DGP that, on average, maximizes the number of successes. What is, instead, the best response of θ_L depends on the cost. In particular, θ_L will also choose DGP A if

$$\sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta_L, A)\mu(\theta_H|\bar{s}) - c > \sum_{\bar{s}=0}^{N} Prob(\bar{s}|\theta_L, B)\mu(\theta_H|\bar{s}).$$

For the above relationship to be true, it must be

$$c < \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right) \mu(\theta_H|\bar{s})$$

And this relationship needs to be satisfied under equilibrium beliefs to be consistent with it. At this point, we can notice that, once we fix $q^*(\theta_H) = 1$, the right-hand side of the above inequality decreases in $q(\theta_L)$. To see that, we can plug in it the posterior consistent with $q(\cdot)$, i.e. $\mu(\theta_H|\bar{s})$ given $q^*(\theta_H) = 1$:

$$\mu(\theta_H|\bar{s}) = \frac{\mu_0 Prob(\bar{s}|\theta_H, A)}{\mu_0 Prob(\bar{s}|\theta_H, A) + (1-\mu_0)\left(q(\theta_L)Prob(\bar{s}|\theta_L, A) + (1-q(\theta_L))Prob(\bar{s}|\theta_L, B)\right)}.$$

At this point we can compute the derivative of $\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right) \mu(\theta_H|\bar{s})$ with respect to $q(\theta_L)$ and we get

$$-\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right)^2 \mu(\theta_H|\bar{s}) \frac{(1-\mu_0)}{Prob(\bar{s}|q)} < 0$$

where

$$Prob(\bar{s}|q) = \mu_0 Prob(\bar{s}|\theta_H, A) + (1-\mu_0)\left(q(\theta_L)Prob(\bar{s}|\theta_L, A) + (1-q(\theta_L))Prob(\bar{s}|\theta_L, B)\right).$$

Given the above argument, $\sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right) \mu(\theta_H|\bar{s})$ is decreasing in $q(\theta_L)$. As in Proposition 2, this is not surprising. θ_L has a larger gain from choosing DGP A when the receiver assigns a low probability to this choice—forming more extreme guesses—and, so, when $q(\theta_L)$ is low. This is because a higher $\bar{s}$ will be stronger evidence that the type is θ_H, making the deviation more profitable. This implies that the right-hand side will reach the lowest possible value when $q(\theta_L) = 1$. Then, if under the assumption that $q(\theta_H) = q(\theta_L) = 1$ the relationship

$$c < \sum_{\bar{s}=0}^{N} \left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right) \mu(\theta_H|\bar{s}) = \underline{c}$$

is satisfied, the cost is low enough, and also type θ_L would best respond to the increasing beliefs by choosing DGP A. In addition to this, it is easy to see that, since $p_A(\theta_H) > p_A(\theta_L)$ it must be true that $\mu^*(\theta_H|\bar{s})$ is increasing in $\bar{s}$.[63] This allows us to conclude that, when the cost is low, we have a monotone PBE with $q^*(\theta_H) = q^*(\theta_L) = 1$.

[63]For the formal steps, see Lemma B.2 with $q(\theta_H) = q(\theta_L) = 1$

Given the same argument, $\sum_{\bar{s}=0}^{N}\left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right)\mu(\theta_H|\bar{s})$ reaches the highest possible value when $q(\theta_L) = 0$. This implies that if, under the assumption $q(\theta_H) = 1$ and $q(\theta_L) = 0$

$$c > \sum_{\bar{s}=0}^{N}\left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right)\mu(\theta_H|\bar{s}) = \bar{c}$$

the cost is too high, and there can never be an equilibrium in which type θ_L chooses DGP A: it must be that $q^*(\theta_L) = 0$. To see if a separating equilibrium can be supported, we need to understand if $q^*(\theta_H) = 1$ and $q^*(\theta_L) = 0$ can be consistent with increasing beliefs. Under our assumptions, it must be that $p_A(\theta_H) > p_B(\theta_L)$, so it is easy to see that $\mu^*(\theta_H|\bar{s})$ will need to be increasing in $\bar{s}$. This implies that the identified combination would constitute a monotone equilibrium.

Finally, there are cases in which the value of the cost is intermediate and does not satisfy either of the inequalities considered before. When this is the case, there will be a value of $\bar{q}$ such that, under the assumption $q(\theta_H) = 1$ and $q(\theta_L) = \bar{q}$

$$c = \sum_{\bar{s}=0}^{N}\left(Prob(\bar{s}|\theta_L, A) - Prob(\bar{s}|\theta_L, B)\right)\mu(\theta_H|\bar{s})$$

with the cost being below the right-hand side for $q(\theta_L) < \bar{q}$ and above the right-hand side for $q(\theta_L) > \bar{q}$. In this case, for the condition to be consistent with equilibrium, we need that θ_L mixes, while, instead, type θ_H always chooses DGP A. For this behavior to be consistent with a monotone PBE, we need the sender's strategy to induce an increasing $\mu^*(\theta_H|\bar{s})$. It is straightforward to see that this is always the case, given the fact that $p_A(\theta_H) > p_A(\theta_L) > p_B(\theta_L)$: a higher number of successes always constitutes good news for the receiver.[64] Given this, we can conclude that when the cost is intermediate, there exists a unique monotone PBE in which type θ_H always chooses DGP A, while type θ_L mixes between the two DGPs. □

Lemma B.3. *A monotone PBE always exists.*

Proof. The proof follows directly from Proposition 1, Proposition 2, and Proposition 3. Indeed, in these propositions we have shown that, no matter the DGP parametrization and the combination of costs considered, there is always at least one monotone PBE.

In Proposition 1, there is a unique (pooling) monotone PBE. In Proposition 3, there is a unique monotone PBE for each level of the cost, and the structure of the equilibrium depends

[64] For a formal argument, see Lemma B.2 under the assumption that $q(\theta_H) = 1$ and $q(\theta_L) \in (0, 1)$. There is no ambiguity here, given that this ranking needs to be satisfied in any parametrization. The only "ambiguous" comparison is the one of $p_B(\theta_H)$ and $p_A(\theta_L)$.

on the cost itself. Finally, in Proposition 2, there is a potential multiplicity of monotone PBE. When, for a certain level of cost, two monotone PBEs exist, they induce different monotonic patterns for the receiver's beliefs. □

C Informativeness

We can consider an alternative game in which, rather than being remunerated according to the inference they make about the pair (θ, i), the receiver is remunerated only for the inference about θ. We can use the same payoff function from before, but adapt it to the new scenario:

$$u_R(a|\hat{\theta}, \bar{s}) = 1 - \frac{1}{2} \sum_{\theta \in \{\theta_H, \theta_L\}} \left(\mathbb{1}\{\theta = \hat{\theta}\} - a(\theta|\bar{s})\right)^2$$

where now, $a(\theta|\bar{s})$ is the probability that the receiver assigns to type θ given the evidence $\bar{s}$ and $\mathbb{1}\{\theta = \hat{\theta}\}$ is equal to 1 when the realized type is θ.

When the payoff function follows this form, we can easily see that the expected payoff of the receiver will be

$$\mathbb{E}_\theta[u_R(a|\hat{\theta}, \bar{s})] = 1 - \frac{1}{2} \sum_{\hat{\theta} \in \{\theta_H, \theta_L\}} Prob(\hat{\theta}|\bar{s}) \sum_{\theta \in \{\theta_H, \theta_L\}} \left(\mathbb{1}\{\theta = \hat{\theta}\} - a(\theta|\bar{s})\right)^2 =$$

$$1 - \frac{Prob(\theta_H|\bar{s})}{2} \left[(1 - a(\theta_H|\bar{s}))^2 + a(\theta_L|\bar{s})^2\right] - \frac{Prob(\theta_L|\bar{s})}{2} \left[a(\theta_H|\bar{s})^2 + (1 - a(\theta_L|\bar{s}))^2\right].$$

Given the definition of $a(\cdot|\bar{s})$, it must be that $a(\theta_L|\bar{s}) = 1 - a(\theta_H|\bar{s})$ and so we can simplify the above expression as

$$\mathbb{E}_\theta[u_R(a|\hat{\theta}, \bar{s})] = 1 - Prob(\theta_H|\bar{s})(1 - a(\theta_H|\bar{s}))^2 - (1 - Prob(\theta_H|\bar{s}))a(\theta_H|\bar{s})^2.$$

It is immediate to see that, in equilibrium, the receiver would choose

$$a(\theta_H|\bar{s}) = Prob(\theta_H|\bar{s}).$$

Plugging the optimal action inside the expected payoff of the receiver, we get:

$$\mathbb{E}_\theta[u_R(a|\hat{\theta}, \bar{s})] = 1 - Prob(\theta_H|\bar{s})(1 - Prob(\theta_H|\bar{s}))^2 - (1 - Prob(\theta_H|\bar{s}))Prob(\theta_H|\bar{s})^2 =$$

$$1 - Prob(\theta_H|\bar{s})(1 - Prob(\theta_H|\bar{s}))\left[1 - Prob(\theta_H|\bar{s}) + Prob(\theta_H|\bar{s})\right] \implies$$

$$\mathbb{E}_\theta[u_R(a|\hat{\theta}, \bar{s})] = 1 - Prob(\theta_H|\bar{s})(1 - Prob(\theta_H|\bar{s})).$$

This implies that $\mathbb{E}_\theta[u_R(a|\hat{\theta},\bar{s})]$ is a linear transformation of the variance of θ given $\bar{s}$. Indeed, θ can only assume two values $\theta_H = 1$ and $\theta_L = 0$ and, given the realization of $\bar{s}$, the respective probabilities are $Prob(\theta_H|\bar{s})$ and $1 - Prob(\theta_H|\bar{s})$ and so the variance of the random variable θ given a realization of $\bar{s}$ will be $Var[\theta|\bar{s}] = Prob(\theta_H|\bar{s})(1 - Prob(\theta_H|\bar{s}))$. This implies that

$$\mathbb{E}_\theta[u_R(a|\hat{\theta},\bar{s})] = 1 - Var[\theta|\bar{s}].$$

The smaller the residual uncertainty about θ after the realization $\bar{s}$, the higher the expected payoff of the receiver. From this expression, we can derive the ex-ante expected payoff for the receiver as

$$\mathbb{E}_{\bar{s}}\left[\mathbb{E}_\theta[u_R(a|\hat{\theta},\bar{s})]\right] = 1 - \mathbb{E}_{\bar{s}}\left[Var[\theta|\bar{s}]\right].$$

The expression above links the expected payoff of the receiver and the expected variance of the state after the additional information is revealed.

It is now useful to look at the formula of the correlation between the state θ and the action taken by the receiver $a(\theta_H,\bar{s})$, which is the most straightforward way of measuring the amount of information contained in $\bar{s}$. We can first write the formula of $Corr(\theta,a)$ as

$$Corr(\theta,a) = \frac{\mathbb{E}\left[\theta\cdot a\right] - \mathbb{E}[\theta]\mathbb{E}[a]}{\sqrt{Var[\theta]Var[a]}} = \frac{\mathbb{E}_{\bar{s}}\left[\mathbb{E}\left[\theta\cdot a(\theta_H|\bar{s})|\bar{s}\right]\right] - \mathbb{E}[\theta]\mathbb{E}_{\bar{s}}[a(\theta_H|\bar{s})]}{\sqrt{Var[\theta]Var_{\bar{s}}[a(\theta_H|\bar{s})]}}.$$

At this point, notice that:

- $\mathbb{E}_{\bar{s}}[a(\theta_H|\bar{s})] = \mu_0$, given Bayes rule;
- $\mathbb{E}[\theta\cdot a] = \mathbb{E}_{\bar{s}}\left[\mathbb{E}\left[\theta\cdot a(\theta_H|\bar{s})|\bar{s}\right]\right] = \mathbb{E}_{\bar{s}}\left[a(\theta_H|\bar{s})\cdot\mathbb{E}\left[\theta|\bar{s}\right]\right] = \mathbb{E}_{\bar{s}}\left[a(\theta_H|\bar{s})^2\right]$ since $\mathbb{E}[\theta|\bar{s}] = a(\theta_H|\bar{s})$;
- $Var[a] = \mathbb{E}_{\bar{s}}\left[a(\theta_H|\bar{s})^2\right] - \mu_0^2$.

This implies that

$$Corr(\theta,a) = \frac{\mathbb{E}_{\bar{s}}\left[a(\theta_H|\bar{s})^2\right] - \mu_0^2}{\sqrt{Var[\theta]Var[a]}} = \frac{\sqrt{Var[a]}}{\sqrt{Var[\theta]}} = \frac{\sqrt{\mathbb{E}[a^2] - \mu_0^2}}{\sqrt{Var[\theta]}}.$$

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

$$-\mathbb{E}_{\bar{s}}[Var[\theta|\bar{s}]] = Var[\theta]\cdot Corr(\theta,a)^2 - Var[\theta]$$

and so

$$\mathbb{E}_{\bar{s}}\left[\mathbb{E}_{\theta}[u_R(a|\hat{\theta},\bar{s})]\right] = 1 + Var[\theta]\cdot Corr(\theta,a)^2 - Var[\theta].$$

where $Var[\theta] = \mu_0(1-\mu_0)$. From the formula, it is clear that the higher is the correlation between θ and a, the higher is this ex-ante expected payoff of the receiver, allowing us to use this quantity as a measure for how much information the sender conveys about θ. Notice that, in the main task, the receiver is trying to guess the probability of each realization of (θ, i), and, for this reason, from the action of the receiver we can derive $a(\theta_H|\bar{s}) = a(\theta_H, A|\bar{s}) + a(\theta_H, B|\bar{s})$.

The same argument can be repeated to measure the informativeness of the data about the selected DGP. Indeed, it is enough to assign a value of 1 to DGP A, and a value of 0 to DGP B, and define the alternative payoff function

$$u_R(a|\hat{i},\bar{s}) = 1 - \frac{1}{2}\sum_{i\in\{A,B\}}\left(\mathbb{1}\{i=\hat{i}\} - a(i|\bar{s})\right)^2$$

where $a(i|\bar{s})$ is the probability that the receiver assigns to DGP i. Again, this probability can be derived from the actions of the receiver in the main game since

$$a(i|\bar{s}) = a(\theta_H, i|\bar{s}) + a(\theta_L, i|\bar{s}).$$

Again, this ex-ante expected payoff of the receiver would allow us to measure how much information the data contains about DGP i.

D Estimation of Receivers' Beliefs

D.1 Suggestive Evidence

Section 4.2.1 and Section 4.2.2 provide suggestive evidence that receivers engage in a form of Bayesian updating. In particular, we find that their guesses systematically vary with the number of signals they observe. This pattern indicates that receivers respond to information in a directionally consistent manner, providing motivation for modeling belief formation through a Bayesian lens—albeit one that may allow for deviations from perfect Bayesian updating.

D.2 Model

We formalize belief formation in this experiment with the following model:

$$p(\theta, DGP \mid \bar{s}) = \frac{P(\theta)\, P(DGP \mid \theta)\, P(\bar{s} \mid \theta, DGP)^{\beta}}{\sum_{\tilde{\theta}, D\tilde{G}P} P(\tilde{\theta})\, P(D\tilde{G}P \mid \tilde{\theta})\, P(\bar{s} \mid \tilde{\theta}, D\tilde{G}P)^{\beta}} \tag{1}$$

where:

- $p(\theta, DGP \mid \bar{s})$ is the (measured) probability of a pair (θ, DGP) given signal $\bar{s}$,
- $P(\theta)$ is the (known) prior probability of type θ, equal to 0.5 for all treatments,
- $P(DGP \mid \theta)$ is the perceived probability (strategic treatments: to be estimated, otherwise assumed to be known) that type θ chooses a certain DGP,
- $P(\bar{s} \mid \theta, DGP)$ is the (known) likelihood of observing $\bar{s}$ positive signals given θ and DGP,
- β is the coefficient of underreaction (< 1) / overreaction (> 1) to information.

This model captures two key sources of heterogeneity in belief formation. First, it allows different participants to hold different priors about the sender's behavior—that is, their perceived probability $P(DGP \mid \theta)$ that a given type chooses a particular data-generating process. This flexibility is central, as differences in priors can lead to fundamentally different inferences even when participants are exposed to the same signal. Second, the model allows for heterogeneity in how participants react to information, captured by the parameter β, which scales the weight placed on the likelihood. This accommodates deviations from full Bayesian updating, such as underreaction ($\beta < 1$) or overreaction ($\beta > 1$), in a disciplined way.

Beyond capturing individual differences, the model provides several practical advantages. It enables us to fully characterize behavior in a compact and interpretable way (3 parameters), reducing the high dimensionality of our data (2 actions x 2 types x 4 signal realizations = 16 distinct belief values) into a small number of parameters. In addition, by recovering participants' priors over sender behavior, we gain a clear counterpart to the sender-side data, allowing us to directly assess whether the structurally recovered beliefs align with observed Sender strategies.

Note that, while Grether-type distortions to priors (via an additional parameter α on the prior) are often considered in the literature, such a parameter is not separately identified in our framework and is therefore excluded.

D.3 Estimation

Estimation is carried out at the individual level using a Minimum Distance Estimator. Specifically, we minimize the Euclidean distance between the model's predicted beliefs and the beliefs reported by each participant across all signal realizations. The data provide $p(\theta, DGP \mid \bar{s})$ directly, while $P(\theta)$ and $P(\bar{s} \mid \theta, DGP)$ are assumed to be known from the experimental design. We estimate the participant-specific parameters β and $P(DGP \mid \theta)$, which are separately identified under our assumptions.

While the model does not include an explicit stochastic component, it nonetheless captures noise in behavior through the estimation procedure. The model is over-identified: each individual provides more belief observations than there are parameters to estimate. Rather than imposing a specific functional form for noise, we adopt an agnostic approach by minimizing the distance between predicted and observed beliefs. It is this estimation procedure that absorbs deviations from the model without restricting how these errors enter. Modeling noise explicitly is non-trivial in our setting, as belief reports are constrained to four non-negative numbers summing to 100. Imposing a specific structure on such data could bias inference, so we rely instead on the flexibility of the distance-minimization approach to account for residual variation. For technical reasons, we cap the value of β to 10. Figure 9 shows that such an assumption is not binding in practice, with almost all the subjects resulting in a value strictly below our upper bound.

D.4 Estimation Results

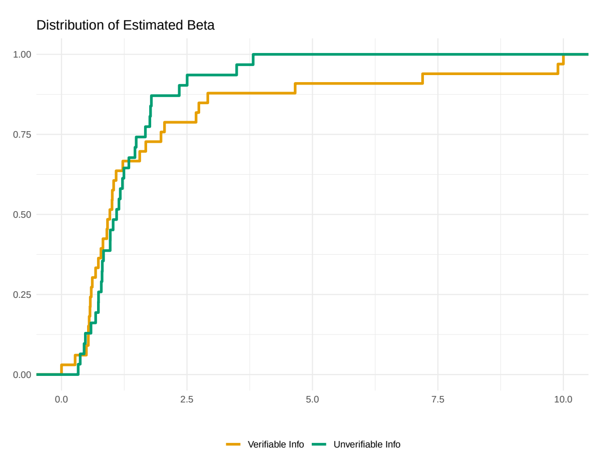


Figure 9: Individual estimates of β.

Figure 9 shows the distribution of β resulting from our estimation exercise.

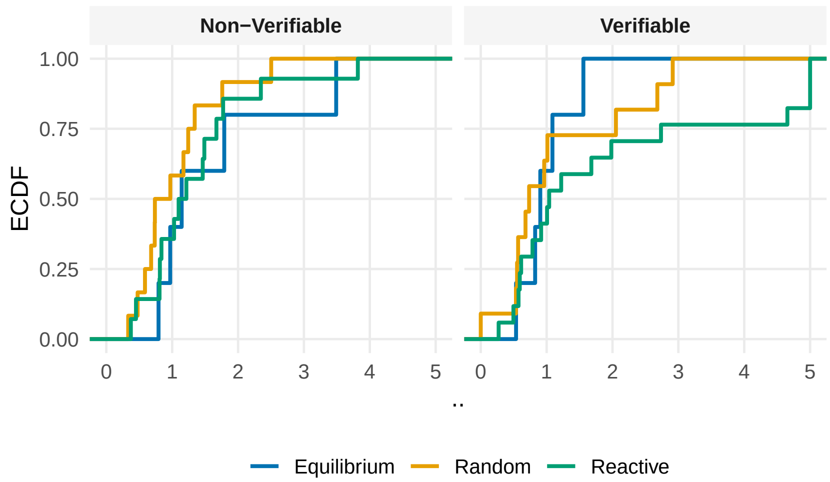


Figure 10: Individual estimates of β by receiver cluster.

Figure 10 reports the distribution of estimated β across receivers' clusters. While there are some differences in the group-specific distribution for the verifiable information treatments, these are not statistically significant. Overall, this pattern confirms that our categorization of receivers properly captures differences in behavior due to their strategic inference, rather than the effects of probabilistic mistakes.

D.5 Fit

The model fits the data well, reinforcing the validity of this approach.

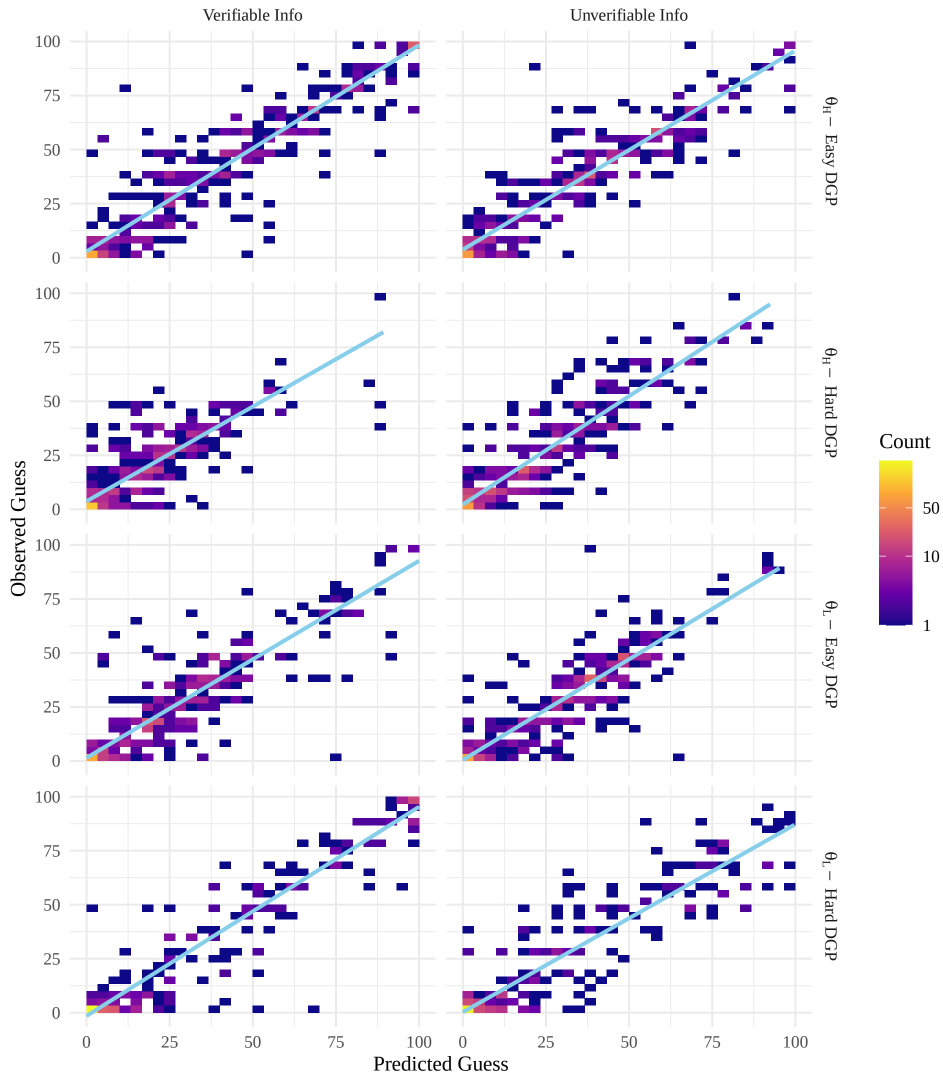


Figure 11: Fit of the estimated model.

E Additional Results

E.1 Learning

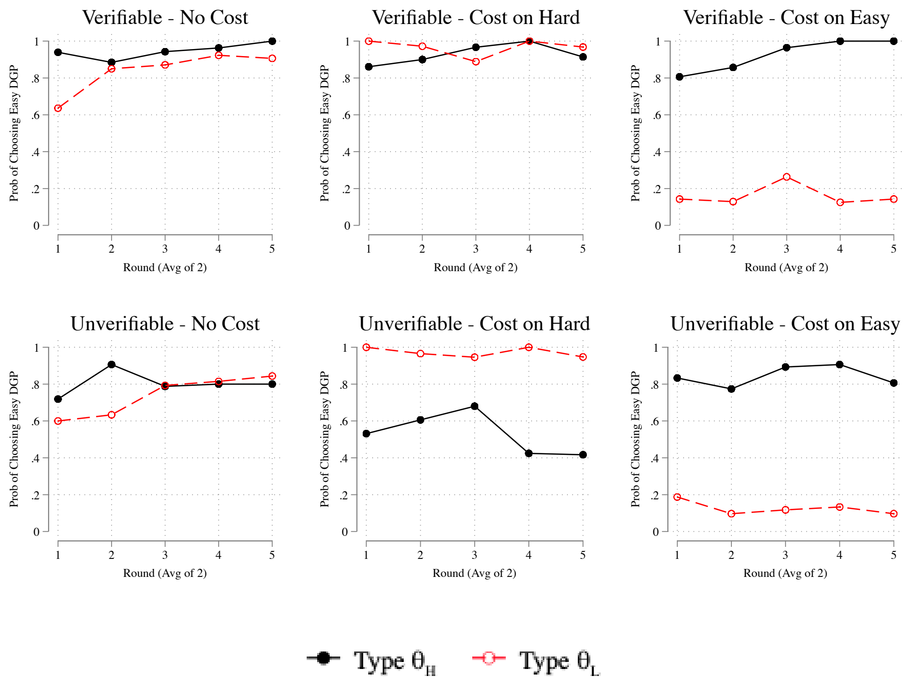


Figure 12: Sender's choice of DGP A over rounds.

Figure 12 plots the evolution over rounds of the senders' average probability of choosing the Easy DGP for each type in every treatment. Figure 13 plots the evolution over rounds of the receivers' guesses for each combination of sender's type and DGP, conditioning on the treatment and on the signal realization. Some learning occurs, mainly for the sender. For this reason, in our main analysis, we focus on the second block for each treatment, but we show in Section G that we get consistent results by including all the rounds.

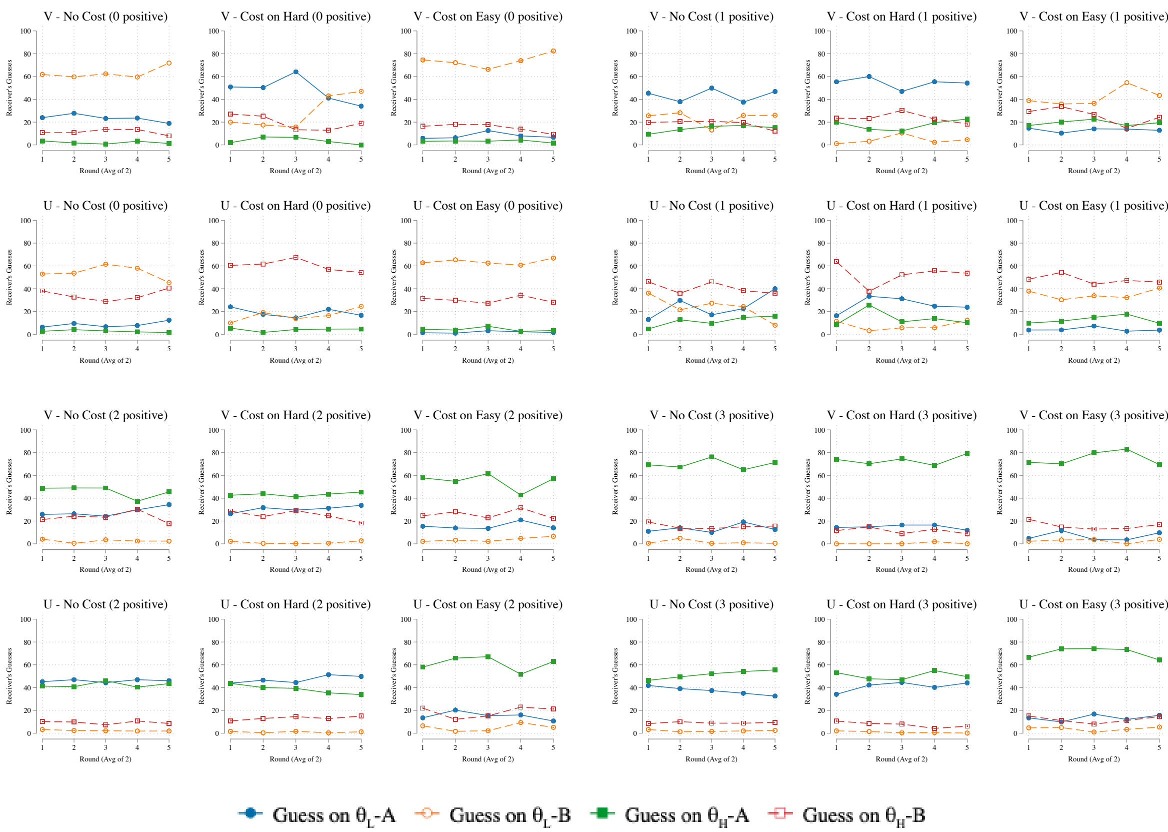


Figure 13: Receivers' average guesses over rounds.

E.2 Exogenous Treatments: Additional Results

Table 8: Information processed, measured as the ex-ante expected payoff of the receiver

	Verifiable Information			Unverifiable Information		
	Joint Info	Info on Type	Info on DGP	Joint Info	Info on Type	Info on DGP
$\theta_H : A - \theta_L : A$	77.98	78.43	99.30	74.52	74.52	100.00
	[81.79]	[81.79]	[100.00]	[75.71]	[75.71]	[100.00]
$\theta_H : B - \theta_L : A$	73.36	73.98	74.05	79.91	79.95	79.95
	[75.23]	[75.23]	[75.23]	[81.75]	[81.75]	[81.75]
$\theta_H : A - \theta_L : B$	92.72	93.50	93.83	89.84	90.03	90.79
	[93.93]	[93.93]	[93.93]	[93.52]	[93.52]	[93.52]
Random	66.77	83.23	77.51	66.48	75.74	84.27
	[69.37]	[83.97]	[78.61]	[69.03]	[76.45]	[85.75]

Notes: A: Easy DGP; B: Hard DGP. In the square brackets, we report the theoretical prediction given the DGP selection rule. Despite some of these gaps being significant, they tend to be smaller than the ones in the strategic treatments (see Table 6).

E.3 Senders: Additional Results

E.3.1 Individual Behavior

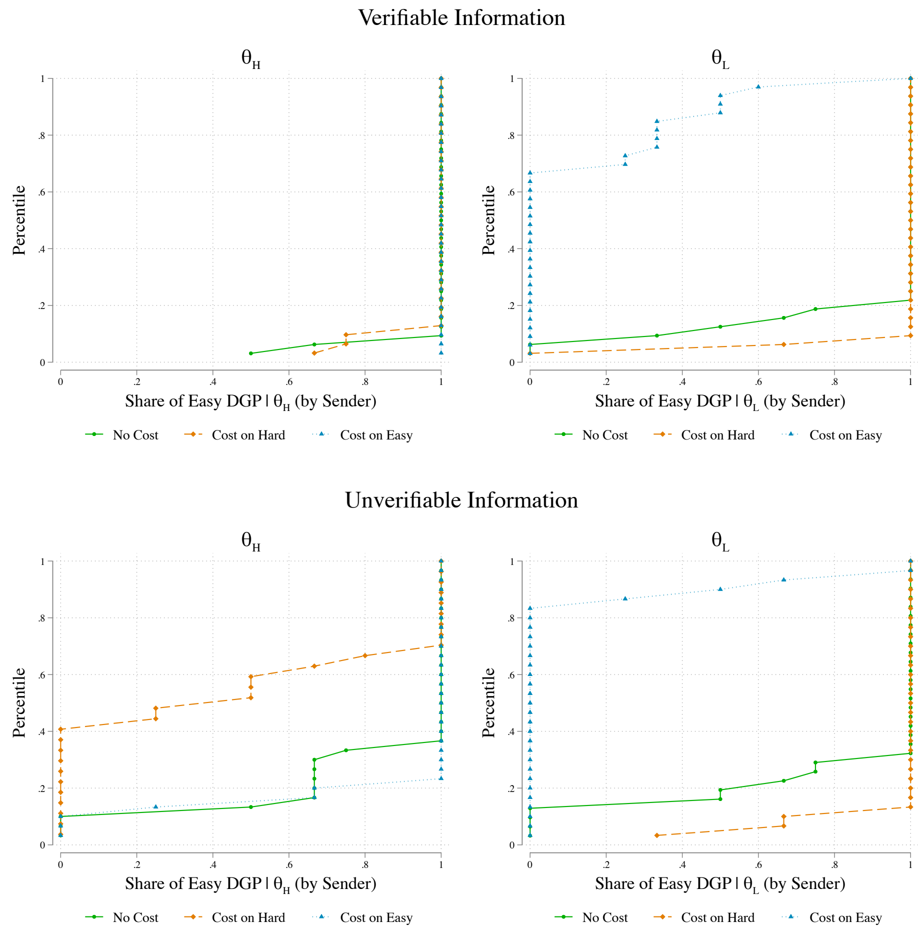


Figure 14: Distribution of sender's DGP choices.

As we could have expected from Table 4, the senders' choices tend to be noisier in the unverifiable information treatments, with some unpredicted variation being detected when there is no selection cost. This pattern is consistent with the existing uncertainty over the receivers' inference, which can heavily depend on the expectations formed on the senders' behavior. For both treatments with a cost on the Easy DGP, we mainly detect some deviations for type θ_L, with some subjects deciding to pay the DGP selection cost. Finally, when looking at the $U - C_{Hard}$ treatment, we can see that the seemingly noisy behavior of type θ_H is again consistent with equilibrium forces. Indeed, roughly a third of the senders choose predominantly the Easy DGP, another third choose the Hard DGP, while the remaining subjects mix between the two. These alternative behaviors are consistent with the possible patterns in receivers' beliefs, due to the multiplicity of equilibria in this setting. Other than this variation in the $U - C_{Hard}$ treatment, Figure 14 does not seem to suggest any other meaningful heterogeneity.

It is important to highlight that, while Figure 14 focuses on the second part of our experiment, in which each sender plays each task for 5 rounds, the same pattern generalizes to different samples. Figure 15 reproduces this figure, restricting attention to senders who played in each type realization at least twice in the last 5 rounds. Figure 32 in Appendix G.1 reproduces the analysis considering all rounds. We find that the overall pattern stays the same.

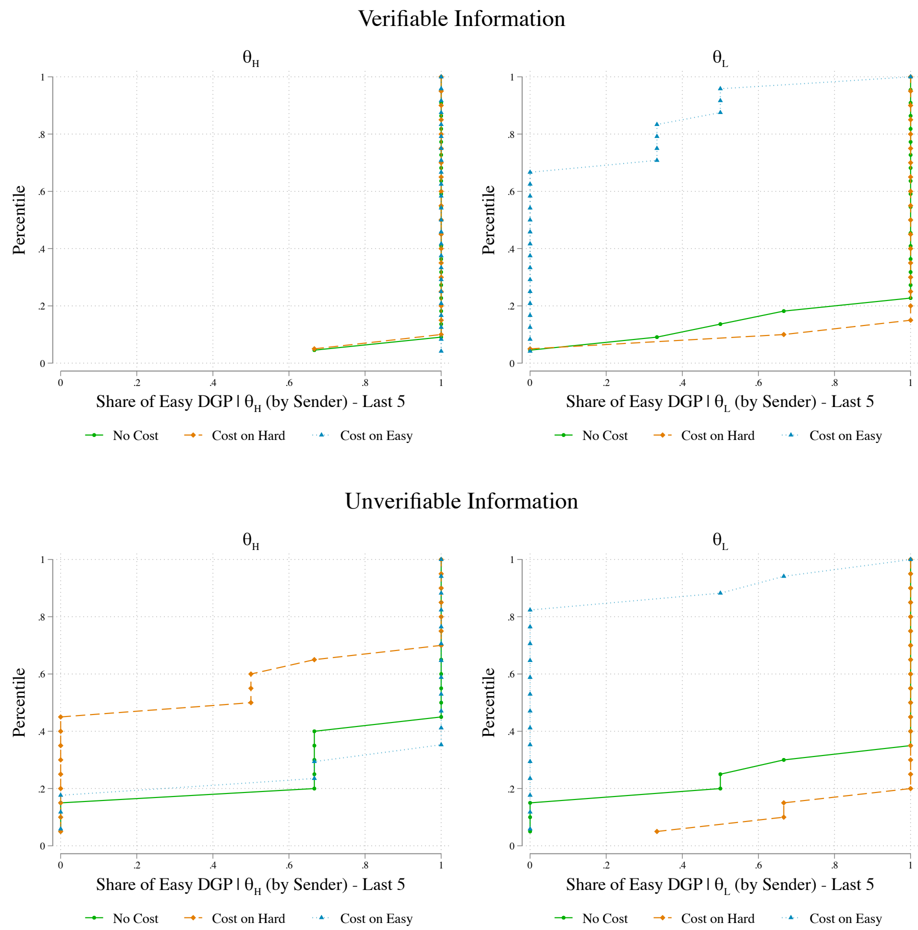


Figure 15: Distribution of sender's DGP choices in the second block, restricting attention to senders who played at least twice in each type realization.

E.3.2 Senders' Beliefs

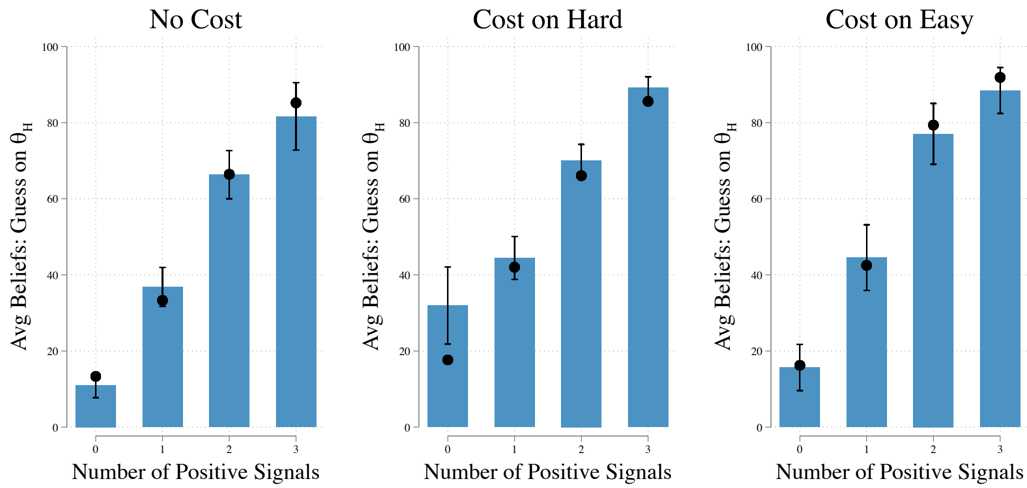


Figure 16: Senders' beliefs in the verifiable information treatments.

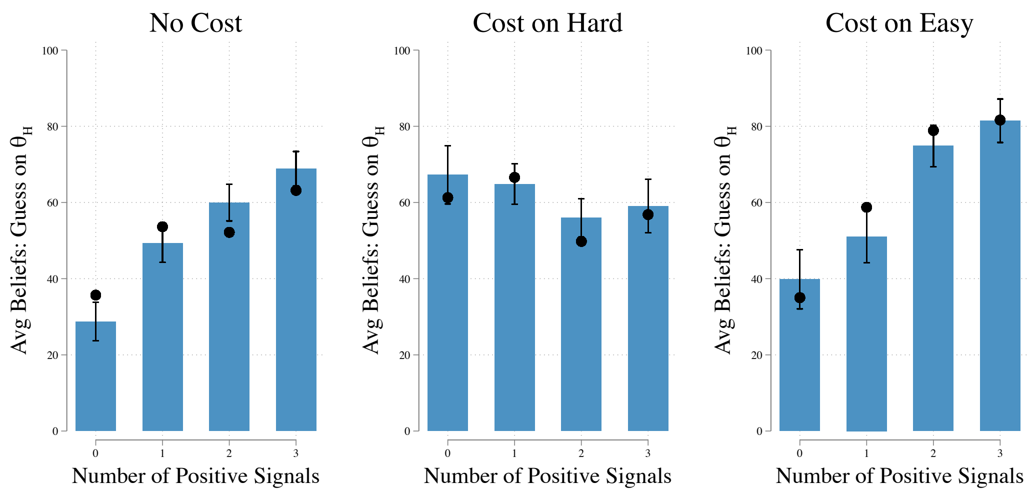
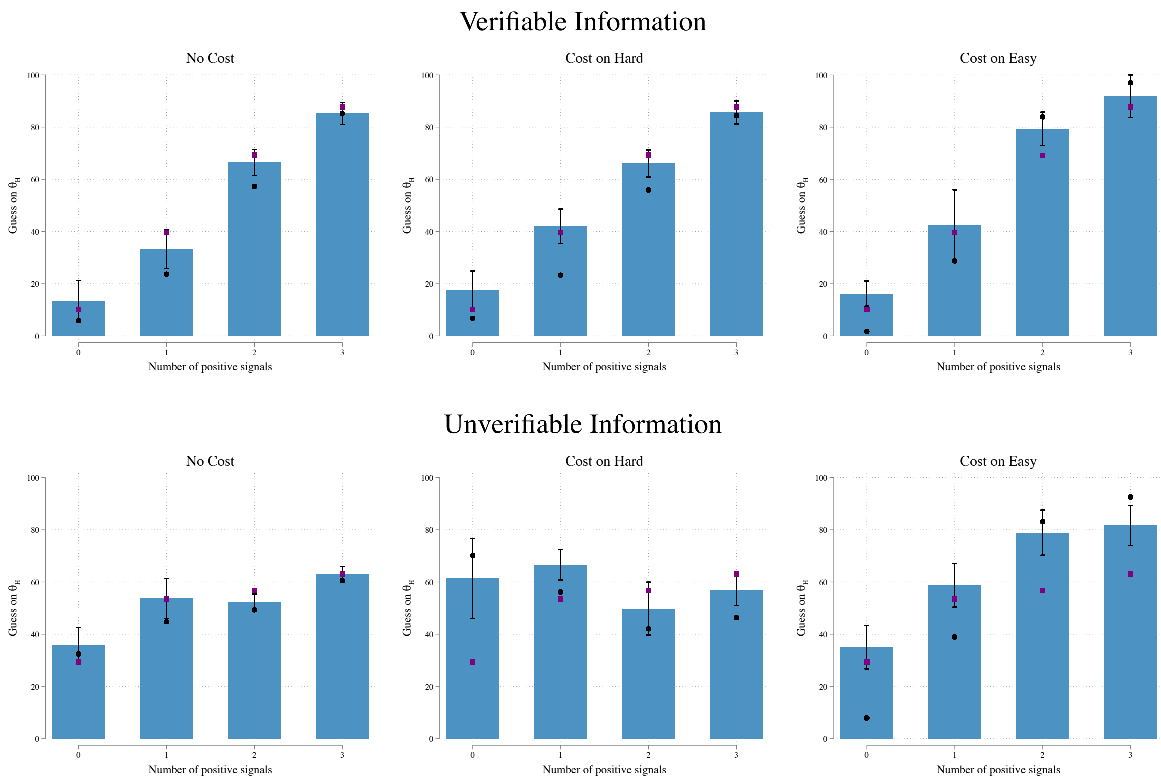


Figure 17: Senders' beliefs in the unverifiable information treatments.

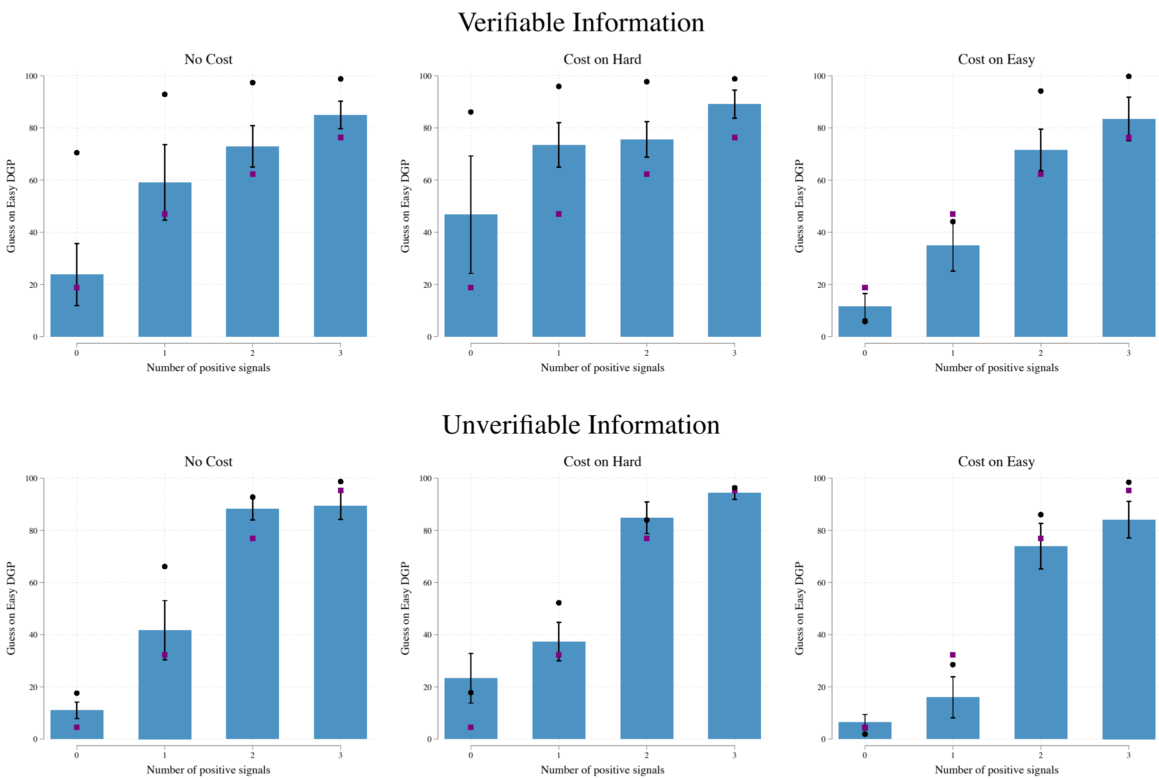
Notes: The dots represent the average receivers' guesses about the type. The bars are the 95% confidence intervals.

E.4 Receivers' Aggregate Guesses: Additional Results



Notes: The black dots represent the Bayesian benchmarks, while the purple squares represent the random DGP benchmark. The bars are the 95% confidence intervals.

Figure 18: Average guesses about the sender's type.



Notes: The black dots represent the Bayesian benchmarks, while the purple squares represent the random DGP benchmark. The bars are the 95% confidence intervals.

Figure 19: Average guesses about the chosen DGP.

E.5 Bayesian and Random DGP Benchmarks

Table 9: Distance from Bayesian and Distance from Random DGP (by treatment).

	Verifiable Information		**Unverifiable Information**	
	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP
C_0	31.53 [0]	24.62 [35.72]	16.09 [0]	15.44 [14.45]
C_{Hard}	26.98 [0]	27.38 [38.18]	20.93 [0]	22.47 [22.05]
C_{Easy}	28.46 [0]	27.77 [36.26]	31.01 [0]	26.30 [34.41]
$\theta_H : Easy - \theta_L : Easy$	12.82 [0]	41.95 [41.31]	8.95 [0]	29.05 [26.04]
$\theta_H : Hard - \theta_L : Easy$	10.40 [0]	43.80 [45.13]	9.63 [0]	46.04 [45.42]
$\theta_H : Easy - \theta_L : Hard$	10.80 [0]	43.22 [44.32]	10.81 [0]	49.03 [48.77]
Random	16.51 [0]		15.63 [0]	

Notes: The top part of the table reports the average distances in the main strategic treatments, while the bottom part reports the average distances in the exogenous treatments. For the random DGP selection treatment, the random DGP benchmark, and the Bayesian benchmark coincide. For reference, the distance implied by the behavior of a Bayesian receiver is reported in square brackets.

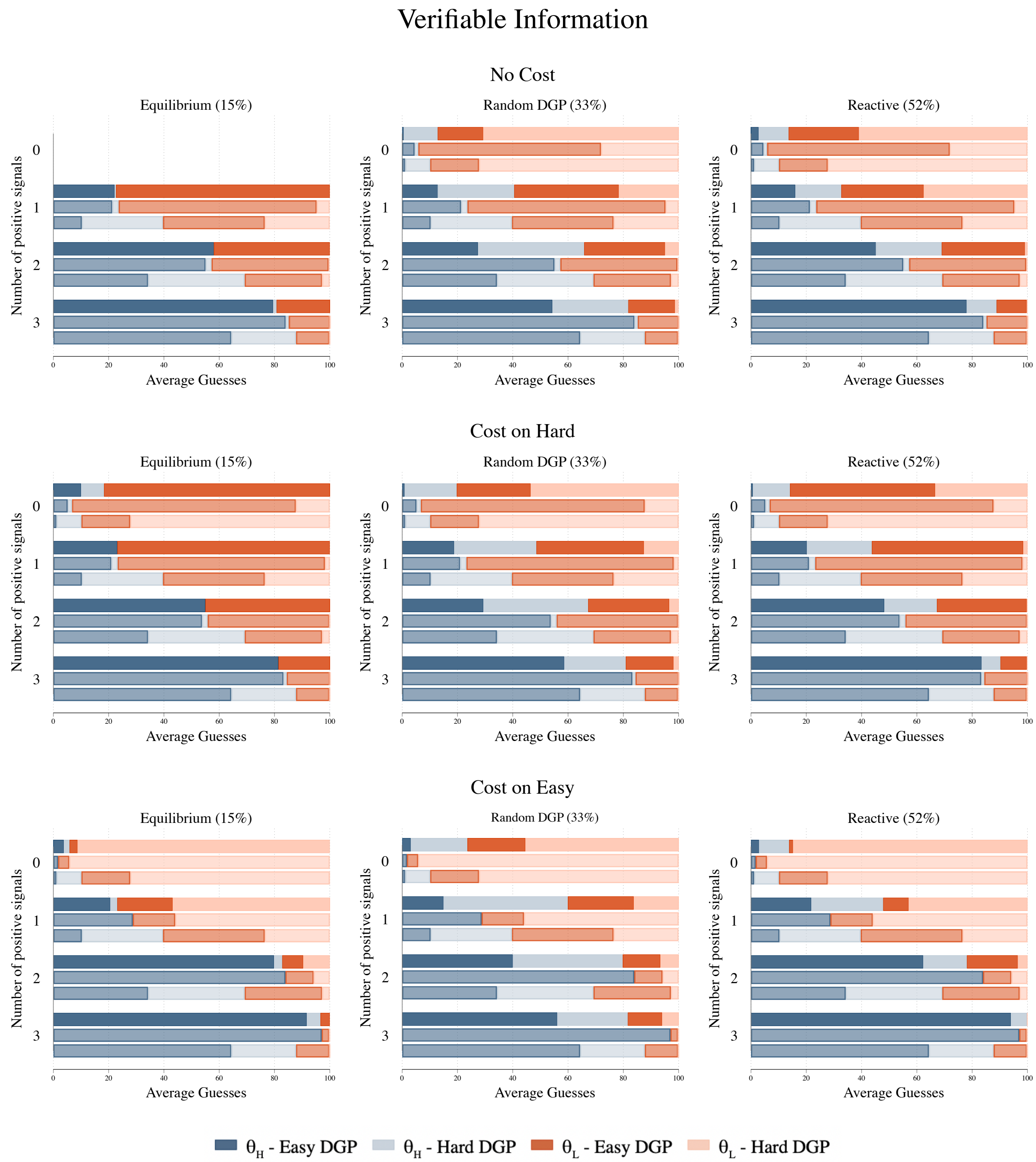
Table 9 reports the average distance from the Bayesian guesses in all our treatments. The first thing we can notice from the table is the fact that subjects induce strictly positive distances from the Bayesian benchmark in all treatments (p-values < 0.01). However, we can also easily see that the deviations in the strategic treatments tend to be higher than those in exogenous treatments (as we could expect from the analysis in Section 4.2.2). In particular, for each information environment, we can directly compare each cost parametrization with the exogenous DGP selection rule aiming to replicate the theoretical predictions. To account for possible differences in the data realizations, we can control for the distribution of positive signals when performing the comparison. We find that the distance from the Bayesian benchmark is significantly higher in the strategic treatments than in the exogenous ones (p-values < 0.01). While not surprising, given the discussion in Section 4.2.1 and Section 4.2.2, this result confirms that, even accounting for possible over-reaction and under-reaction to the inference—which would be partially offset by looking at the average guesses—we find that subjects are much closer to the Bayesian guesses in the deterministic treatments than in the strategic ones. The same ranking holds if we compare the strategic treatments with the random ones (p-values < 0.01, except for $U - C_{Hard}$ and Random for which p-value < 0.05). The only comparison for which we detect no significant difference is the one between $U - C_0$ and Random. However, in this treatment, the Bayesian guesses are closer than they should be to the theoretical predictions of the random treatment, making it more challenging to detect a difference.

Table 9 also reports the average distance from the random DGP benchmark in all our treatments. Again, we can compare such measures in the strategic treatments with those of the corresponding deterministic treatments. As expected, we find that the distance from the random benchmark is always significantly smaller in the treatments with the strategic interaction (p-values < 0.01). This supports the qualitative conclusion in Section 4.2.2 and our intuition that the guesses under random DGP selection are a useful benchmark to consider.

E.6 Receivers' Heterogeneity: Additional Results

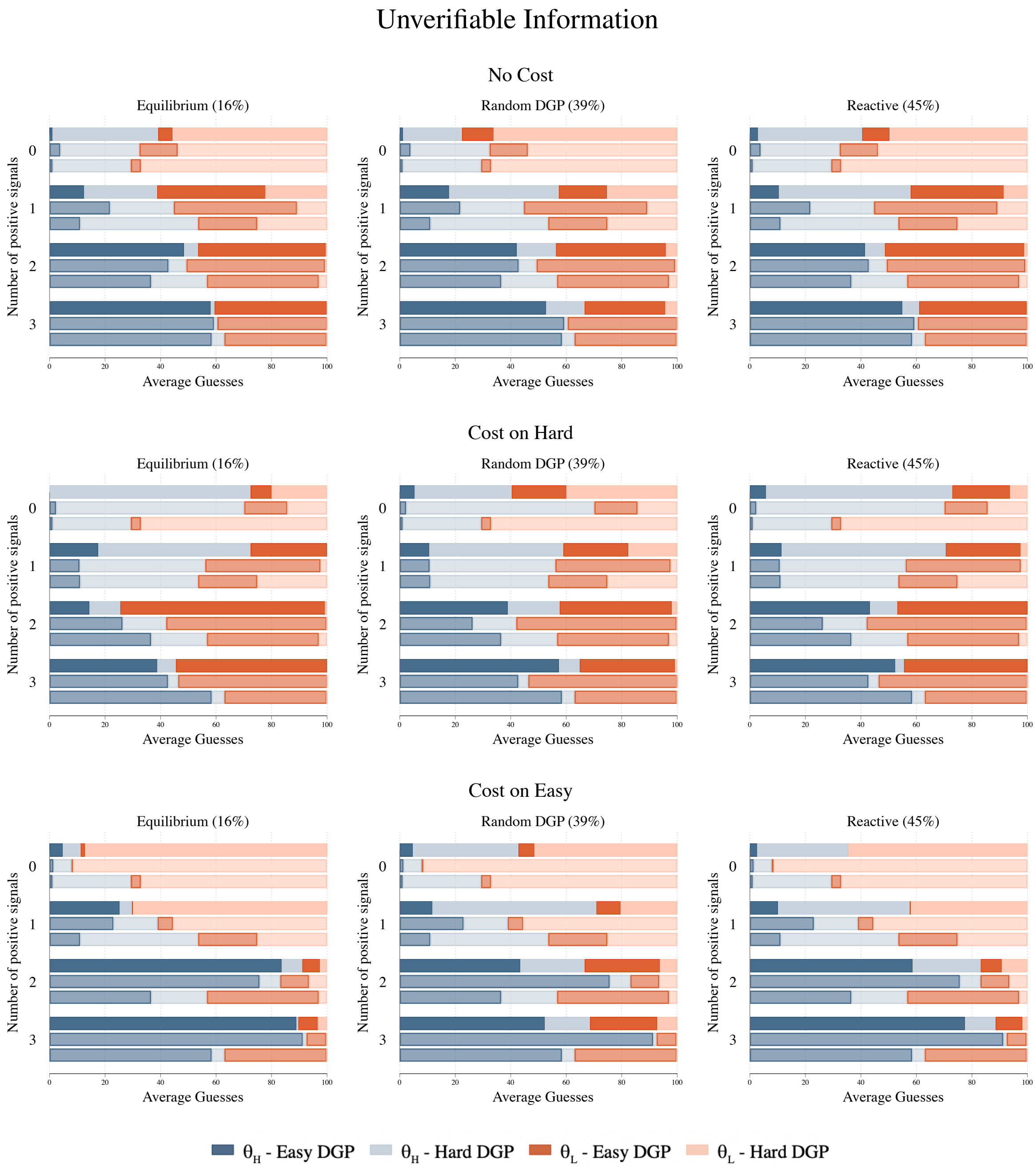
E.6.1 Receivers' Guesses

Our interpretation of the three groups comes directly from the analysis of their behavior. For each group, Figure 20 and Figure 21 summarize the average guesses conditional on a certain realization $\bar{s}$ (top bar) and compare them to the Bayesian guesses and to the ones consistent with the random DGP benchmark (respectively, middle and bottom bar). It is easy to see how, in all treatments, the equilibrium group is indeed the one exhibiting a behavior consistently in line with the theoretical predictions, no matter the specific cost parametrization. On the other hand, the random DGP group is the one whose choices tend to be more in line with the random selection of the data-generating process (see Figure 3). Finally, our third group, the reactive one, displays intermediate behaviors, with a tendency to react to the existence of a cost. While the interpretation of the three groups is consistent across information environments, there are also some important differences. In the verifiable information environment, the behavior of the reactive group when costs are absent is closer to the random DGP benchmark than to the



Notes: For each $\bar{s}$, the top (darker) bar represents receivers' guesses in the data, the middle (lighter) bar represents the Bayesian benchmark, and the bottom (lighter) bar represents the random DGP selection benchmark. There are no bars for $\bar{s} = 0$ in the equilibrium group of the $V - C_0$ treatment, since in the second block, subjects in this group never saw that realization (which is unlikely given the sender's behavior). A caveat of this figure is that each bar is computed on a subset of the observations, and this reduces the sample mostly for the equilibrium group, which is our smallest one.

Figure 20: Average receivers' guesses by number of positive signals and cluster in the verifiable information treatments.



Notes: For each $\bar{s}$, the top (darker) bar represents receivers' guesses in the data, the middle (lighter) bar represents the Bayesian benchmark, and the bottom (lighter) bar represents the random DGP selection benchmark. A caveat of this figure is that each bar is computed on a subset of the observations, and this reduces the sample mostly for the equilibrium group, which is our smallest one.

Figure 21: Average receivers' guesses by number of positive signals and cluster in the unverifiable information treatments.

Bayesian one. This is, instead, not the case in the unverifiable information environment, where the reactive group tends to be more similar to the Bayesian benchmark. This difference can be partially explained by the fact that Bayesian and random predictions are closer in the second case, making the distinction between the two benchmarks harder to achieve. When a cost is imposed on the Hard DGP, we find again that the reactive group in the unverifiable information environment tends to be more consistent with Bayesian predictions than the one in the verifiable information case. Indeed, given the effect of the cost, this group tends to respond to signaling content of low realizations of $\bar{s}$, which is present even with noisy beliefs about the high type—as long as the incentives of the low type are correctly inferred. This substantially affects the direction of the inference, in a direction consistent with the optimal one.[65] Finally, when a cost is imposed on the Easy DGP, the three groups follow extremely similar behavior across environments, and in both cases, this treatment is the one in which we can observe the starkest difference between groups—and this pattern is particularly helpful to achieve our classification, mainly in the unverifiable information environment.

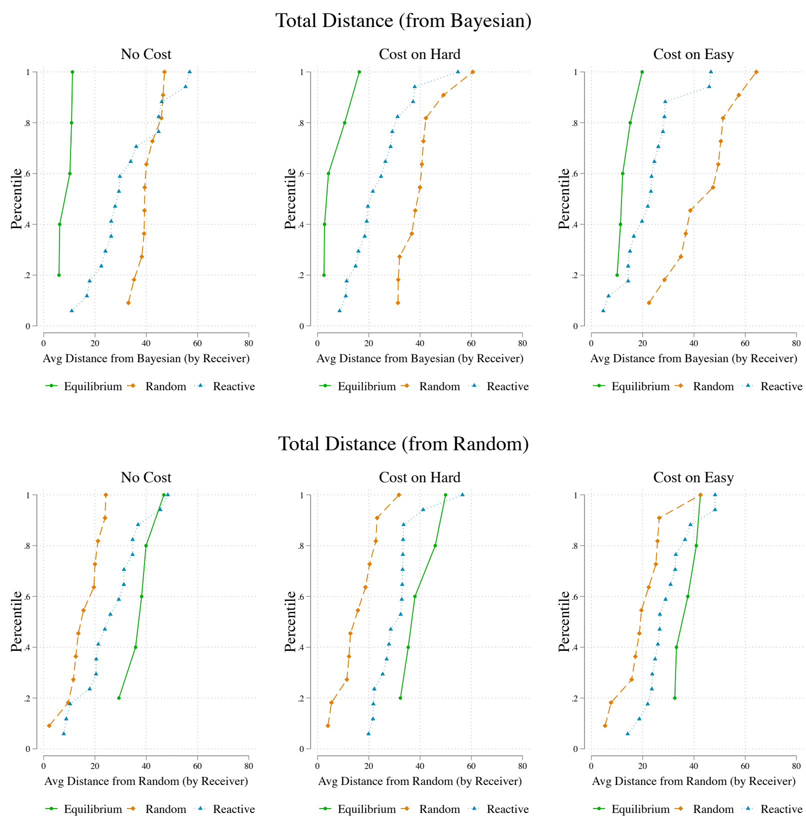


Figure 22: Distribution of receiver's distances from the Bayesian and from the Random DGP benchmark for each cluster in the verifiable information treatments.

[65] We find that 2 subjects in this treatment behave as if the sender played a perfect separating equilibrium. These subjects are categorized as equilibrium by our algorithm, and this partially inflates the average deviations of this group, given the sender's behavior reported in Table 4.

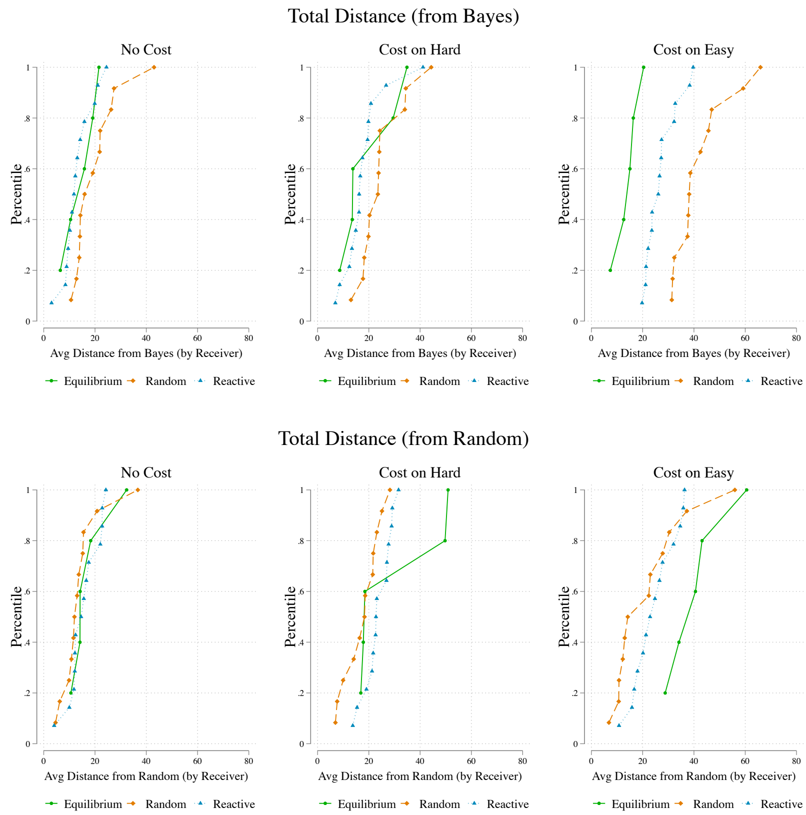


Figure 23: Distribution of receiver's distances from the Bayesian and from the Random DGP benchmark for each cluster in the unverifiable information treatments.

E.6.2 Receivers' Beliefs

Figure 26 shows that receivers' elicited beliefs are consistent with our interpretation of the three groups. Other than looking qualitatively different in ways that are consistent with our analysis, we can also detect meaningful statistical differences between the estimated beliefs. In particular, for the verifiable information treatment, we find that all comparisons yield statistical differences,[66] with two exceptions: in the $V - C_0$ treatment, the comparison of $P(Easy|\theta_L)$ between the random group and the reactive group is not significant, and in the $V - C_{Easy}$, the comparison of $P(Easy|\theta_L)$ between the equilibrium group and the reactive group are not statistically significant.[67] These insignificant comparisons are consistent with our interpretation of the groups, since in the $V - C_0$ treatment, no specific difference is predicted between the equilibrium and the random group, while in the $V - C_{Easy}$ treatment, it is reasonable that the introduction of a cost improves more heavily the inference about the strategy of type θ_L. Similar meaningful patterns are present in the unverifiable information treatment, where we detect differences in the $U - C_0$ treatment between the random and the equilibrium group (p-values < 0.05). We also observe significant differences for $P(Easy|\theta_L)$ between the reactive and the random group (p-value < 0.05). As we could have expected from Figure 21, no difference in beliefs is detected between the equilibrium and the random group in this treatment. Moving to the treatments with a cost, we find that the only insignificant comparisons are the ones comparing $P(Easy|\theta_H)$ for the random and reactive group, and the ones comparing $P(Easy|\theta_L)$ for the equilibrium and reactive group.[68] Again, this is consistent with our interpretation of the clusters and, in particular, of the reactive group, which is directly affected by the cost when making the inference.[69]

[66]The difference in $P(Easy|\theta_H)$ for the random and the reactive group in the $V - C_0$ treatment, the difference in $P(Easy|\theta_H)$ for the equilibrium and the reactive group in the $V - C_{Easy}$ treatment and the difference in $P(Easy|\theta_L)$ for all groups in the $V - C_{Hard}$ treatment have p-values < 0.1. The other p-values < 0.05.

[67]We can repeat the same tests using the elicited beliefs. We find similar patterns, even if the comparisons are noisier. In particular, we do not detect differences across groups for $P(Easy|\theta_L)$ in the $V - C_{Easy}$ treatment, while we detect differences in the beliefs $P(Easy|\theta_H)$, which is considered higher for the equilibrium and the reactive group than for the random. For the $V - C_{Hard}$ treatment, we do not detect differences between the equilibrium and the reactive group, while the beliefs elicited for the random group differ from the others. Again, we define statistically significant comparisons with p-values strictly below 0.1.

[68]For the other comparisons, p-values < 0.1. Clustering standard errors at the session level makes some comparisons not significant in the $U - C_{Hard}$ treatment, but this does not affect our overall interpretation.

[69]Again, similar patterns can be detected when we compare the elicited beliefs. In particular, in the $U - C_0$ and $U - C_{Easy}$ treatment, we can only derive significant differences between the equilibrium and the random group, while for $U - C_{Hard}$ treatment, we mainly find significant changes between the reactive and the random group.

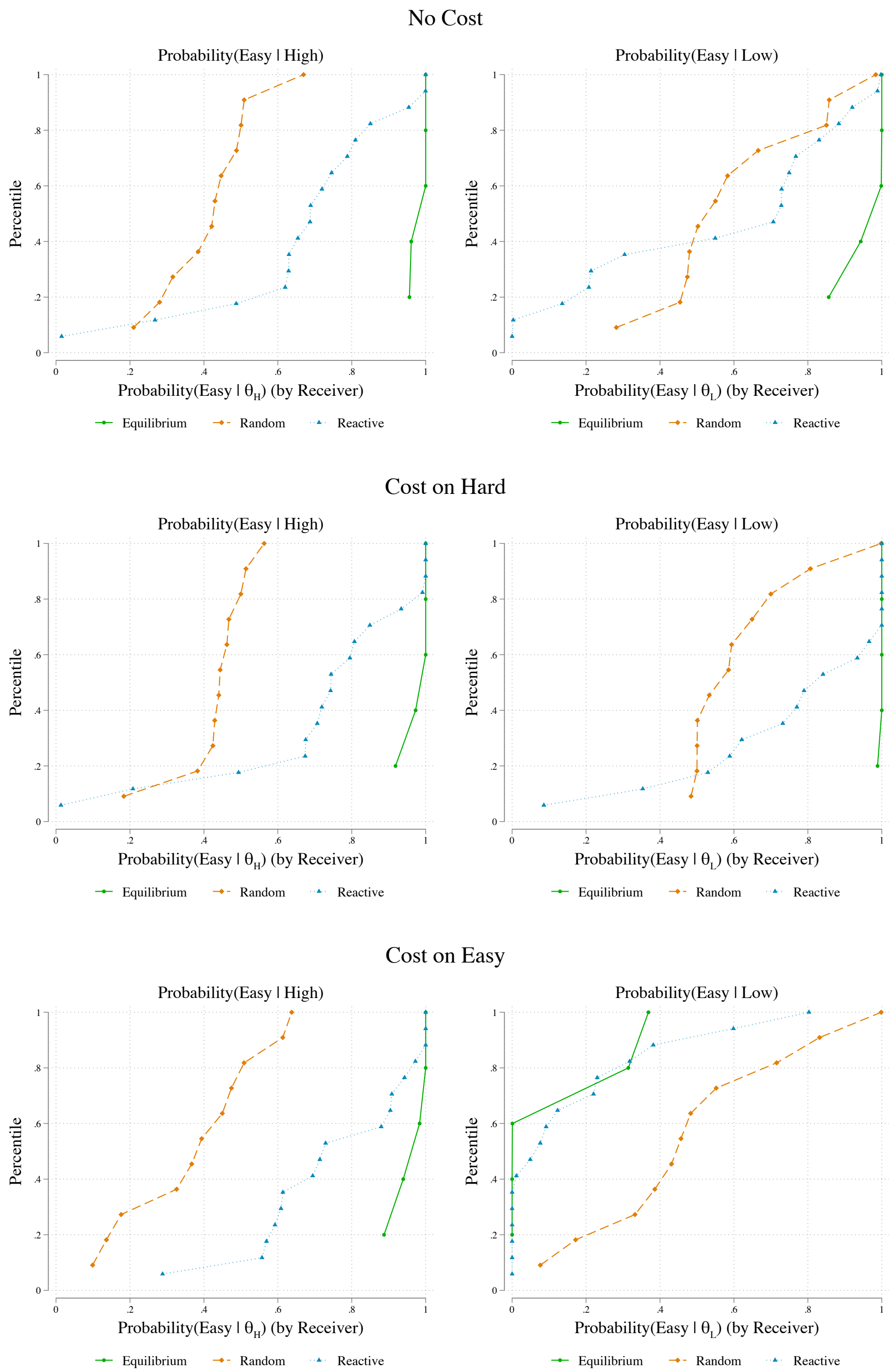


Figure 24: Distribution of receiver's estimated beliefs by cluster in the verifiable information treatment.

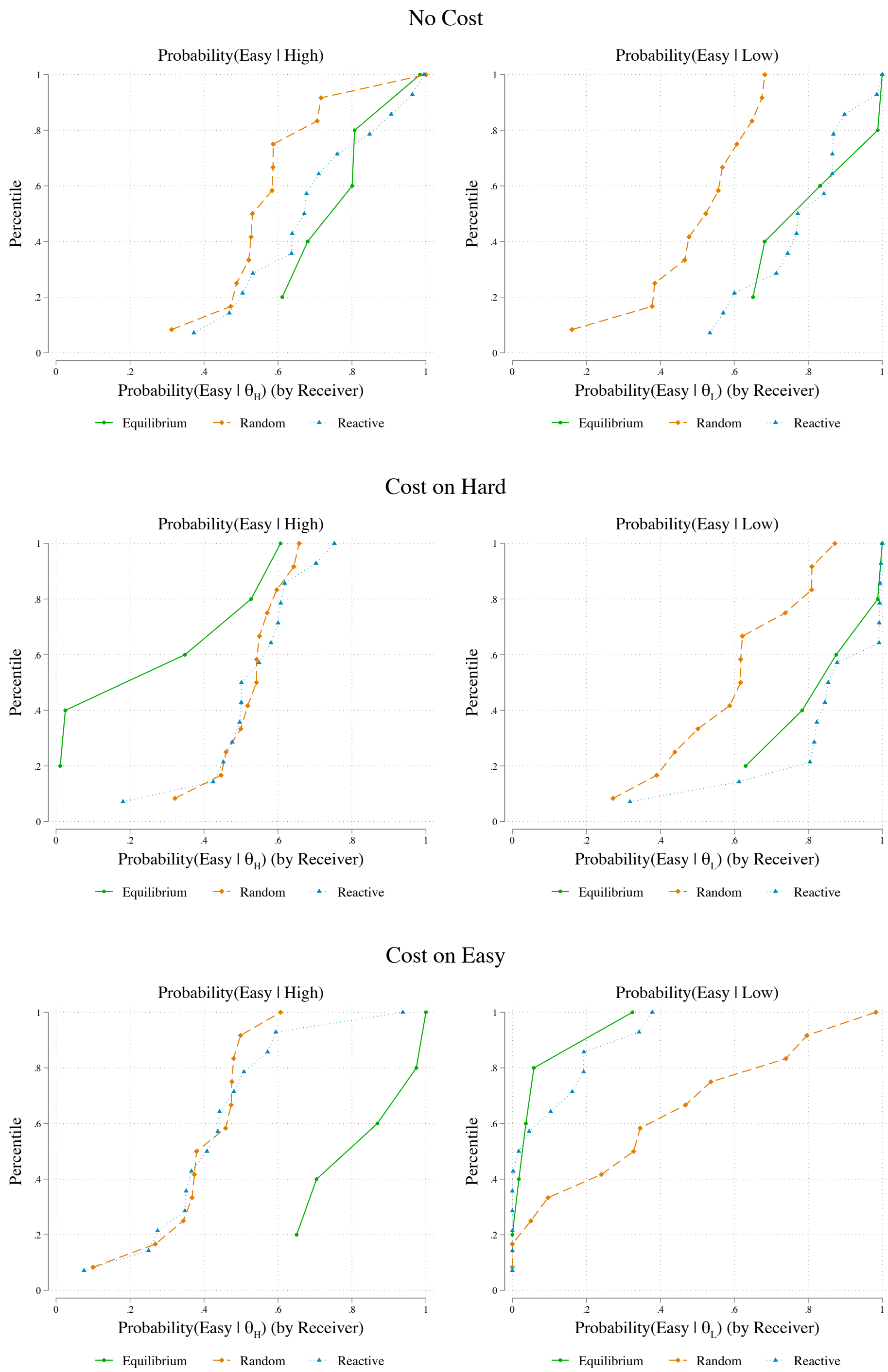
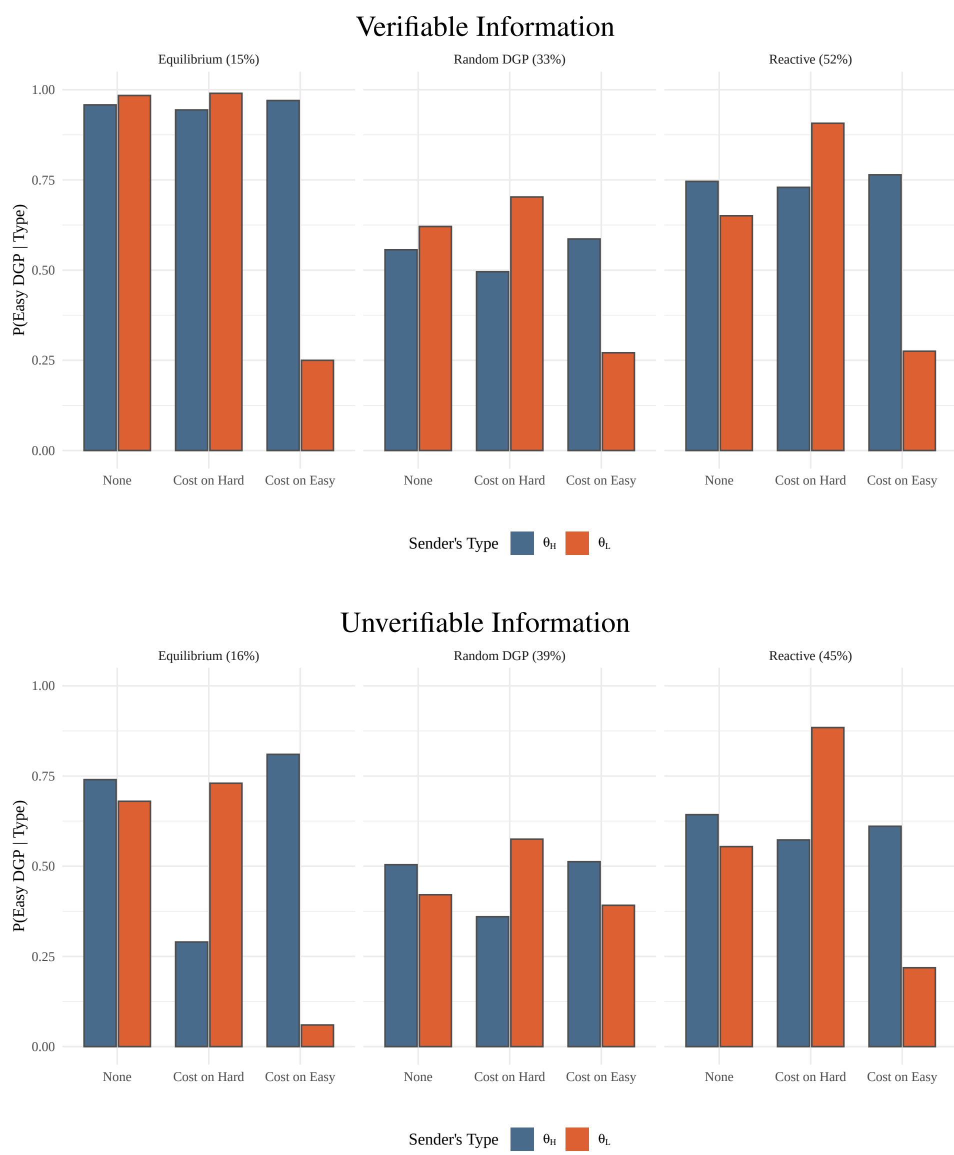


Figure 25: Distribution of receiver's estimated beliefs by cluster in the unverifiable information treatment.



Notes: The top panel reports the elicited beliefs in the verifiable information treatments and the bottom panel the elicited beliefs in the unverifiable information ones.

Figure 26: Average elicited beliefs by receiver cluster.

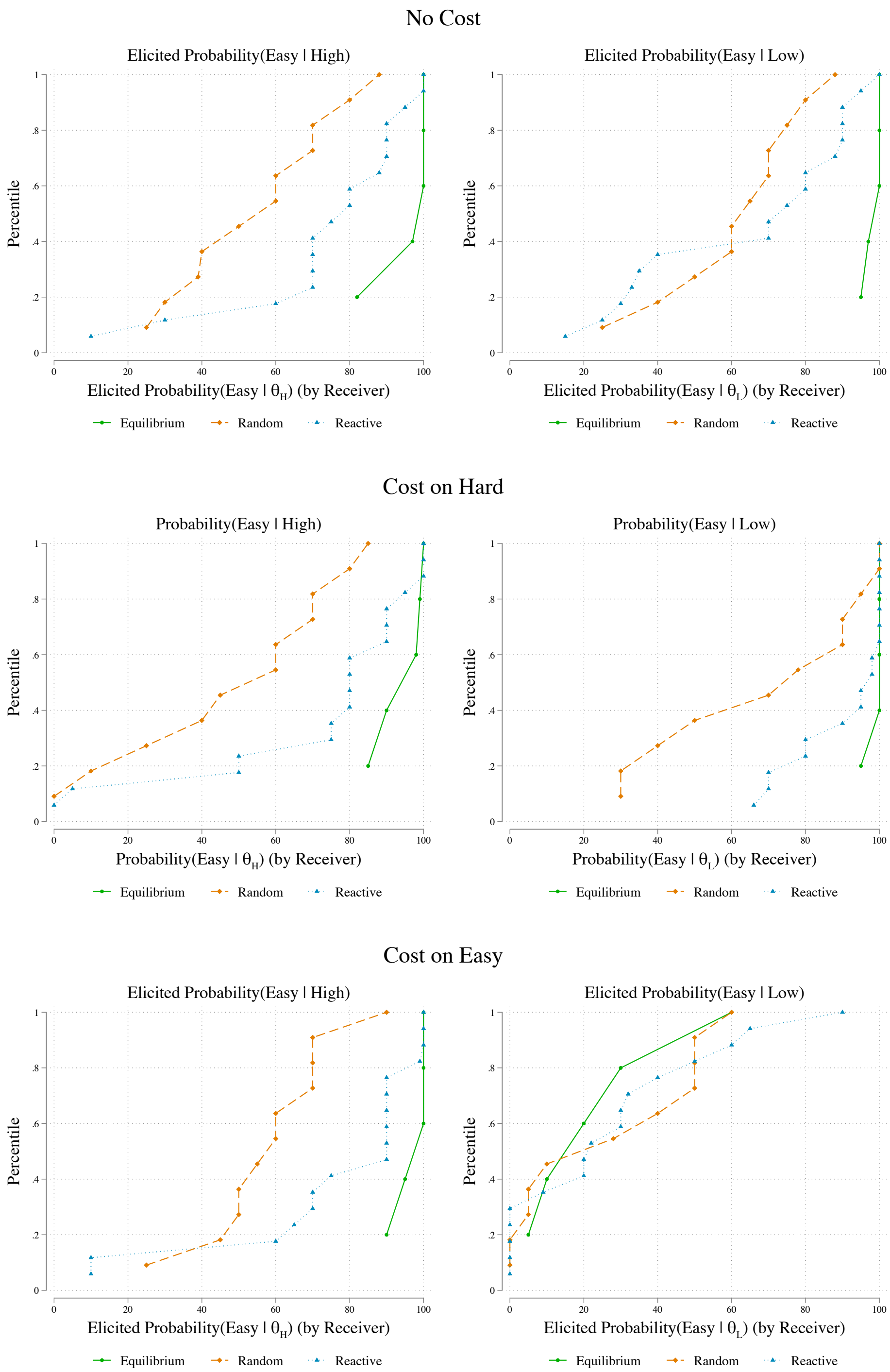


Figure 27: Distribution of receiver's elicited beliefs by cluster in the verifiable information treatment.

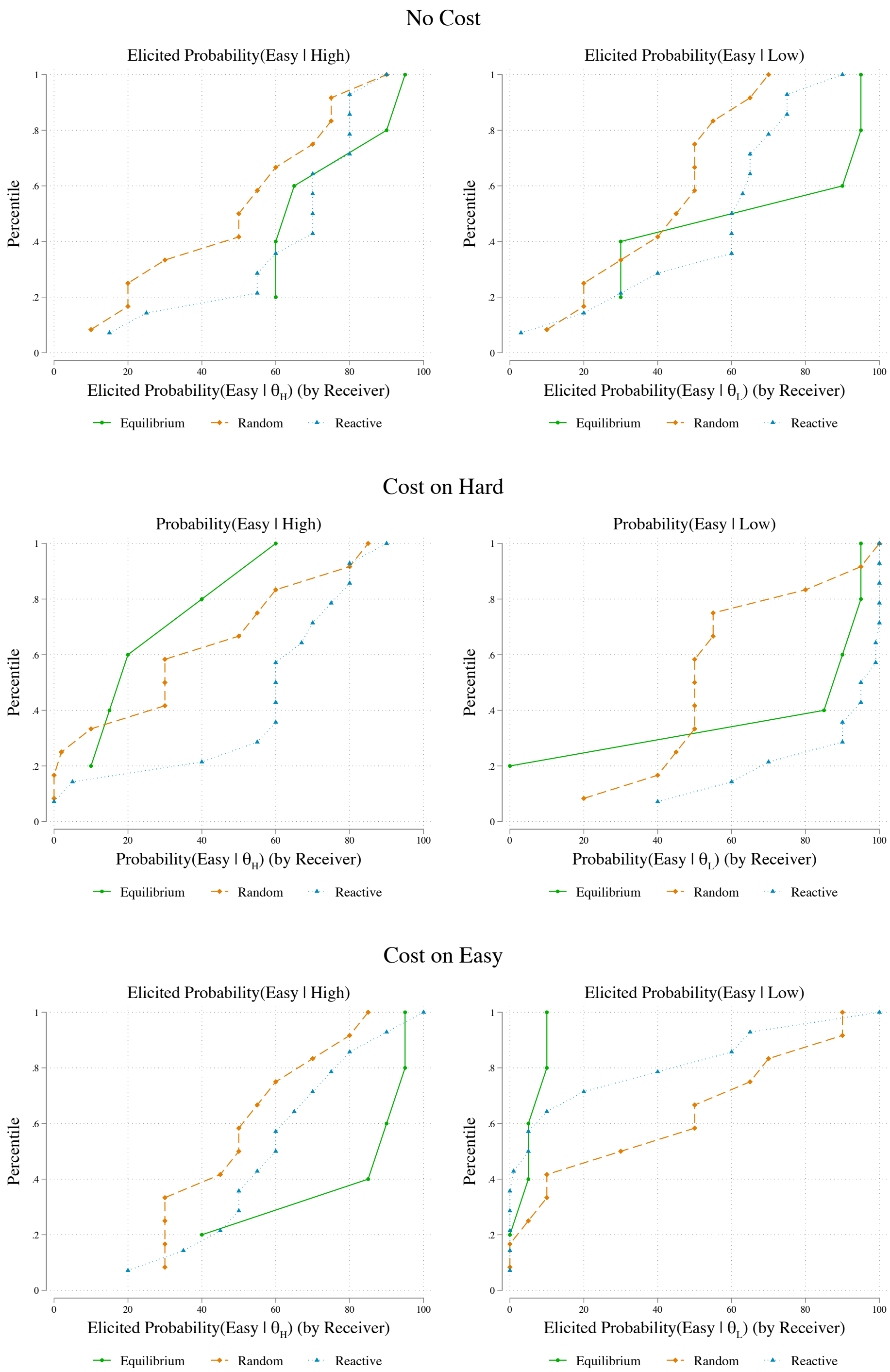
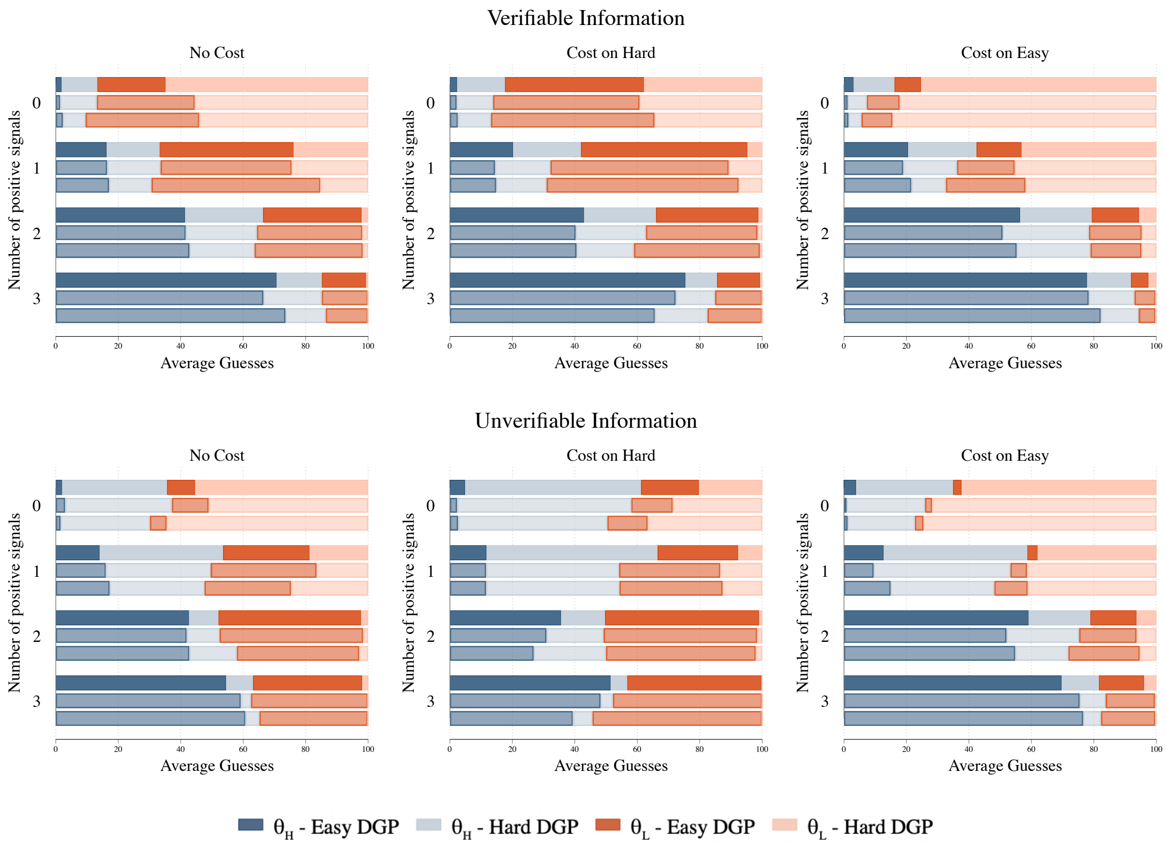


Figure 28: Distribution of receiver's elicited beliefs by cluster in the unverifiable information treatment.

E.6.3 Counterfactuals: Bayesian Guesses Given Beliefs



Notes: For each number of positive signals, the top (darker) bar represents receivers' observed behavior, the middle (lighter) bar represents the Bayesian guesses given the estimated beliefs, and the bottom (lighter) bar represents the Bayesian guesses given the elicited beliefs.

Figure 29: Average receivers' guesses given estimated and elicited beliefs.

E.6.4 Counterfactuals: Distance from the Bayesian Benchmark

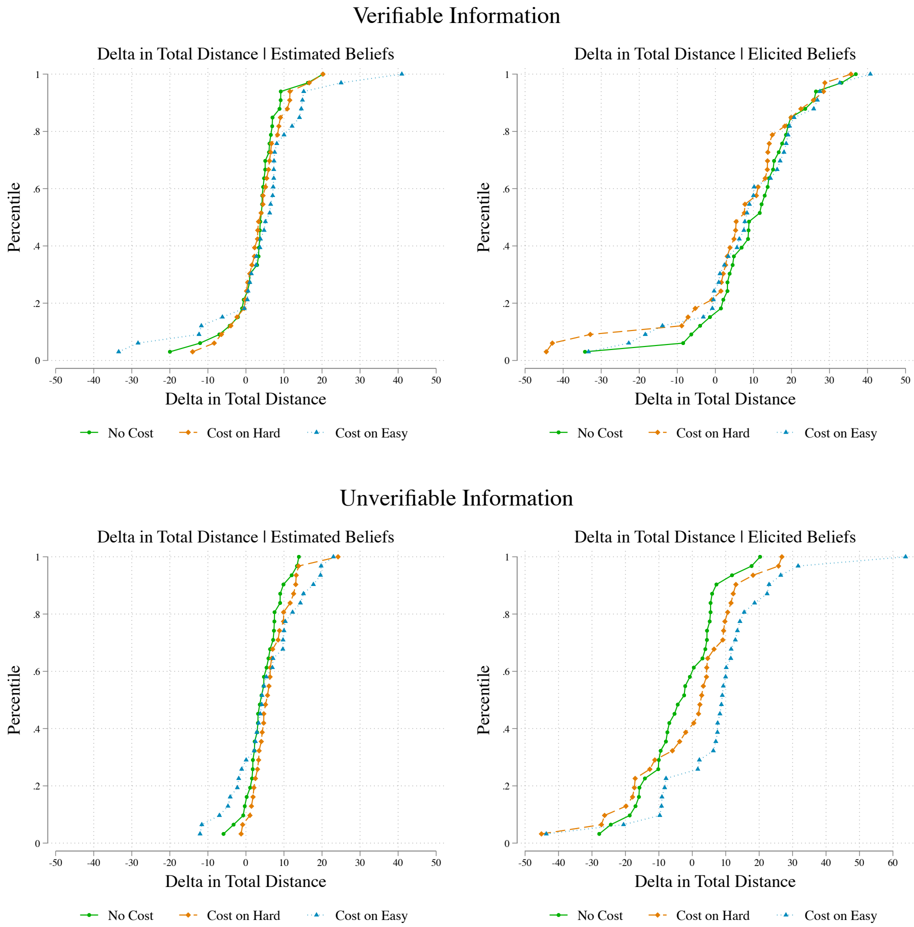


Figure 30: Distribution of difference between the total distance from the Bayesian benchmark in the data and the total distance from the Bayesian benchmark in the counterfactuals analysis (by receiver).

Figure 30 reports the subject-specific distribution of the difference between the total distance from the Bayesian benchmark in the data and the total distance from the Bayesian benchmark in our counterfactuals analysis. In the verifiable information treatments, this gap is positive for over 20% of the subjects in each cost parametrization. The same is true in the unverifiable information treatments when we compute our counterfactuals using the estimated beliefs. More variation is, instead, detected when we use the elicited beliefs in the unverifiable information cases. However, this result is consistent with the belief distributions reported in Appendix E.6.2, where we could clearly see the increased impact of noise in the elicitation—probably due to the higher variation in the sender's behavior throughout the sessions.

E.7 Information Transmitted and Processed

Table 10: Information transmitted, measured as the expected payoff of a Bayesian receiver.

	Verifiable Information			Unverifiable Information		
	Joint Info	Info on Type	Info on DGP	Joint Info	Info on Type	Info on DGP
C_0	81.24	85.21	94.38	72.57	75.72	92.43
	[82.22]	[82.22]	[100]	[76.12]	[76.12]	[100]
C_{Hard}	80.67	83.04	96.76	72.21	75.40	85.98
	[82.22]	[82.22]	[100]	[76.12 / 81.89]	[76.12 / 81.89]	[100 / 81.89]
C_{Easy}	85.85	89.45	89.75	85.25	88.88	92.46
	[93.86]	[93.86]	[93.86]	[93.86]	[93.86]	[93.86]

Notes: Theoretical predictions are reported in square brackets.

Table 11: Information processed, measured as the expected payoff of the receivers in our data.

	Verifiable Information			Unverifiable Information		
	Joint Info	Info on Type	Info on DGP	Joint Info	Info on Type	Info on DGP
C_0	67.92	82.45	82.48	69.69	74.48	90.39
	[81.24]	[85.21]	[94.38]	[72.57]	[75.72]	[92.43]
C_{Hard}	70.38	77.65	88.61	66.32	70.85	84.14
	[80.67]	[83.04]	[96.76]	[72.21]	[75.40]	[85.98]
C_{Easy}	77.53	85.88	84.21	74.48	79.91	89.86
	[85.85]	[89.45]	[89.75]	[85.25]	[88.88]	[92.46]

Notes: The information transmitted, measured as the expected payoff of a Bayesian receiver, is reported in the square brackets. The no information benchmark is 62.5 for the joint informativeness and 75 for the informativeness on θ and the one on the DGP.

E.7.1 Implications of Behavior in the Different Clusters

Table 12: Average receiver's distance from the Bayesian benchmark by cluster.

	Verifiable Information								
	Equilibrium (15%)			Random DGP (33%)			Reactive (52%)		
	Total	Type	DGP	Total	Type	DGP	Total	Type	DGP
C_0	8.89	6.89	3.53	40.58	11.66	39.07	32.33	14.43	26.76
C_{Hard}	7.26	5.77	2.66	40.33	14.33	37.72	24.15	15.11	17.12
C_{Easy}	13.70	8.97	8.09	43.81	17.93	27.26	22.86	14.55	10.84
	Unverifiable Information								
	Equilibrium (16%)			Random DGP (39%)			Reactive (45%)		
	Total	Type	DGP	Total	Type	DGP	Total	Type	DGP
C_0	14.69	8.44	11.66	20.11	8.44	17.92	13.14	8.01	8.53
C_{Hard}	20.07	16.89	7.02	24.80	18.26	12.07	17.92	12.67	10.39
C_{Easy}	14.35	11.02	8.16	42.31	28.94	15.49	27.28	19.09	12.66

Notes: The table reports the average total distance, the average type distance, and the average DGP distance from the Bayesian benchmark in each receiver's group.

Table 13: Simulations of information processed by each receiver cluster (100,000 rounds).

	Verifiable Information								
	Joint Info			Info on Type			Info on DGP		
	Equilibrium	Random DGP	Reactive	Equilibrium	Random DGP	Reactive	Equilibrium	Random DGP	Reactive
C_0	76.19	64.45	65.16	81.89	80.36	78.94	92.32	77.73	79.50
		[78.72]			*[82.74]*			*[94.55]*	
C_{Hard}	79.28	66.62	71.16	81.51	78.69	77.88	96.40	81.01	89.23
		[80.02]			*[82.26]*			*[96.58]*	
C_{Easy}	86.61	70.44	82.33	90.51	84.95	87.37	90.75	82.57	89.07
		[88.31]			*[91.45]*			*[91.91]*	
	Unverifiable Information								
	Joint Info			Info on Type			Info on DGP		
	Equilibrium	Random DGP	Reactive	Equilibrium	Random DGP	Reactive	Equilibrium	Random DGP	Reactive
C_0	68.95	67.14	68.83	74.56	74.28	74.36	87.86	85.34	88.57
		[71.33]			*[75.77]*			*[90.80]*	
C_{Hard}	67.70	64.85	68.27	71.77	69.85	73.16	84.80	83.05	84.01
		[71.93]			*[75.79]*			*[86.07]*	
C_{Easy}	82.19	68.34	77.24	85.77	76.69	82.47	90.60	87.65	88.75
		[84.45]			*[87.78]*			*[91.98]*	

Notes: The simulated information transmitted by the sender strategies, measured as the expected payoff of a Bayesian receiver, is reported in the square brackets. We simulate the information transmission for each group rather than using the measure in our data since some of the clusters are very small. This would leave us with a biased measure, also given the non-linearity in the receiver's expected payoff.

F Strategic Treatments with Observable DGP

For consistency, we label the treatments in this Section as the unobservable DGP ones, being careful in distinguishing them when any comparison between the two sets of treatments is made.

F.0.1 Sender's Choice of the Data-Generating Processes

Table 14: Sender's choice: share of Easy DGP.

	Verifiable Information		**Unverifiable Information**	
	θ_H	θ_L	θ_H	θ_L
C_0	0.64	0.79	0.38	0.57
C_{Hard}	0.44	0.89	0.25	0.79
C_{Easy}	0.69	0.25	0.68	0.20

Table 14 shows that the sender makes less extreme choices when the DGP is observable than when it is not (See Table 4). In $V - C_0$ both sender types choose the Easy DGP with a smaller probability ($p-$value < 0.1 for θ_L and $p-$value < 0.01 for θ_H), showing how the use of the Hard DGP can be consistent with a rational choice as well. When a cost is introduced on the Hard DGP, the high type chooses the Easy DGP roughly half of the time compared to before ($p-$value < 0.01). Finally, when the cost is introduced on the Easy DGP, the low type chooses the costly action more often (but this difference is not significant), while the high type chooses the Easy DGP less often ($p-$value < 0.01). Again, this pattern can be consistent with the existence of more equilibria, which makes the DGP selection harder for the sender, even when signaling opportunities are available.

We have similar patterns for the unverifiable information case. As for the other information environment, in $U - C_0$ both sender types choose the Easy DGP with a smaller probability ($p-$value < 0.01). When a cost is introduced on the Hard DGP, both sender types choose the Easy DGP less often ($p-$value < 0.01 for θ_L and $p-$value < 0.1 for θ_H), consistently with an attempt to use the cost to signal their type. Finally, when the cost is imposed on the Easy DGP, we detect a decrease in the use of the Easy DGP from the high type ($p-$value < 0.1) and a (non-significant) increase in the choice of the costly DGP for the low type.

Overall, this behavior shows how senders adapt to the change in the game structure, sometimes understanding the possible signaling value of the DGP choice when there is a positive DGP cost. However, the observability of the DGP can substantially affect the receivers' beliefs and induce multiple equilibria in a way that makes it harder for the sender to identify the correct action to take. Indeed, we find that subjects best respond less often than in the unobservable DGP treatment to the receivers' data (roughly 70% of the observations).

Despite the behavior being less extreme, we still find, as in the main treatment, that senders adapt to the different strategic incentives induced in the different treatments, allowing for some variation in the meaning of the data. This makes us able to test the effect on the receivers' guesses, this time accounting for the observability of the DGP.

F.1 Receiver's Guesses

In these treatments, plotting the receivers' guesses and comparing them with the Bayesian benchmark would require twice as many figures as before, since it is now important to condition on the DGP choice. In addition, some events—i.e., some specific combinations of DGP and number of signals—happen very rarely, mainly in the unverifiable information environment, making the numbers we represent quite noisy. For this reason, we study the consistency with equilibrium patterns aggregating over signal realizations and comparing the guesses made for a particular (θ, DGP) combination across cost parametrizations. In doing so, we can control for the composition of signals subjects saw when making their guesses, and extract the average changes due to variations in the treatment.

Starting with the verifiable information treatments, we find that, on average, after observing the Easy DGP, subjects believe the high type is more likely in $V - C_{Easy}$ than in $V - C_0$ ($p-$value < 0.1)—accounting for the effect of the cost. However, the magnitude of this change is much smaller than what the Bayesian guesses would predict. Similarly, while comparing the same treatments, subjects should place less weight on the type being high in $V - C_{Easy}$ when the Hard DGP is used; we detect an extremely low change that lacks significance. We can then compare the guesses made in $V - C_{Hard}$ and $V - C_0$. Given the cost of the Hard DGP for the low type, we should detect an increase in the inference for the high type when the Hard DGP is used. Such an increase is detected in the data, with a magnitude very close to the theoretical one ($p-$value < 0.05). Similarly, comparing the same treatments, we should observe lower weights attached to the type being high in $V - C_{Hard}$ when the Easy DGP is used. While the magnitude of this change resembles the predicted one, it lacks significance ($p-$value = 0.121). Finally, we can compare the guesses made in the treatments with positive costs. When a cost is imposed on the Hard DGP, subjects should be, on average, more optimistic about the sender's type being high after seeing the Hard DGP. Data confirms this pattern, but the magnitude of the change is below the predicted one ($p-$value < 0.1). In contrast, when a cost is imposed on the Easy DGP, subjects should be, on average, more optimistic about the sender's type after seeing the Easy DGP. Again, this is confirmed in the data, but the magnitude of the change is smaller than the Bayesian benchmark ($p-$value < 0.01).

We can perform the same type of comparisons in the unverifiable information treatments. On average, subjects believe the high type is more likely in $U - C_{Easy}$ than in $U - C_0$ when the Easy DGP is used ($p-$value < 0.01), but the average change is smaller in magnitude than the predicted one. Similarly, while comparing the same treatments, subjects should place less weight on the type being high in $U - C_{Easy}$ when the Hard DGP is used. While such a change is detected ($p-$value < 0.1), the magnitude is much smaller than the Bayesian predictions. We can then compare the guesses made in $U - C_{Hard}$ and $U - C_0$. Given the cost of the Hard DGP for the low type, we should observe a significant increase in the inference of the high type when the Hard DGP is used. That is indeed the case, with magnitude even exceeding the Bayesian one ($p-$value < 0.01). Similarly, we detect a lower weight attached to the type being high when the Easy DGP is used, but again with smaller magnitudes compared to the Bayesian benchmark ($p-$value < 0.05). Finally, as we did before, we can compare the two treatments with positive costs. When a cost is imposed on the Hard DGP, subjects should attach a higher probability to the sender's type being high after seeing that the Hard DGP was indeed chosen. Data confirms this pattern ($p-$value < 0.01), and the magnitude of the change is not very far from the predicted one. In contrast, when a cost is imposed on the Easy DGP, subjects should be, on average, more optimistic about the sender's type if the Easy DGP is chosen. Again, data reveal a pattern consistent with this change, but with overreaction compared to the Bayesian benchmark ($p-$value < 0.01).

Overall, these results show that subjects react, on average, to the changes in the environment that trigger changes in the senders' strategy. Despite this positive result, we also find some systematic deviations in the way receivers make their guesses. First, we find that subjects tend to overweight, on average, the probability of the type being high after observing the choice of the Easy DGP when no cost is imposed and when the cost is imposed on the Hard DGP. When the low type needs to pay a cost to choose the Hard DGP, the Easy DGP choice is a possible signal that the type is low and is not willing to pay the signaling cost. The fact that receivers tend to put too much weight on θ_H suggests a partial failure to make this reasoning. We find significant positive gaps between the average guess on the high type and the average optimal one in $V - C_0$ ($p-$value < 0.01), $U - C_0$ $p-$value < 0.1, $V - C_{Hard}$ ($p-$value < 0.05) and $U - C_{Hard}$ ($p-$value < 0.01). Consistent with this argument, we find that when the Hard DGP is chosen, subjects underweight the likelihood of a high type in three of these treatments in which the choice would potentially have a signaling value: $V - C_0$ ($p-$value < 0.01), $V - C_{Hard}$ ($p-$value < 0.01), and $U - C_0$ ($p-$value < 0.05). In $U - C_{Hard}$ we find no significant gap. Smaller systematic deviations are, instead, detected in the treatments in which a cost on the Easy DGP is imposed. Here, we find that subjects tend to place too much weight on the high type when the Hard DGP is chosen. While the magnitudes are comparable (roughly 8 percentage points), the gap is only significant for $V - C_{Easy}$ ($p-$value < 0.1).

This discussion suggests that, similarly to what we found in the unobservable DGP treatments, while receivers respond to changes in the information environment and in the incentives, they tend to make mistakes that are consistent with an imperfect assessment of the sender's strategy, mostly when such a strategy involves type-dependent choices.

F.1.1 Receivers' Elicited Beliefs

As we did for the main strategic treatments, at the end of the session, we elicited receivers' beliefs.[70] We can compare the beliefs elicited in these new treatments (reported in Figure 31), and in the main ones. We find that, despite changes in the sender's behavior, there are very small differences in the elicited beliefs. This suggests that the possible biases in accounting for DGP selection lead to similar aggregate biases in the assessment of the possible DGPs.[71] Hence, while at a first glance beliefs seem more accurate in the treatments with observable DGP, this may be merely the result of a less extreme sender's behavior, rather than a more accurate receiver's assessment. Since it is not possible to separate these forces, we try to study whether some meaningful heterogeneity exists in the way receivers behave.

[70] Our structural approach cannot be used to estimate beliefs in these new treatments. In this case, subjects only guess the sender's type, and not the chosen action, making it impossible to separately identify the probability they attach to a certain action being taken by each type.

[71] Plotting the distributions of such beliefs, we find again that they are extremely similar across treatments.

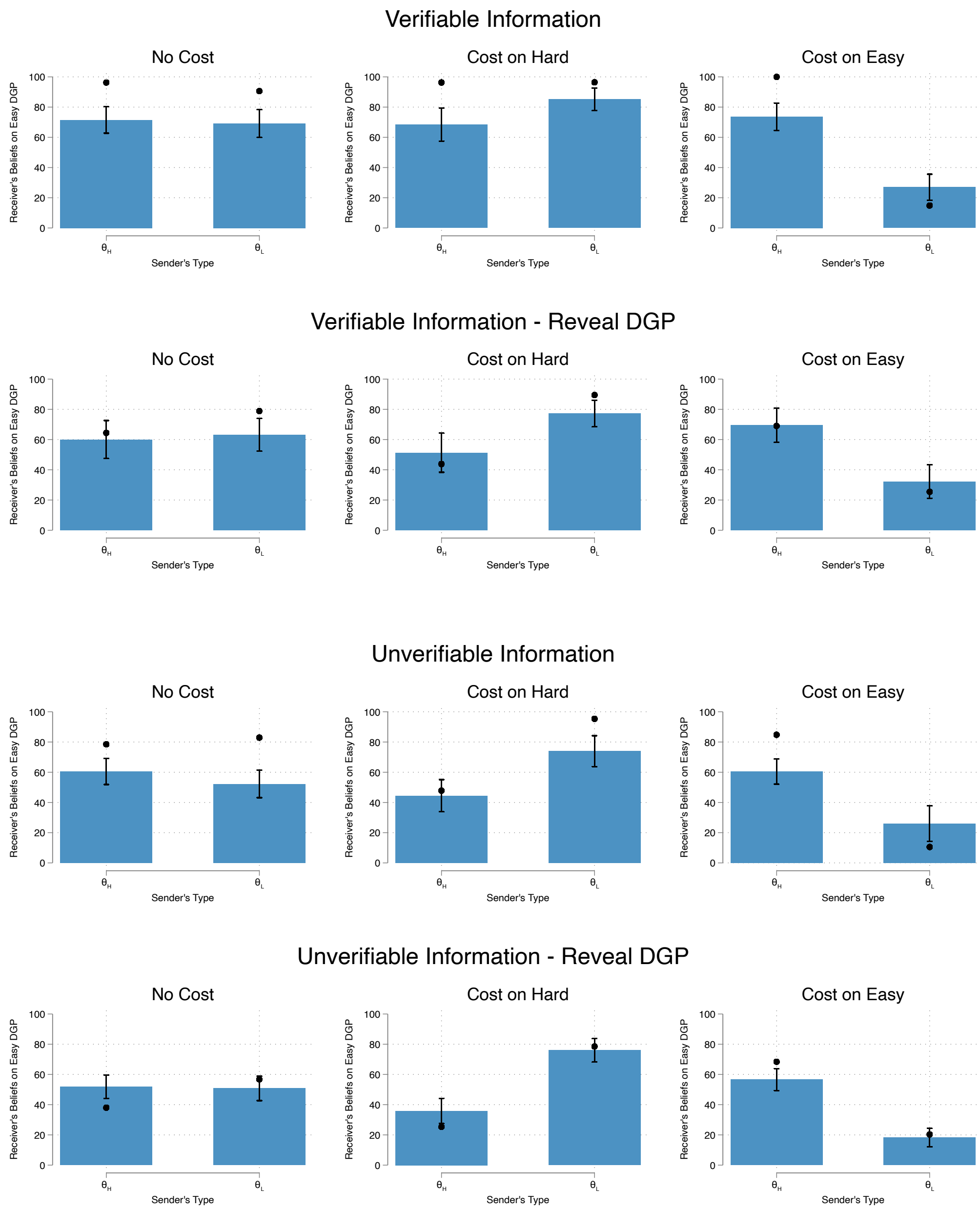


Figure 31: Receivers' elicited beliefs.

Notes: The dots represent the sender's strategy.

F.1.2 Discussion of Receivers' Heterogeneity

As we did for the main treatments, we try to classify subjects from their behavior, comparing their guesses to two benchmarks: the Bayesian benchmark and the random DGP one. Differently from before, subjects only provide us with a guess about the sender's type, preventing us from using our total distance measure. For this reason, we perform the same classification discussed in Section 4.2.3 using the type distance instead.

While our procedure identifies some heterogeneity across participants, the results are much less clean than in our main treatments. We detect the existence of two groups of receivers in each information environment. The first (roughly 45% of the receivers) exhibits behavior closer to equilibrium. However, given the almost random behavior of the sender in the no cost treatments, this mainly captures an appropriate reaction to the costs, making it hard to separate between an equilibrium and a reactive group as we did in the main strategic game. The second group presents a more inaccurate behavior. However, while we can explain the deviations with possible random DGP selection in the unverifiable information environment, this explanation fits the data less in the verifiable information treatments—where the elicited beliefs of the two groups tend to be more similar.

The classification exercise is much more challenging in this treatment for two reasons. First, the fact that receivers only report beliefs about the type allows for much less variation in their guesses, making it harder to detect relevant patterns. Second, the sender's behavior is less selective, making it harder to distinguish between optimal responses to an intermediate behavior and, instead, neglect of the sender's strategy. However, our results suggest that again we can detect heterogeneity in the way in which subjects play the game and, mostly in the unverifiable information treatments, differences in beliefs are partially predictive of behavior and of the accuracy of receivers' guesses.

G Robustness: All Rounds

G.1 Senders' Results

G.1.1 Choice of the Data-Generating Processes

Table 15: Sender's DGP choice: share of Easy DGP (all rounds).

	Verifiable Information		**Unverifiable Information**	
	θ_H	θ_L	θ_H	θ_L
C_0	0.95	0.84	0.80	0.74
	[1]	[1]	[1]	[1]
C_{Hard}	0.93	0.96	0.53	0.97
	[1]	[1]	[0 / 1]	[1]
C_{Easy}	0.92	0.16	0.84	0.13
	[1]	[0]	[1]	[0]

Notes: Theoretical predictions in square brackets.

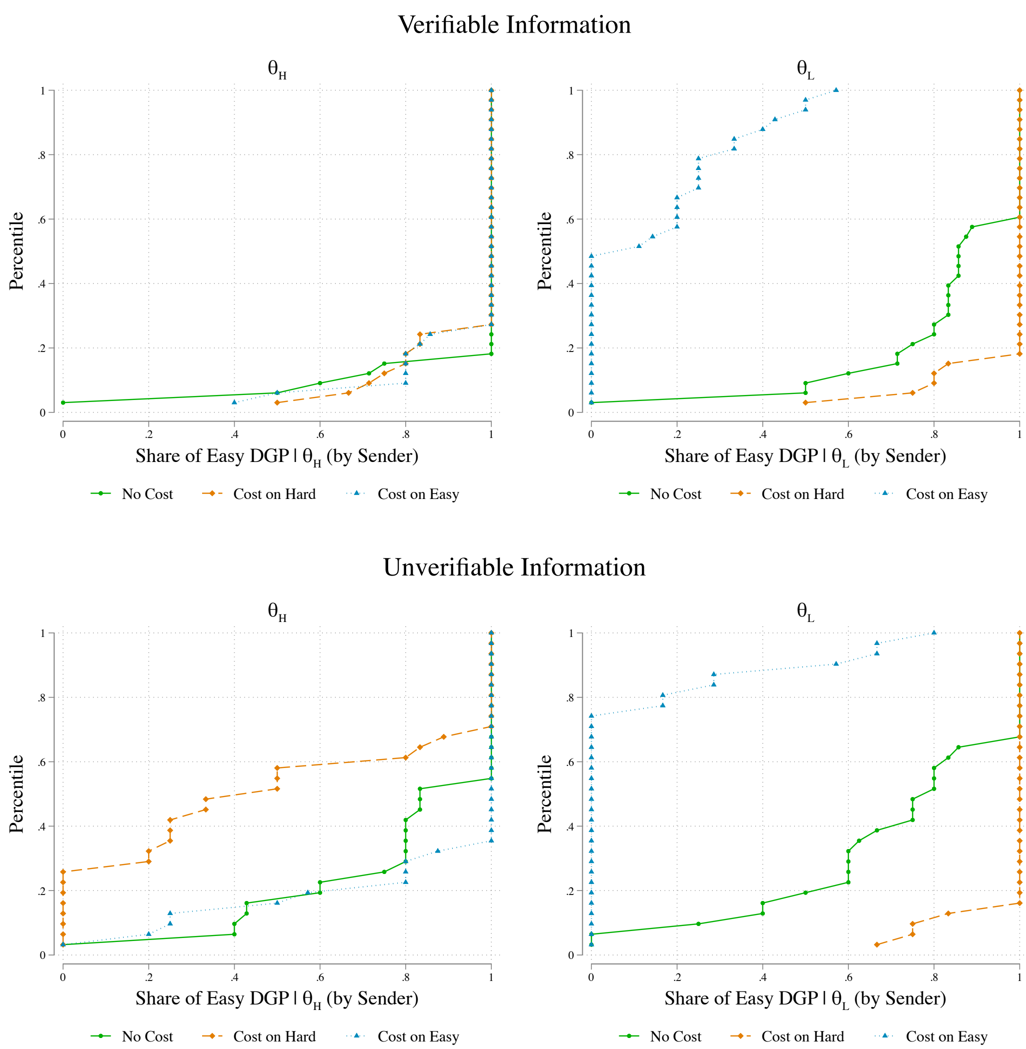
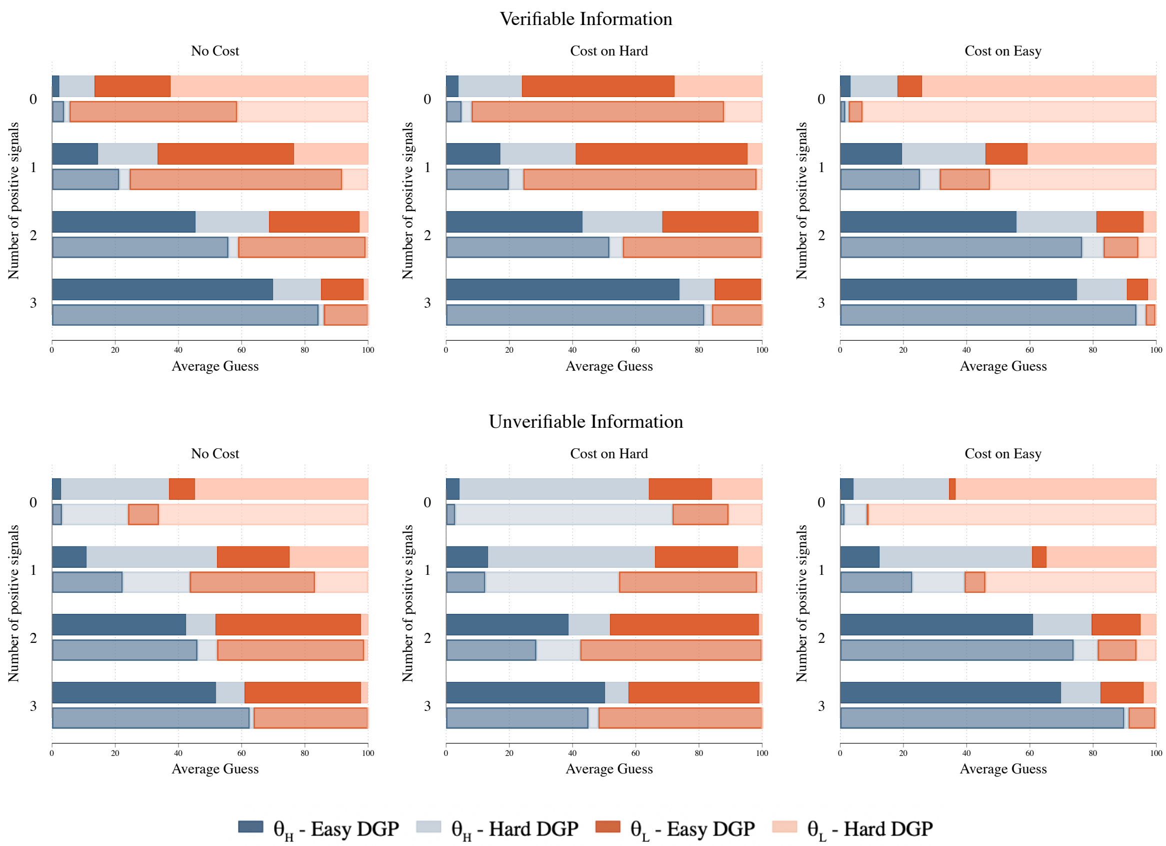


Figure 32: Distribution of sender's DGP choices (all rounds).

G.2 Receivers' Results

G.2.1 Receivers' Inference



Notes: For each number of positive signals, the top (darker) bar represents receivers' observed behavior, while the bottom (lighter) bar represents the optimal guesses implied by the senders' behavior.

Figure 33: Average receiver guesses (all rounds).

G.2.2 Deviations from Bayesian and Random Benchmark

Table 16: Average distance implied by the receivers' behavior, both relative to the Bayesian benchmark and to the random DGP selection benchmark (all rounds).

	Verifiable Information		**Unverifiable Information**	
	Distance from Bayesian	Distance from Random DGP	Distance from Bayesian	Distance from Random DGP
C_0	28.72 [0]	21.74 [33.42]	17.47 [0]	15.73 [13.60]
C_{Hard}	25.66 [0]	25.96 [37.18]	19.06 [0]	21.70 [22.26]
C_{Easy}	27.12 [0]	26.23 [32.87]	29.22 [0]	25.13 [33.21]

Notes: The distance implied by the behavior of a Bayesian receiver is reported in square brackets.

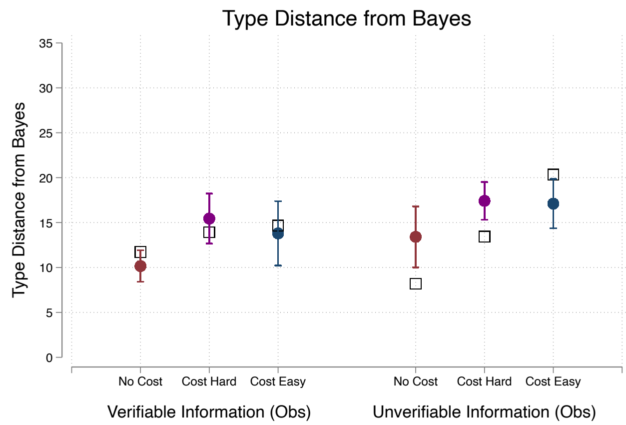
G.3 Observable DGP

G.3.1 Senders' Results

Table 17: Senders' DGP choice: share of Easy DGP (all rounds).

	Verifiable Information		**Unverifiable Information**	
	θ_H	θ_L	θ_H	θ_L
C_0	0.62	0.66	0.45	0.59
C_{Hard}	0.48	0.92	0.32	0.83
C_{Easy}	0.73	0.28	0.70	0.23

G.3.2 Inference and Information Transmission



Notes: The hollow squares report the average type distance from the Bayesian benchmark for the corresponding strategic treatment with unobservable DGP.

Figure 34: Average type distances from the Bayesian benchmark for the observable DGP treatments (all rounds).

Table 18: Information transmitted and processed about the type (all rounds).

	Verifiable Information		**Unverifiable Information**	
	Unobservable DGP	Observable DGP	Unobservable DGP	Observable DGP
C_0	81.09	81.20	75.68	76.44
	[84.43]	[84.73]	[76.69]	[80.14]
C_{Hard}	77.56	81.44	72.19	76.93
	[82.48]	[84.53]	[75.43]	[83.28]
C_{Easy}	84.29	84.13	78.97	77.34
	[88.77]	[88.48]	[87.29]	[81.95]

Notes: The information transmitted, measured as the expected payoff of a Bayesian receiver, is reported in the square brackets.

H Additional Elicitations

We report the details of the additional elicitation tasks performed at the end of our experiment.

Beliefs (strategic treatment only). We elicit cross-role beliefs from both roles. Receivers report, for each cost configuration, their beliefs about each sender type's DGP choice. These reports are incentivized with a binarized scoring rule (BSR) using a randomly matched sender. Senders report, for each cost configuration and for each possible signal realization (0–3 positive signals)—and DGP choice in the strategic treatment with observable DGP—, their beliefs about the receiver's belief that the sender's type is θ_H. These reports earn a 100-point bonus if the matched receiver's guess lies within 10 percentage points of the sender's stated belief.

Preferred signal realizations (main strategic treatment only). We elicit preferences over the realization of signals for the senders. Senders select their preferred realization (0–3 positive signals); payment is determined by matching to a round with that realization under the corresponding cost.

Preferred information (all treatments with unobservable DGP). We elicit preferences over information. Receivers, within a given setting (cost or selection rule), complete an additional task that mirrors the main task: the selected DGP is that of the matched sender or machine from the corresponding main round. Before submitting their belief report, they may choose one of two information options: (i) observe nine additional signals drawn at random from the same DGP as before, or (ii) learn the identity of the DGP (DGP A or DGP B). After observing the chosen information, they complete the task under the same incentives as in the main task.

Risk aversion and altruism. At the end of each session, risk aversion is elicited with two investment tasks following Gneezy and Potters (1997). In strategic treatments, altruism is elicited with two dictator tasks.[72]

[72]In the investment tasks, participants receive 100 points to allocate between a safe investment (returning one point per point invested) and a risky investment that, in Task 1, returns 2.5 times the invested amount with probability 50%, and in Task 2, returns 3 times the invested amount with probability 40%. In the dictator tasks, participants received 100 points to allocate between themselves and another participant; in the first elicitation, each transferred point yielded one point to the recipient, while in the second, each transferred point was doubled for the recipient. Participants were paid for one randomly selected investment task, one randomly selected dictator task, and one randomly selected recipient task.

I Interface and Additional Elicitations

(a) Verifiable Information

(b) Unverifiable Information

Figure 35: Information environments: DGP compositions.

I.1 Interface Strategic Treatments

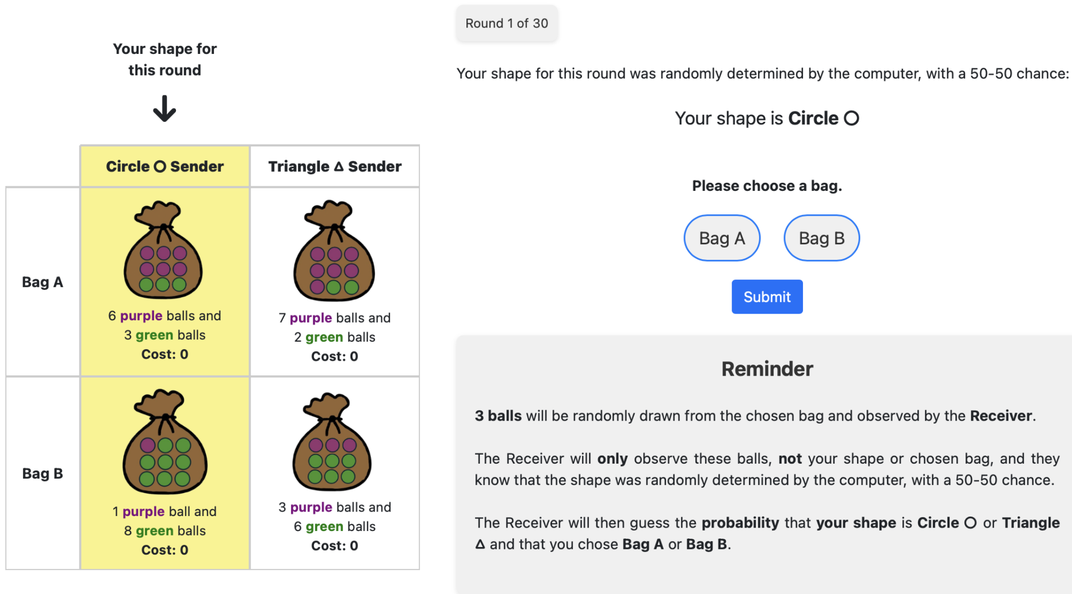


Figure 36: Sender's interface (main strategic treatments)

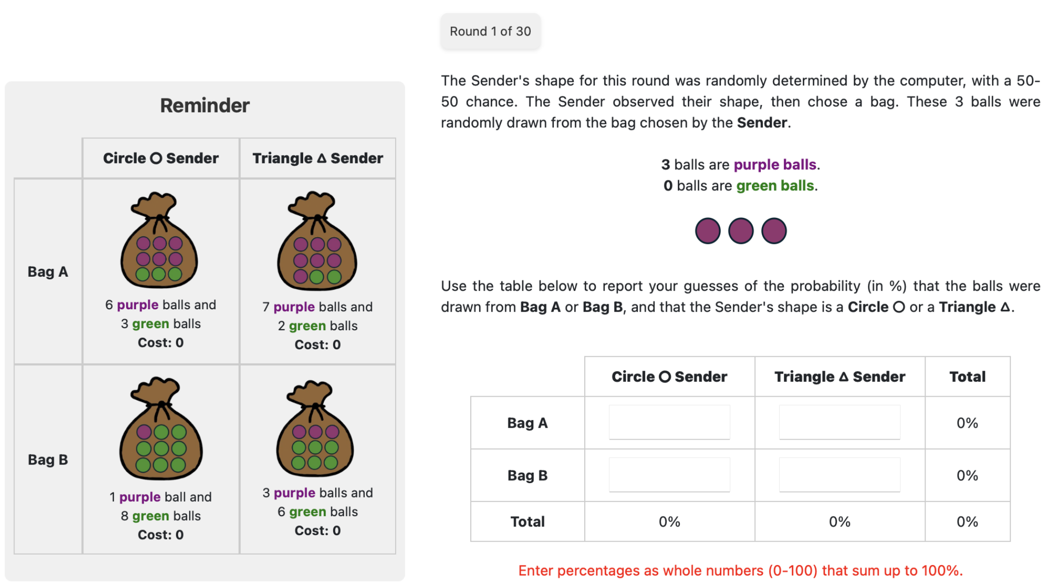


Figure 37: Receiver's interface (main strategic treatments).

I.2 Interface Exogenous Treatments

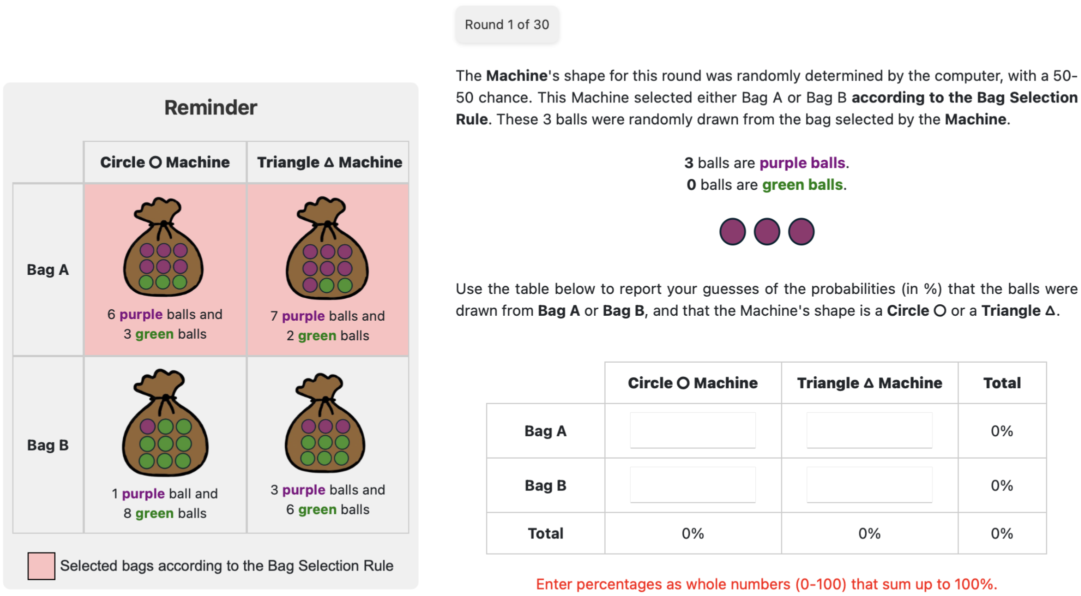


Figure 38: Sender's interface (deterministic treatments)

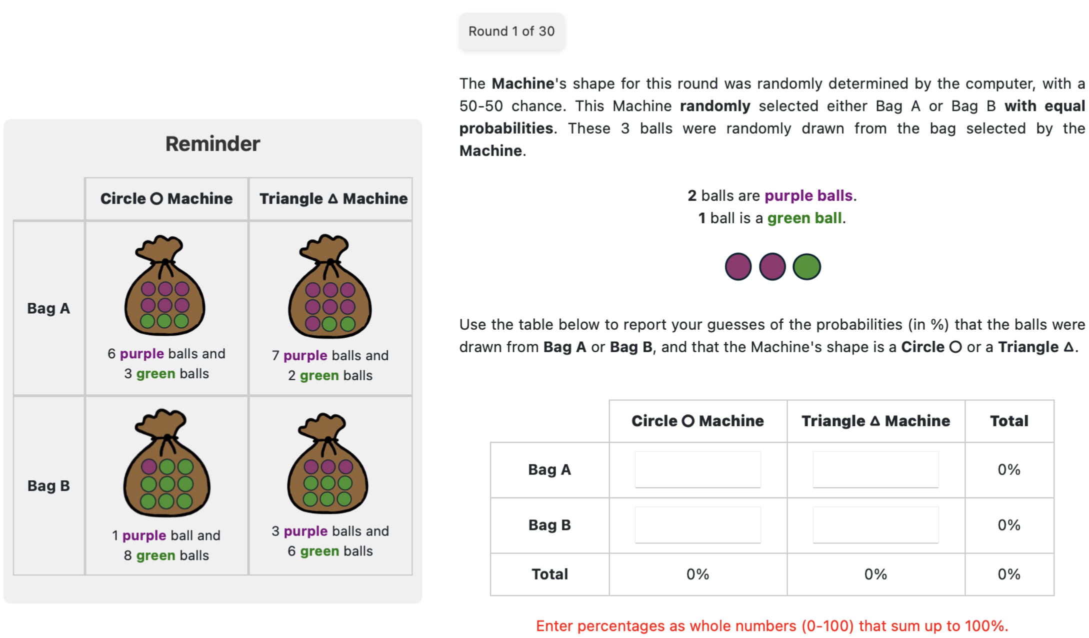


Figure 39: Sender's interface (random treatments).

I.3 Interface Strategic Observable Treatments

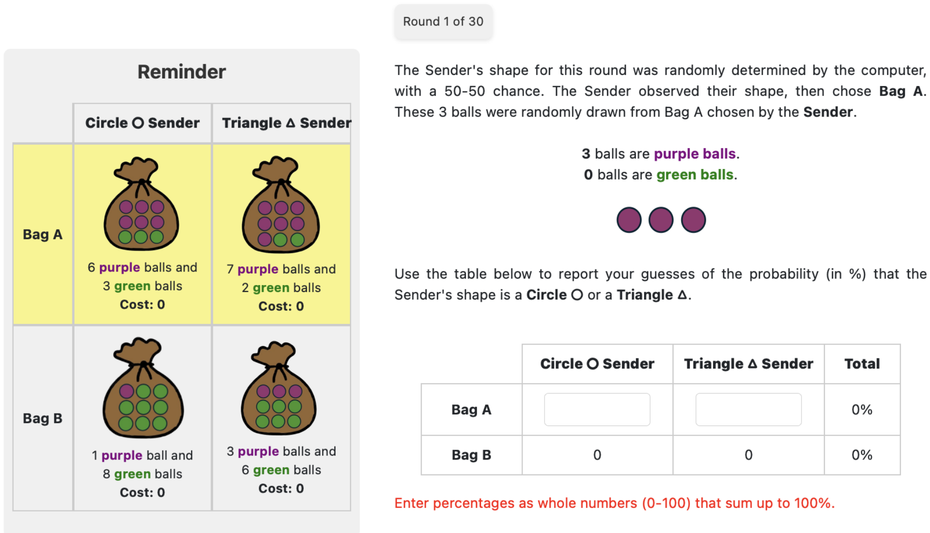


Figure 40: Receiver's interface (observable DGP treatments).

J Sample Instructions: Main Strategic Treatment

General Information about the Game

Today's game is played in groups of two players, a Sender and a Receiver. You will play for 30 rounds in one of two roles: either Sender or Receiver. At the end of each round, you will be randomly paired with a new participant, but your role will remain fixed throughout the session.

At the beginning of each round, the Computer will randomly assign each Sender a shape, either Circle ◯ or Triangle △, with equal probabilities. This means that, in each round, the Sender has a 50% chance of being a Circle ◯ Sender and a 50% chance of being a Triangle △ Sender. **The Sender learns their shape, while the Receiver does not**.

Each Sender has access to two bags, **Bag A** and **Bag B**, each containing **nine balls**. The balls can be either **purple** or **green**. The exact composition of each bag depends on the Sender's shape.

If the Sender is a **Circle ◯ Sender**:

- **Bag A** contains **6 purple balls** and **3 green balls**;
- **Bag B** contains **1 purple ball** and **8 green balls**.

If the Sender is a **Triangle △ Sender**:

- **Bag A** contains **7 purple balls** and **2 green balls**;
- **Bag B** contains **3 purple balls** and **6 green balls**.

Player's Tasks

Sender

After learning their shape, the **Sender chooses one of the two available bags, Bag A or Bag B**. The Sender may incur a cost for choosing either bag. This cost may depend on the Sender's shape. Such costs will vary throughout the session, and all players (Senders and Receivers) will know these costs and will be informed whenever they change.

After the Sender chooses the bag, the Computer will randomly draw **three balls** from the chosen bag and show them to the Receiver. After each draw, the ball is placed back into the bag, ensuring that the total number of balls and their composition remains unchanged across all three draws. As a result, **the chance of drawing a purple or a green ball is the same across all three draws**.

Receiver

The Receiver observes the three balls but is uninformed about the Sender's shape and the chosen bag. Given the available information, **the Receiver will assess how likely it is that the Sender's shape is Circle ◯ or Triangle △ and that the chosen bag is either Bag A or Bag B**.

Specifically, the Receiver must assign a probability to each combination of the Sender's shape and chosen bag. To complete their task, the Receiver will enter their guesses of the probability of each shape and bag combination as a whole number between 0 and 100, representing a percentage. The four percentages in the table must sum to 100%. Given the Receiver's guesses, the software will automatically provide the sum of each column, which indicates the total probability assigned to the Sender's shape being Circle ◯ or Triangle △. Similarly, the software will automatically provide the sum of each row, which indicates the total probability assigned to the chosen bag being Bag A or Bag B.

Earnings

You will earn points in each round which will be converted into cash at the end of the experiment.

Sender

The number of points the Sender earns in a round depends on the Receiver's guesses and on the cost of the chosen bag. Specifically, the number of points earned in a round is equal to **the Receiver's guess of the probability that the Sender's shape is Triangle $\triangle$, minus the cost of the chosen bag**.

In other words, the higher the probability the Receiver assigns to the Sender's shape being Triangle $\triangle$, the higher the number of points potentially earned by the Sender. However, the cost of the chosen bag may reduce that amount. The final number of points depends on both the Receiver's beliefs and the cost associated with the bag selection. Note that the sender's payoff in a round may be negative.

Receiver

The number of points earned by the Receiver in a round depends on the **accuracy of their guesses** regarding the Sender's shape and the chosen bag. Specifically, given the actual Sender's shape and chosen bag and the Receiver's guesses, a function is used to determine a score, which will be a whole number between 0 and 100. The score is the probability (in percentage) of winning a prize of 100 points in a lottery. Indeed, after the score is computed, the Computer will randomly draw a whole number from 1 to 100, with each number being equally likely. If the score is greater than or equal to the randomly drawn number, the Receiver wins a prize of 100 points, otherwise, they win nothing. **The more accurate the Receiver's guesses are, the higher the score and the**

higher their chances of winning the prize.

If you wish, you can refer to the cheat sheet for more information on the function used to compute the score and on the compensation mechanism.

What matters is that this compensation rule is purposefully designed so that the Receiver has the greatest chance of winning the 100 points prize when their guesses equal their true belief for each combination of the Sender's shape and the chosen bag.

Round Summary

At the end of each round, both Sender and Receiver will see a screen summarizing all the important information from that round. They will be informed about the shape of the Sender (Circle ◯ or Triangle △), the bag chosen by the Sender (either Bag A or Bag B), the three balls drawn from the chosen bag, and the Receiver's guesses. Also, both participants will receive information about the payoffs in that round.

Additional Tasks

After playing 30 rounds of this game, you will be asked to complete some additional tasks. The instructions for these tasks will be provided on the screen. Please read them carefully and raise your hand if you have a question.

These tasks will give you the chance to earn more points. The number of points you may earn will be clearly stated in the instructions.

Final Payments

You will be compensated for all rounds and for the additional tasks completed at the end of the experiment. Your total number of points will be converted into dollars at the rate of $1 per 100 points ($0.01 per point). In addition, you will receive a participation fee of $10.

Practice Rounds

You will first complete two unpaid practice rounds, where you will play both the role of Sender and Receiver. These rounds aim to make you familiar with the interface. We will guide you through the first practice round, and you will complete the second one by yourself. After the practice rounds, you will be informed of your role for the subsequent rounds directly on the interface.

K Sample Cheat Sheet: Main Strategic Treatment

Cheat Sheet: Formula for Receiver Earnings

In each round of the experiment, after having observed the randomly drawn balls, the Receiver will be asked to guess the probability that the Sender's shape is Circle $\bigcirc$ or Triangle $\triangle$ and that the chosen bag is either Bag A or Bag B.

Let's denote the four guesses using the function $p(\cdot,\cdot)$, where

- $p(\bigcirc, A)$ is the guess of the probability (in %) that the Sender's shape is Circle $\bigcirc$ and the chosen bag is Bag A;
- $p(\bigcirc, B)$ is the guess of the probability (in %) that the Sender's shape is Circle $\bigcirc$ and the chosen bag is Bag B;
- $p(\triangle, A)$ is the guess of the probability (in %) that the Sender's shape is Triangle $\triangle$ and the chosen bag is Bag A;
- $p(\triangle, B)$ is the guess of the probability (in %) that the Sender's shape is Triangle $\triangle$ and the chosen bag is Bag B.

Given the actual Sender's shape, the actual chosen bag, and the Receiver's guesses in the round, the Receiver's score S is defined as follows.

- If the Sender's shape is is Circle $\bigcirc$ and the chosen bag is Bag A
$$S = 100 - 50 \cdot \left[\left(1 - \frac{p(\bigcirc, A)}{100}\right)^2 + \left(\frac{p(\bigcirc, B)}{100}\right)^2 + \left(\frac{p(\triangle, A)}{100}\right)^2 + \left(\frac{p(\triangle, B)}{100}\right)^2\right]$$
- If the Sender's shape is is Circle $\bigcirc$ and the chosen bag is Bag B
$$S = 100 - 50 \cdot \left[\left(\frac{p(\bigcirc, A)}{100}\right)^2 + \left(1 - \frac{p(\bigcirc, B)}{100}\right)^2 + \left(\frac{p(\triangle, A)}{100}\right)^2 + \left(\frac{p(\triangle, B)}{100}\right)^2\right]$$
- If the Sender's shape is is Triangle $\triangle$ and the chosen bag is Bag A
$$S = 100 - 50 \cdot \left[\left(\frac{p(\bigcirc, A)}{100}\right)^2 + \left(\frac{p(\bigcirc, B)}{100}\right)^2 + \left(1 - \frac{p(\triangle, A)}{100}\right)^2 + \left(\frac{p(\triangle, B)}{100}\right)^2\right]$$
- If the Sender's shape is is Triangle $\triangle$ and the chosen bag is Bag B
$$S = 100 - 50 \cdot \left[\left(\frac{p(\bigcirc, A)}{100}\right)^2 + \left(\frac{p(\bigcirc, B)}{100}\right)^2 + \left(\frac{p(\triangle, A)}{100}\right)^2 + \left(1 - \frac{p(\triangle, B)}{100}\right)^2\right]$$

The computed score S will be rounded to be a whole number between 0 and 100 and will represent the probability of winning 100 points in a lottery. Indeed, after the score is computed, the Computer will randomly draw a whole number from 1 to 100, with each number being equally likely. The Receiver will win the 100 points prize if their score S is greater than or equal to the randomly drawn number, and 0 points otherwise.

Given this rule, it is immediate to see that the higher the score, the higher the probability that the Receiver wins the 100 points. As explained in the video instructions, the more accurate the Receiver's guesses are, the higher the score. **In particular, this compensation rule is purposefully designed so that the Receiver has the greatest chance of winning the 100 points prize when their guesses equal their true beliefs for each combination of the Sender's shape and the chosen bag**.

In the additional tasks, the Receiver will be asked to guess the probability of some other unknown events. When that is the case, the Receiver will be, again, remunerated according to a rule equivalent to the one explained above. If the number of probabilities to assess in the new task changes, the main mechanism for computing the score will remain the same, but the details of the formulas will reflect this new set of probabilities. **Again, what matters is that these compensation rules are designed to remunerate accuracy and to give the Receiver the greatest chance of winning 100 points when their guesses equal their true beliefs.**