Higher-order Learning

Piotr Evdokimov* and Umberto Garfagnini†

September 10, 2019

Abstract

We use a novel experiment to study how players form and update their beliefs about the beliefs of others. The experiment is designed so that players' choices correspond to higher-order expectations about a common state of the world (e.g., economic fundamentals). We find that higher-order expectations reflect a certain degree of sophistication. In particular, they respond to information about how lower-order expectations are formed and are updated more slowly when a player faces uncertainty about others' information. However, we also identify a novel behavioral bias: Higher-order expectations diverge over time, even as more information becomes available. This finding points to a failure of theory of mind, i.e., a failure of subjects to correctly predict the way in which others process information.

JEL Classification: C90, D83, D89

Keywords: Higher-order expectations, learning, theory of mind

*Department of Applied Economics, Higher School of Economics.

†School of Economics, University of Surrey.

‡We are grateful to Ryan Oprea and seminar participants at the ESA World Meeting (Berlin) and the University of Southampton. Financial support from the Asociación Mexicana de Cultura A.C. is acknowledged.

1 Introduction

Strategic decision-making requires agents to form beliefs about how others process information.[1] For example, one's expectations about what expectations other agents in the economy form about key variables like inflation, the state of the economy, etc., following the release of new economic data, matter for individual investment decisions. Higher-order beliefs arise naturally in games of incomplete information, (see, e.g., Harsanyi, 1967; Mertens and Zamir, 1985), figure prominently in epistemic models (e.g., Aumann and Brandenburger, 1995; Chen et al., 2016), and play an important role in financial markets (Keynes, 1936; Morris and Shin, 2002; Banerjee et al., 2009).

The following simple example will help illustrate this point.[2] Two agents must decide whether or not to invest in a project. The return of the project depends on whether both agents invest and an underlying state of the world (e.g., the economic fundamentals). If the agents receive signals about the state over time, will they be able to coordinate on an efficient course of action? The answer depends on whether each agent assigns a high enough probability to the event that her partner assigns a high enough probability to the same state, that her opponent does the same, that both agents believe that their opponent does the same, and so on. If the underlying state of the world does not become approximate common knowledge, agents may avoid investing even if they are individually quite certain that a good state has realized. In his study of the effect of disinflationary policies, Phelps (1983) also emphasizes the importance of uncertainty due to the "forecasts people will make of the other people's forecasts, their forecasts of other forecasts, and so on ad infinitum." Building on Phelps' work, Woodford (2001) shows that monetary disturbances can have real effects if agents are uncertain about the higher-order expectations of others.[3]

We introduce a novel experimental design to shed light on how players form and update their beliefs about the beliefs of others. The key component is a network game which we call the *chain game* (Figure 1). In an n-player chain game, Player k's payoff increases as her action better matches the action of her predecessor in the network, Player $k-1$, whose payoff in turn increases as her action better matches Player $k-2$'s action, and so on. Finally, Player 1's payoff increases as her action better matches an underlying state of the world. In the unique equilibrium of this game, Player k's action corresponds to her k-th order expectation about the state of the world. This is defined as Player k's expectation about Player $k-1$'s expectation about Player

[1] Many studies in economics and psychology have explored how individual beliefs about an underlying state of the world are updated in response to information. Early studies documenting violations of Bayes' rule include Tversky and Kahneman (1971), Tversky and Kahneman (1974), and Grether (1980).

[2] The example is borrowed from Cripps et al. (2008).

[3] See also Angeletos and La'O (2013) and Angeletos et al. (2018) for recent contributions to this literature.

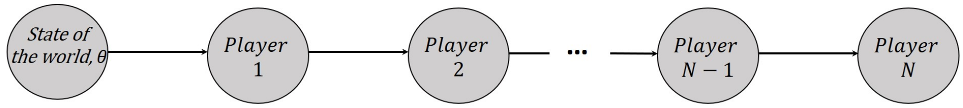


Figure 1: The chain game. An arrow from i to j means that i affects j's payoffs.

$k-2$'s expectation, ..., about Player 1's expectation about the realized state of the world. Thus, the chain game naturally triggers higher-order thinking with a degree that increases in a player's position in the chain while anchoring expectations of *any* degree to uncertainty about an exogenous state of the world.

When players receive the same information about the state of the world, equilibrium behavior predicts that first- and all higher-order expectations are equal to each other. This seemingly intuitive result relies on non-trivial assumptions about the depth of reasoning of players in the higher-order roles. As Player 1 observes a signal about the state, she updates her beliefs to better match her action to the state. Player 2 only cares about the state indirectly through its impact on Player 1's choice. Therefore, if she observes the same signal as Player 1 *and* correctly anticipates how Player 1 will process this piece of information, she will take an action which exactly matches Player 1's.[4] Thus, her second-order expectations about the state will correspond to Player 1's first-order expectations. This logic easily extends to higher-order roles but with stricter assumptions about the depth of reasoning imposed upon players.

Our experimental design allows us to study the evolution of higher-order expectations as progressively more information about the state is observed and whether higher-order learning—that is, convergence of higher- to lower-order expectations—is achievable. This is behaviorally important, as cognitive limitations could constrain a player's ability to engage in higher-order reasoning in any particular chain game; nevertheless, more information about the state of the world could still lead all players to form progressively similar expectations about the expectations of others.[5]

In the experiment, we implement the chain game as a simple sequential betting task. Subjects are randomly matched into groups of three. Before any decision is made, an unknown state of the world is drawn at random and held fixed for the duration of the game. It is common knowledge that the state is the same for every player in the group. Over the course of 30 periods, group members observe signals about the state of the world, as in standard belief updating tasks (e.g. Holt and Smith, 2009). After receiving her signal, Player 1 is incentivized to choose an action

[4]Recall that Player 2's earnings decrease in the distance between Player 1's action and her own.

[5]This conclusion seems behaviorally plausible, and certainly empirically testable, even if players are non-Bayesian, as long as they hold an approximately correct theory of mind about the way in which other players process information.

as close as possible to the realized state (her *first-order expectations*), Player 2 is incentivized to choose an action as close as possible to Player 1's action (her *second-order expectations*), and Player 3 is incentivized to choose an action as close as possible to Player 2's action (her *third-order expectations*).[6] Subjects receive no feedback about the behavior of their matched partners for the duration of the experiment,[7] and each subject is paid only on her performance in one randomly chosen period.[8]

The experiment is designed around two treatments. In the **public treatment**, all players in the same team observe the exact same signal in each period of the game, as described above.[9] In the **private treatment**, Player 1, Player 2, and Player 3 observe conditionally independent signals from the same distribution and each player can only observe her own signal. It should be clear from the preceding discussion that a comparison of actions across player roles in the public treatment can shed light on what expectations players form about other players' expectations and whether these beliefs are correct. A wedge between higher- and lower-order expectations would be consistent either with the hypothesis that higher-order players assume that their counterparts are less able to process information (i.e., are less rational) than they themselves are or simply that players recognize and anticipate the heterogeneity in other players' updating rules. The first concern is motivated by recent experimental literature (e.g., Weizsäcker, 2003) which finds that subjects underestimate the rationality of their opponents on average, thus believing that they are facing subjects who are less rational than they are themselves.[10] The second concern is motivated by a longstanding literature studying information processing in belief updating tasks.[11]

On the other hand, the lack of a wedge between higher- and lower-order expectations could have multiple interpretations. First, it could reflect sophisticated behavior wherein the players are able to make correct inferences about how other players' expectations are formed. Second, if Player $i > 1$ fails to take the reasoning processes of the other players into account, she might simply perceive the game as taking an action which matches the state of the world, that is, by behaving as a Player 1. In the public treatment, this naive behavior is indistinguishable from

[6]Thus, each player only takes one action. This is to avoid hedging across multiple decisions and to keep each player's task as simple as possible. In particular, the focus of the paper is to study how players' mental models about other players accommodate the arrival of new information, rather than providing a typology of how many subjects are capable to engage in k levels of reasoning, for any k, which has been the focus of the level-k literature.

[7]Furthermore, Player 1 is only informed about her own task, Player 2 is informed both about her own and Player 1's task, while Player 3 is the only player informed about all tasks. This is done because it both simplifies the design and makes social preferences theoretically irrelevant.

[8]This mutes any intertemporal incentives and thereby ensures that each period of the game is a static game with only information about the state changing across iterations.

[9]The computer draws a new signal at the beginning of every period.

[10]Weizsäcker (2003) finds substantial heterogeneity in subjects' beliefs, as we do.

[11]We refer to Holt and Smith (2009) for a recent review of the literature.

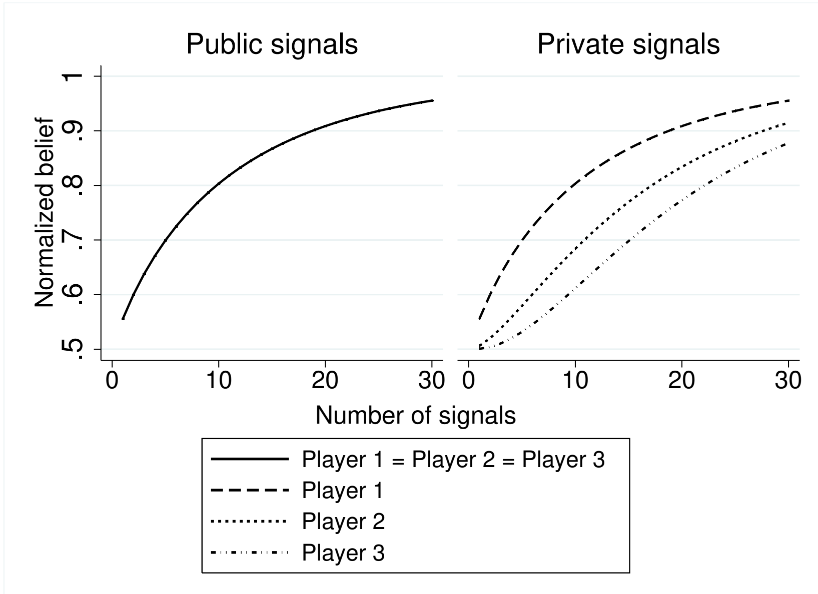


Figure 2: The predicted evolution of expected first-, second-order, and third-order expectations.

the more sophisticated one. This concern motivates our second treatment in which the players' signals are private (conditionally independent).

In the unique equilibrium of the chain game with private signals, players' uncertainty about the information of others creates a positive wedge between the k-th order and $k-1$-th order expectations, for any $k \geq 2$ (Figure 2). Player 2 uses her private history of signals to update her belief about the state of the world. This information is used to form a belief about the histories Player 1 might have observed. To account for the difference between her own and Player 1's observed histories, and conditional on the true state of the world, Player 2's expectation about Player 1's expectation is shaded down compared to Player 1's own expectation about the state. As Player 3 faces even more uncertainty, she also shades down her own expectation about Player 2's expectation, and so on. Thus, a direct comparison between behavior with public and private signals can shed light on whether subjects' behavior is naive or reflects the evolution of higher-order beliefs.[12] For instance, statistically indistinguishable higher- and lower-order beliefs in the public but not the private treatment would suggest that behavior in the chain game is not naive. A greater wedge in the private than the private treatment would also suggest that behavior is not driven by naïvité alone.

[12]While behavior differs depending on the informational condition, differences in higher-order expectations vanish as players observe more information with both cases leading to the same long-run prediction of *higher-order learning*–that is, Player k's k-th order expectation converges to Player $k-1$'s $k-1$-th order expectation, for any $k \geq 2$. An even stronger theoretical result holds, due to Cripps et al. (2008), that the state of the world becomes approximate common knowledge. Cripps et al. refer to this case as *common learning*.

We find that higher-order expectations respond to the public vs. private nature of information: Average first- and higher-order expectations are statistically indistinguishable when learning occurs through public signals, while higher-order expectations are updated more slowly with private signals (Result 1). This effect can be entirely explained by a shift toward conservatism in belief updating (Result 2), highlighting the (predicted) increased difficulty of updating higher-order expectations in the presence of private information. We also observe substantial heterogeneity in belief formation processes. Thus, the higher-order players are not able to correctly model the beliefs of other players in the experiment (Result 3). This failure of theory of mind is consistent with the hypothesis that the higher-order players assume their lower-order counterparts to adopt a similar updating rule to theirs, as opposed to allowing for the possibility that different individuals might follow different updating rules.[13]

Our main finding points to a failure of higher-order learning: The accuracy of higher-order expectations declines over time, even as more information about lower order expectations becomes available (Result 4). The reasons underlying this divergence of higher- and lower-order expectations are different in the public and the private treatment. In the public treatment, the divergence is driven by the players whose behavior responds little to information, while in the private treatment it is driven by the players that are better described as being Bayesian. The intuition in the public treatment is straightforward. As lower-order expectations converge to the truth, the higher-order expectations that are not updated become less accurate over time. The intuition in the private treatment is more subtle. The updating of lower-order expectations is sufficiently slow and the private signals are sufficiently uncertain that Bayes' rule provides an overly optimistic benchmark. As a result, the prediction of a close-to-Bayesian player becomes less accurate over time.

This paper complements several strands of literature. A longstanding literature has investigated the question of belief updating in the presence of new information (e.g., Grether, 1980, 1992; Holt and Smith, 2009) finding violations of Bayes' rule as well as substantial heterogeneity in updating rules. More recent contributions also find that decision makers process information about personal characteristics such as IQ asymmetrically, overweighting good news and underweighting bad news (e.g., Eil and Rao, 2011; Mobius et al., 2014; Coutts, 2018). By contrast, we experimentally manipulate and investigate the updating of higher-order expectations, thus shedding light on information processing in strategic settings.

We also contribute to the growing experimental literature on higher-order reasoning. Nagel (1995) introduces level-k thinking in the context of guessing games and finds mostly two levels

[13] We refer to theory of mind as "the ability to think about others' thoughts and mental states to predict their intentions and actions" (Coricelli and Nagel, 2009, pag. 9163).

of depth of reasoning.[14] Recently, Arad and Rubinstein (2012) use a novel game which naturally triggers level-k thinking to identify subjects' depth of reasoning and find higher levels of sophistication.[15] To the best of our knowledge, this paper is the first to investigate how mental models of other players are affected by the arrival of new information.[16] To this literature, we contribute the finding that higher-order expectations do not converge over time even as more information becomes available.

The paper is organized as follows. Section 2 provides a formal definition of the chain game and proves our key theoretical predictions. Section 3 contains the experimental design. Section 4 explains the main results. Section 5 provides a discussion of our experimental findings, and Section 6 concludes.

2 Theoretical Background

This section formally introduces our theoretical framework. We first define an N-player chain game.

Definition 1. *An N-player chain game* Γ *is a tuple* $\Gamma = \left\langle I; \Theta; (A_i, T_i, \tau_i, K_i, \pi_i, u_i, p_i)_{i \in I} \right\rangle$, *where:*

- $I = \{1, ..., N\}$ *is a finite set of players;*
- Θ *is a finite set of states of the world, with the true state of the world* $\theta \in \Theta$*;*
- $A_i = [0, 1]$ *is the set of actions for Player i;*

[14]Early contributions to the level-k literature also include Stahl and Wilson (1994, 1995). Huck and Weizsäcker (2002) study whether subjects' elicited beliefs about their opponents' behavior are correct on average. Although subjects are able to correctly predict the choice frequencies of other subjects on average, they find a significant and systematic bias toward a uniform prior. Kübler and Weizsäcker (2004) study how subjects process information generated through their predecessors' choices in a social learning framework. Using an error-rate model which allows to estimate how subjects reason about other subjects' behavior, they also find that subjects underestimate the rationality of their immediate predecessors (similar to what found by Weizsäcker 2003) and that the average subject's reasoning does not exceed two steps.

[15]Using a clever experimental design and a novel game, Kneeland (2015) is able to recover the implied level of higher-order rationality from subjects' choices. She finds higher levels of rationality in her experimental subjects compared to previous studies. Similarly, Lim and Xiong (2016) provide a different experimental design than Kneeland's to identify a subject's level of rationality.

[16]Nyarko and Schotter (2002) investigate whether elicited beliefs can better explain behavior than alternative models like fictitious play. Thus, their subjects need to update their beliefs about their opponent's future action. In order to recover a subject's mental model of how her opponent might be processing new information about the game, we instead focus on introspective learning by introducing uncertainty about an underlying state of the world. Our scope and results are therefore quite different.

- $\tau_i : \Theta \to T_i$ *is Player i's signal function which maps a state of the world into a (possibly multi-dimensional) signal* $\mathbf{t}_i \in T_i$;[17]

- K_i *is an individual-specific random variable uniformly distributed over the interval* $[0, 1]$;

- $\pi_i : A_i \times A_{i-1} \times [0, 1] \to \mathbb{R}$ *is Player i's payoff function, with* $A_0 \equiv \Theta$, *and*

$$\pi_i(a_i, a_{i-1}, k_i) = \begin{cases} R_1, & \text{if } (a_i - a_{i-1})^2 \leq k_i, \\ R_0, & \text{otherwise,} \end{cases} \tag{1}$$

  with $R_1 > R_0 \geq 0$.

- $u_i : \mathbb{R} \to \mathbb{R}$ *is Player i's utility function over payoffs, which is assumed to be increasing;*

- $p_i \in \Delta(\Theta)$ *is Player i's prior probability measure over states of the world.*

The chain game is a simultaneous-move directed network game of asymmetric information in which, for any $i \geq 2$, Player i's payoffs depend on her action, the action of her immediate predecessor in the network, Player $i - 1$ (Nature in the case of Player 1), and a uniformly distributed random variable on the interval $[0, 1]$.[18] Player i obtains a reward R_1 if the squared distance between her action and her predecessor's action is less than or equal to the random draw k_i. Players are allowed to observe different information about the state of the world through the signal functions $\{\tau_i\}_{i \in I}$. To simplify notation, we further define $\hat{u}_i = u_i \circ \pi_i$.

A strategy for Player i is a mapping $s_i : T_i \to A_i$. Given a strategy profile $\mathbf{s}_{-i} = (s_j)_{j \neq i}$ for Player i's opponents, the expected utility of a type $\mathbf{t}_i$ from choosing an arbitrary action $a_i \in A_i$ is given by

$$U_i(a_i, \mathbf{s}_{-i}; \mathbf{t}_i) = \int_0^1 \sum_{\theta \in \Theta} p_i(\theta | \mathbf{t}_i) \hat{u}_i \left(a_i, s_{i-1}(\tau_{i-1}(\theta)), k\right) dk \tag{2}$$

with the convention that $s_0(\tau_0(\theta)) = \theta$. A Bayesian equilibrium of this game is a strategy profile $\mathbf{s}$ such that, for any $i \in I$ and $\mathbf{t}_i \in T_i$,

$$s_i(\mathbf{t}_i) \in \underset{a_i \in A_i}{\operatorname{argmax}} \, U_i(a_i, \mathbf{s}_{-i}; \mathbf{t}_i) \tag{3}$$

Before we characterize the equilibria of this game, it is useful to formally define a player's k-th order expectations.

---

[17] We interchangeably refer to $\mathbf{t}_i$ as Player i's *type*.

[18] The use of a uniformly distributed draw is a trick borrowed from Hossain and Okui (2013) which makes the equilibrium predictions independent of any attitudes toward risk. This is formally shown in the proof of Proposition 1 below.

Definition 2. *Fix an N-player chain game. For any $k > 1$ and type $\mathbf{t} \in T_i$, a player's k-th order expectations about the state of the world, conditional on being of type $\mathbf{t}$, are defined as*

$$\bar{E}^k[\theta|\mathbf{t}] = \underbrace{E[E[...E[E[}_{k \text{ times}}\theta|\tau_1(\theta)]|\tau_2(\theta)]...|\tau_{k-1}(\theta)]|\mathbf{t}] \tag{4}$$

with $\bar{E}^1[\theta|\mathbf{t}] = E[\theta|\mathbf{t}]$.

A player's first-order expectations correspond to her expectations about the state of world given her own type $\mathbf{t}$, that is, $\bar{E}^1[\theta|\mathbf{t}] = E[\theta|\mathbf{t}]$. Her second-order expectations correspond to her expectations, given her own type, about Player 1's expectations about the state of the world conditional on Player 1's own type, which is given by $\tau_1(\theta)$, and so on.

Proposition 1. *The N-player chain game has a unique Bayesian equilibrium in which $s_i^*(\mathbf{t}_i) = \bar{E}^i[\theta|\mathbf{t}_i]$, for any $i \in I$.*

The equilibrium action of Player i corresponds to her i-th order expectations about the state of the world. Thus, despite the fact that only the action of Player $i-1$ has a direct payoff consequence for Player i, equilibrium behavior requires Player i to think about the action taken by $i-1$'s predecessors as well in the order given by the network structure of the game. The regression stops at Player 1, whose action is anchored to her expectations about the state of the world. Thus, the chain game requires greater depth of reasoning as a player's index increases. For example, Player 2's equilibrium action corresponds to her expectations about Player 1's expectations about the state of the world (that is, her *second-order expectations*). Player 3's equilibrium action corresponds to her expectations about Player 2's expectations about Player 1's expectations about the state of the world (that is, her *third-order expectations*), and so on.

This feature of the chain game makes it appealing for experimental manipulations. By varying the signal functions, the chain game allows us to affect higher-order beliefs and consequently behavior. Furthermore, we can investigate how players' higher-order expectations vary as the amount of information provided increases by increasing the informativeness of the signal $\mathbf{t}_i$. Next, we prove additional properties of the chain game which we are going to experimentally test.

Consider a sequence of chain games $\Gamma^\infty \equiv \{\Gamma_n\}_{n=1}^\infty$. Γ^∞ is said to be *feasible* if: i) the state of the world is persistent across games, that is, for any n, $\theta_n = \theta$; ii) players share a common prior, that is, for each n and i, Player i's prior belief $p^n = p$; iii) for any n and i, $T_i^n = \times_{k=1}^n T$, for some finite set T. Thus, for fixed n, Player i observes an n-dimensional signal $\mathbf{t}_i^n = (t_i^1, t_i^2, ..., t_i^n)$. Our favorite interpretation, and the one used in our experimental implementation, is that t_i^k would be the signal observed by Player i in period k if she observed each one-dimensional signal sequentially.

Proposition 2 characterizes properties of the expected actions played by each player, conditional on the true state of the world, given two different informational regimes.

PROPOSITION 2. *Suppose that Γ^∞ is a feasible sequence of chain games with $|\Theta| = |T| = 2$. Then, for any n, the following holds true:*

a) ***Public signals:*** *If it is common knowledge that the signals in chain game n are all perfectly positively correlated, then $s_i^n(\mathbf{t}^n) = E_n[\theta|\mathbf{t}^n]$ for any $\mathbf{t}^n$. Thus, for any $i \in I$, $E_n[s_i^n(\mathbf{t}^n)|\theta] = E_n[s_1^n(\mathbf{t}^n)|\theta]$.*

b) ***Private signals:*** *If it is common knowledge that the signals in chain game n are private and conditionally independent,*[19] *then $E_n[s_i^n(\mathbf{t}_i^n)|\theta] < E_n[s_{i-1}^n(\mathbf{t}_{i-1}^n)|\theta]$, for any $i \in I$.*

When signals are public, all players' actions (and thus average actions) coincide as they all observe the same information in each given game. Given the mapping between equilibrium actions and expectations about the state of the world, this means that first and higher-order expectations are identical. While Proposition 2.*a*) is intuitive, it is important to point out that its simplicity relies on nontrivial assumptions about higher orders of rationality *and* subjects' ability to engage in higher-order reasoning. Formally, the result holds if: *i*) Player 1 is Bayesian; *ii*) Player 2 is Bayesian and thinks that Player 1 is Bayesian; and *iii*) Player 3 is Bayesian, thinks that Player 2 is Bayesian, and thinks that Player 2 thinks that Player 1 is Bayesian.[20] Empirically, Proposition 2.*a*) is also consistent with a more naive kind of behavior. If Player $i > 1$ fails to take the reasoning processes of the other players into account, she might simply perceive the game as taking an action which matches the state of the world, that is, by behaving as a Player 1. This naive behavior is indistinguishable from the more sophisticated one.

By contrast, when signals are private, higher-order expectations are progressively lower. The private nature of information leads a player to shade down her action, on average, as her role in the game increases. For example, Player 2 uses her own observed history of signals to update her belief about which state has realized. This information is used to form a belief about the history observed by Player 1 but the two histories are likely to be different. This uncertainty leads Player 2 to lower her equilibrium action in expectation, and so on. Thus, if a player behaves naively, we should also observe no average difference between her action and that of a player in a lower role. Therefore, changing the type of information that players observe about the state of the world can help to disentangle naive from sophisticated behavior.

[19] Let $T = \{x, y\}$. Here, conditional independence means that $\text{Prob}_i(\mathbf{t}_i^n|\theta) = \text{Prob}(x|\theta)^{n_x}(1 - \text{Prob}(x|\theta))^{n-n_x}$, where n_x denotes the number of signals of type x out of n total (unidimensional) signals in $\mathbf{t}_i^n$.

[20] These requirements can be somewhat relaxed even without the assumption of Bayesianism, however it is still needed that players in higher-order roles are capable of greater depth of reasoning.

We can go one step further and show that higher-order expectations converge to their lower-order counterparts as all players observe progressively more information about the state of the world. While we do not test the following proposition directly, as the experiment only includes a finite number of periods, it provides an interesting contrast to our main result:

PROPOSITION 3 (HIGHER-ORDER LEARNING). *For any* $i > 2$, $|E[s_i^n(\mathbf{t}_i^n)|\theta] - E[s_{i-1}^n(\mathbf{t}_{i-1}^n)|\theta]| \to 0$*, and* $E[s_1^n(\mathbf{t}_1^n)|\theta] \to 1$ *as* $n \to +\infty$*, regardless of whether signals are private or public.*

3 Experimental Design

Our experimental design consists of a three-player version of the chain game. At the beginning of each session, the subjects are matched in teams of three. Within a team, each subject is randomly assigned to one of three roles: Player 1, Player 2, or Player 3, with exactly one subject in each role. The roles and teams stay fixed for the duration of the session. A session consists of three incentivized rounds,[21] with each round lasting for 30 periods.[22] Each subject could participate in only one session.

Each subject is paid on the basis of one randomly chosen period of a randomly chosen round. Azrieli et al. (2018) show theoretically that selecting one task at random is the only incentive compatible way to pay subjects under a mild monotonicity assumption on subjects' preferences. Paying for only one randomly chosen round breaks any risk and intertemporal hedging across periods thus making each period of a round into a standalone chain game.

Subjects are told that there are two urns, *Orange* and *Purple*, each containing 3 balls. The *Orange* urn contains 2 orange balls and 1 purple ball, while the *Purple* urn contains 1 orange ball and 2 purple balls. At the beginning of each round, the computer selects one of the two urns with equal probability for each three-player team. None of the subjects are told which urn is selected for their team.

In the **public treatment**, the computer draws a ball with replacement from the selected urn in every period of a round. The color of the ball is shown to all the subjects in the same team. I.e., the signals in this treatment are public. In the **private treatment**, the computer draws a ball from the selected urn with replacement for each subject in the same team. Each subject is shown the color of her selected ball but not the color of the selected ball for her matched partners. I.e., the signals in the private treatment are private and conditionally independent.

[21] Due to an experimenter error, the first session of the experiment was programmed for five instead of three rounds. For that session, we only analyze the first three rounds to be consistent with all the remaining sessions.

[22] An experiment like ours faces a trade-off between having a large number of periods and a large number of rounds; we favored a large number of periods to address the questions of learning.

The chain game in a period n can be easily implemented through the use of the "binarized scoring rule" of Hossain and Okui (2013). The rule works as follows:

1. Player i takes an action $a_i \in [0, 1]$;[23]
2. A random variable of interest, Z, is realized;
3. The player's loss is computed according to a loss function $L(a_i, z)$, where z is the realization of Z;
4. The computer draws a number k uniformly at random from the interval $[0, K]$, for some commonly known $K \geq \max_{a,z} L(a, z) > 0$;
5. If $L(a_i, z) \leq k$, Player i receives a monetary reward of R_1, otherwise the player receives $R_0 < R_1$.

We employ a quadratic loss function $L(a, z) = (a - z)^2$, and set $K = 1$. For Player 1, Z is either 1, if the selected urn is *Orange*, or 0 otherwise. For Player $i \in \{2, 3\}$, Z corresponds to the action chosen by Player $i - 1$. Regardless of attitudes to risk, the optimal response of Player 1 is to report her probability of the urn being *Orange*. For Player 2 and Player 3, the optimal response corresponds to the expectation of the action chosen by the preceding player, that is, $a_i = E[a_{i-1}|\mathcal{I}_i]$,[24] where $\mathcal{I}_i$ is Player i's information set, $i = 2, 3$, which correspond to the equilibrium actions of the chain game.

Each subject takes one action, as opposed to three actions, to avoid possible contamination between lower- and higher-order expectations. Also, subjects in the role of Player 1 are told that they will be matched to two other subjects in the role of Player 2 and Player 3 but that the decisions made by those players will be inconsequential for her own performance. Moreover, Player 1 is not told what tasks Player 2 and Player 3 are given. Player 2 is explained Player 1's task but not Player 3's task. Player 3 is the only player with full information about the structure of all tasks.[25] Finally, as is standard practice in experimental economics, the experiment is framed neutrally in that it avoids any reference to guesses, beliefs, or expectations, instead explaining the task as a betting problem.

Subjects receive no feedback about their own performance or the performance of their matched partners for the entire duration of the experiment. At the end of each round, the correct composition of the urn is revealed to all subjects in the same team and no other information is

[23]The action set in the experiment was a discretized version of [0,1] with $A_i = \{0, 0.01, 0.02, ..., 0.98, 0.99, 1\}$.
[24]This follows immediately from Theorem 1 in Hossain and Okui (2013) because $K \geq \max_{a,z\in[0,1]^2} L(n, z)$.
[25]Notice that this makes social preferences theoretically irrelevant.

disclosed.[26] While lack of feedback is common in experiments measuring subjects' beliefs about other subjects' beliefs (see e.g., Stahl and Wilson, 1995; Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006; Costa-Gomes and Weizsäcker, 2008), we refrain from providing subjects with feedback for two main reasons. First, our experiment tries to identify a subject's mental model of other subjects and the possible effect of introspective learning rather than her response to reinforcement learning. Second, it would not be possible to implement the private treatment with period-to-period feedback. Observing a partner's action would reveal information about that partner's private signals, thus affecting a subject's choice of action in subsequent periods.

3.1 Implementation

The experiment was conducted at Instituto Tecnológico Autónomo de México in Mexico City between October and December 2017 using the software z-Tree (Fischbacher, 2007). Data was collected from 120 subjects in 7 sessions for the public treatment and from 129 subjects in 8 sessions for the private treatment. A session lasted 75 minutes on average. All subjects were undergraduate students recruited from the general student population.

Each session starts with subjects signing the consent forms, reading the instructions, and completing an incentivized quiz.[27] Every subject is guaranteed a 100 Mexican pesos show-up fee (≈US$5.26 at the time of the experiment) in addition to the earnings from the quiz (2 Mexican pesos for each correct answer). These earnings are called the subject's "guaranteed earnings." Each subject is also given an initial endowment of 80 pesos which the subject has a chance to either double or lose completely according to the following procedure based on the binarized scoring rule.[28] The computer randomly selects a period, for each subject, from among all the periods in all rounds of the experiment. Given a subject's loss for the period from her decision, the computer independently draws a number, k, uniformly between 0 and 1 and the subject's

[26]At the end of a round of the private treatment, a subject is also shown the cumulative number of balls drawn for each of her matched partners, but still no information about her partners' guesses. This is chosen to further emphasize to a subject the privacy of signals and that her matched partners can observe different histories than her own. While doing this provides additional information to a subject about likely histories, we believe this issue to be minor as we only give a subject a snapshot of the cumulative number of balls of each color seen by her partners at the end of the round rather than the sequence of balls drawn from the urn in each period, which we believe it would be more likely to inform decision making in subsequent rounds.

[27]The online appendix with the instructions of all our treatments can be found here. While the sample instructions are in English, the actual instructions were administered in Spanish. The answers to all of the quiz questions are incentivized.

[28]At the time of the experiment, the minimum wage in Mexico was about 70 pesos per day, which is arguably a poor reference point for students at a private research university such as ITAM. For a better one, consider that the cost of a 15km Uber ride was around 80 pesos.

"additional earnings" are determined as follows:

$$Additional\ earnings = \begin{cases} 2 * \text{Initial endowment}, & \text{if } (n-z)^2 \leq k, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The payment rule is clearly explained to the subjects in the instructions, and several examples are provided.[29]

4 Results

Figure 3 shows the average reported actions of subjects in all player roles and all treatments, as well as the benchmark equilibrium actions given the observed signal histories in the experiment. The actions are normalized by the correct state so that the data being plotted is x when the state is orange and $1-x$ when the state is purple, where x is the subject's response. Overall, subjects' average actions show sizable deviations from the equilibrium predictions. The first column of Table 1, which regresses the normalized actions against their equilibrium counterparts, shows a coefficient of 0.516.[30] This result is in line with previous studies showing deviations from Bayesianism in tasks similar to ours (e.g., Grether, 1980). However, unlike these prior studies, our focus is not Bayesian behavior *per se*. Rather, we are interested in studying whether higher-order expectations respond differently to the nature of information (private vs. public) and whether higher- and lower-order converge as more information becomes available. We turn to the first of these two questions next.

In line with the equilibrium predictions, higher-order expectations evolve more slowly with private than public signals. The significance of this result is assessed in the second column of Table 1, where the normalized expectations of Players 2 and 3 are regressed against a dummy variable for the treatment with private signals. The private dummy is negative and significant ($P < 0.01$), suggesting that higher-order expectations are able to account for the uncertainty associated with private information. We highlight this as follows:

Result 1 (Effect of Private Information). *Higher-order expectations respond to the type of information received and evolve more slowly with private than public signals.*

[29]The software also provides an initial screen that a subject can use to experiment with the calculation of the loss based on the distance between a hypothetical guess (entered by the subject) and possible target numbers.

[30]Since expectations are always between 0 and 1, we can also perform a fractional logit regression (Papke and Wooldridge, 1996) which leads to the same qualitative and quantitative results as shown in Table 6 in the appendix.

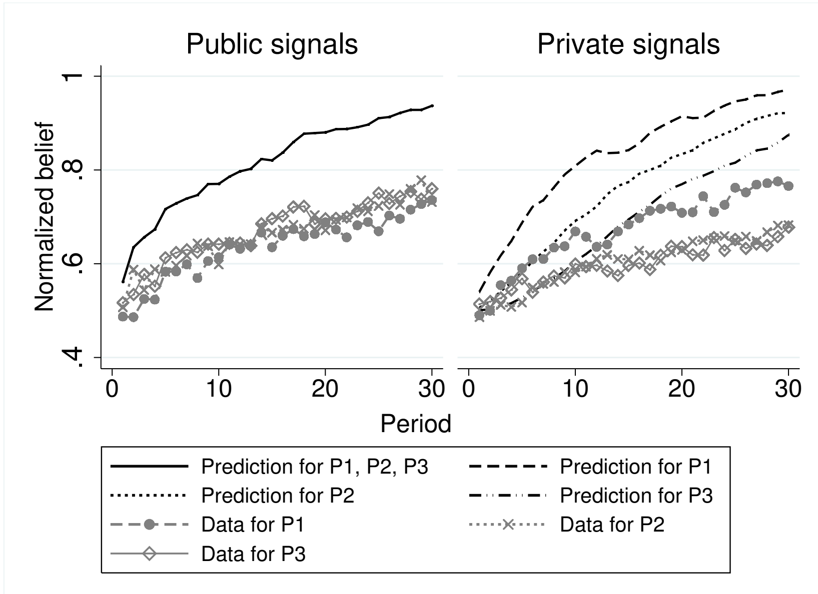


Figure 3: Observed and predicted lower- and higher-order expectations for all treatments and player types.

While equilibrium predicts identical expectations for Player 1, Player 2, and Player 3 after every public history, the belief of Player k might differ from that of Player $k-1$ if Player k believes that Player $k-1$ is less able to process information than Player k is. We do not find evidence of such failure of belief in information processing abilities at the aggregate level. Figure 3 suggests that Player 1, Player 2, and Player 3 take similar actions on average when information is public. This is confirmed in the third column of Table 1, where the normalized expectations of Players 1, 2 and 3 in the public treatment are regressed against a Player 2 dummy and a Player 3 dummy. We find that neither dummy variable is significant (smallest $P = 0.3539$) and the dummy for Player 2 is not significantly different from that for Player 3 ($P = 0.8115$).[31] Thus, at least at the aggregate level, players in higher-order roles do not seem to assume that their lower-order counterparts are more or less able to process information (i.e., are more or less rational) than they themselves are. This observation contrasts with a recent experimental literature based on choice data which finds that subjects underestimate the rationality of others (see, e.g., Weizsäcker, 2003; Kübler and Weizsäcker, 2004), although Costa-Gomes and Weizsäcker (2008) find that subjects show more strategic thinking when stating beliefs about their opponents' actions.[32]

[31] An additional regression including a period variable and interactions with the player dummies shows that only the period variable is significant ($P < 0.001$).

[32] Overall, Figure 3 and Table 1 obscure substantial heterogeneity in information processing. For instance, while the average actions of Player 1, Player 2, and Player 3 were not significantly different in the public treatment, some subjects in the higher-order roles overestimated the lower-order actions, while others underestimated them. We discuss these observations in detail when we analyze behavior at the level of individual subjects.

	Dependent variable: Normalized observed expectations			
	(1)	(2)	(3)	(4)
	All	Players 2 and 3 (private and public)	Players 1, 2 and 3 public	Players 1, 2 and 3 private
Bayesian prediction	0.516****			
	(0.0319)			
Private		-0.0651***		
		(0.0221)		
Player 2			0.0264	-0.0697*
			(0.0373)	(0.0358)
Player 3			0.0341	-0.0742**
			(0.0366)	(0.0343)
Constant	0.235****	0.666****	0.636****	0.673****
	(0.0191)	(0.0160)	(0.0292)	(0.0276)
Observations	22410	14940	10800	11610

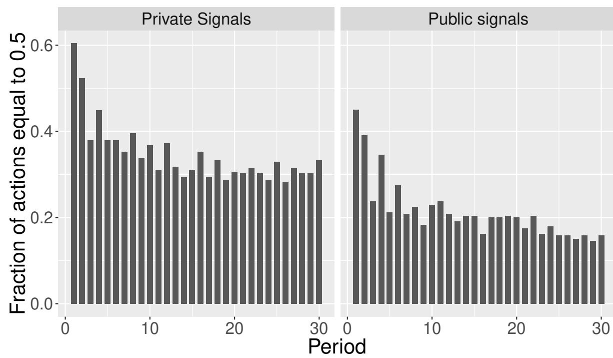
Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 1: Analysis of average normalized observed expectations.

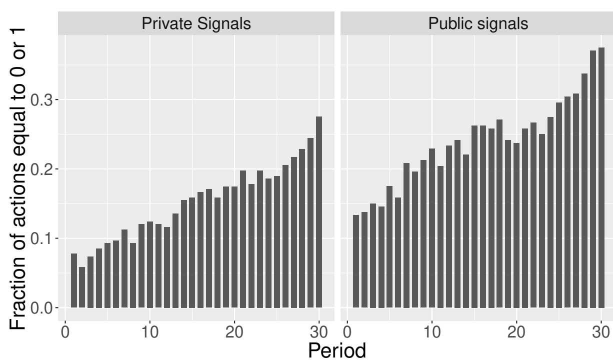
Turning to the private treatment, the equilibrium benchmark predicts that the action of Player 3 is closer to 0.5 than that of Player 2. Provided that Player 3 has the correct model of how Player 2's beliefs are formed, this prediction holds even if Player 2's beliefs are non-Bayesian. As seen in Figure 3, we do not find evidence of such behavior in the data. In a regression of the normalized expectations of Players 1, 2 and 3 in the private treatment against a Player 2 dummy and a Player 3 dummy (fourth column of Table 1), the two dummy variables are not significantly different ($P = 0.883$ in both the OLS and fractional logit regressions).[33] This result suggests that the actions of Player 3 do not optimally respond to those of Player 2 on average. One possibility is that because Player 3 faces a more difficult information processing task than Player 2 in the private treatment, her behavior is further away from best responding than that of Player 2. As we argue in Section 4.2, this is not the case. When behavior is analyzed at the level of individual subjects, both Player 2 and Player 3 show substantial deviations from best-responding.

Before analyzing behavior in more detail, we highlight an additional stylized fact: A large fraction of subjects' actions equal the initial prior belief of 0.5. Figure 4 reports the proportion of Player 2's and Player 3's actions being exactly equal to the initial prior of 0.5 as a function

[33] A regression that includes a period variable and interactions with the player dummies shows that the coefficient on the period variable is smaller for both Player 2 and Player 3 compared to Player 1 although only significantly so for Player 3 ($P = 0.133$ for Player 2, and $P < 0.05$ for Player 3), while the trend is insignificantly different between Player 2 and Player 3 ($P = 0.3674$). This suggests slower updating for higher-order players compared to Player 1 and similar updating speeds for Player 2 and Player 3.



(a) 0.5 actions



(b) 0 or 1 actions

Figure 4: The incidence of actions equal to the period one prior of 0.5, or either 0 or 1, for Player 2 and Player 3.

of period number. Remarkably, we find that around 15.83% of actions in the public treatment and 33.33% of the actions in the private treatment are exactly equal to the initial prior in period 30 (at the end of the round).[34] One possibility is that these non-updated actions are driven by subjects with high information processing costs.[35] If the cost of extracting information from the received signal exceeds the benefit of increasing the probability of attaining the prize, the subject's optimal response could be to set her posterior equal to the prior. That the fraction of non-updated higher-order expectations is higher in the private than the public treatment might reflect the fact that the information processing problem is more difficult (and the costs of extracting information are higher) in the private case. On the other hand, as we show below, the higher-order players who change their actions less in response to information in the private treatment are closer to best-responding than the higher-order players who are more responsive to information. It follows that the increased fraction of non-updated higher-order expectations in the private treatment could also reflect a rational response. We discuss these issues in more detail in Section 5.

To study heterogeneity in behavior, we estimate the following model at the subject level:[36]

$$ObsTruth_{it} = (1 - \beta_i^C) EqmTruth_{it} + \beta_i^C 0.5 + \epsilon_{it}. \tag{6}$$

$ObsTruth_{it}$ is i's normalized action (i.e., expectations), which equals her action when the urn is orange and one minus her action when the urn is purple, where i denotes the subject ID, and t denotes the observation. Similarly, $EqmTruth_{it}$ is the normalized equilibrium prediction. The parameter $\beta_i^C \in [0,1]$ captures subject i's *degree of conservatism*.[37] This specification is motivated by the observation in Figure 4 that a large fraction of higher-order expectations are exactly equal to the period one prior of 0.5.[38] The model is flexible enough to capture subjects who are unresponsive to information ($\beta_i^C \approx 1$), subjects who choose equilibrium actions ($\beta_i^C \approx 0$), as well as subjects who may respond to information but less so than predicted by equilibrium.

[34] The treatment effect is significant with $P < 0.01$ according to a t-test with subject-clustered standard errors. Note that only a few subjects always respond with a 0.5 action in every period (8/249 with 7 subjects from the private treatment). While the number of non-updated expectations is high, Mobius et al. (2014) also find that approximately 16% of their subjects did not update their beliefs over the course of their experiment ($\approx 169/1058$ subjects).

[35] Recall that the binarized scoring rule is immune to risk-aversion. Thus, hedging considerations cannot directly explain the large fraction of 0.5 actions observed throughout the experiment.

[36] For this analysis, we perform a nonlinear least squares estimation.

[37] Call an action *conservative* if it is closer to the initial prior of 0.5 than the equilibrium prediction. Our notion of conservatism follows Edwards (1968) who introduced it to refer to belief distortions toward a uniform prior.

[38] In Appendix C.1, we estimate a more general model than (6) in which 0.5 is replaced by a subject-specific constant which is also estimated from the data. We find qualitatively and quantitatively similar results regarding the estimated degrees of conservatism in our dataset.

In addition, this model can be applied to subjects across all of our treatments without imposing strong structural assumptions about the way in which each subject processes information.

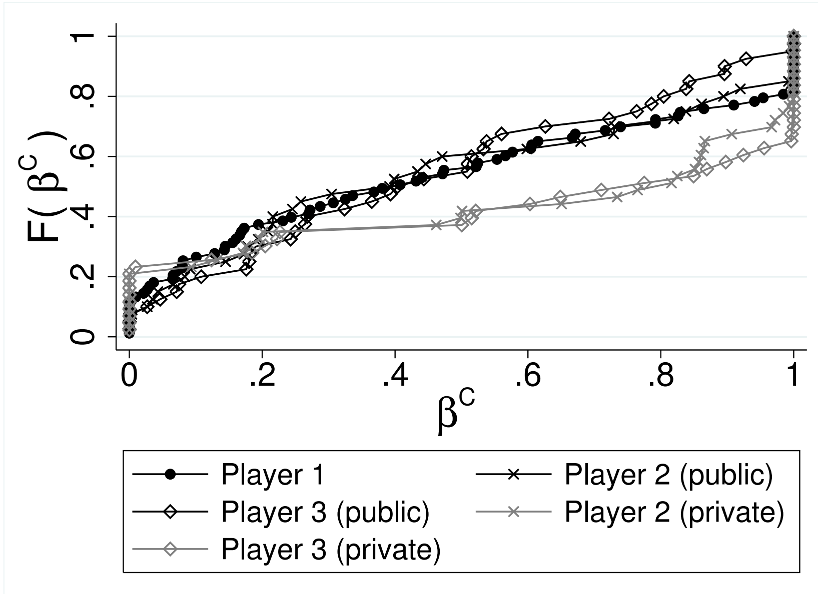


Figure 5: The distribution of the β_i^C parameter, estimated at the subject level using the model $ObsTruth_{it} = (1 - \beta_i^C)EqmTruth_{it} + \beta_i^C 0.5 + \epsilon_{it}$.

The estimated distributions of β_i^C are shown in Figure 5.[39] Using a Wilcoxon rank-sum test, we find no pairwise differences in β_i^C for Player 1, 2, and 3 in the public treatment (smallest $P = 0.8688$).[40] This is consistent with the already noted fact that the average beliefs of Player 2 and Player 3 in the public treatment are no closer to the prior than those of Player 1 (Figure 3). By contrast, the estimated distributions of β_i^C are markedly different for the higher-order players in the private treatment compared to all others. Specifically, β_i^C is significantly higher for Player 2 and Player 3 in the private treatment than for all other player types across the two treatments ($P < 0.05$ according to a Wilcoxon rank-sum test). This suggests that higher-order players with private signals were more conservative than either lower-order players or higher-order players with public signals ($P < 0.05$ and $P < 0.1$, respectively, according a Wilcoxon rank-sum test).

We can use the estimates of the β_i^C parameter to classify subjects into information processing types as shown in Table 2. Calling a subject More Conservative if she is better described by

[39]Recall that Player 1's task is identical in both treatments, as this player is not informed of the game Players 2 and 3 are playing. For this reason, we pool together Player 1's observed choices in both the public and the private treatment. A regression of the normalized action of Player 1 against a dummy equal to one for the treatment with private signals produces a statistically insignificant coefficient ($P = 0.3595$), we also obtain a similar result from a fractional logit regression ($P = 0.354$).

[40]The results of t-tests are qualitatively similar.

the conservative benchmark than the equilibrium one ($\beta_i^C \geq 0.5$) and More Bayesian if she is better described by the equilibrium benchmark ($\beta_i^C < 0.5$), we find that 44.58% of subjects in the role of Player 1 are More Conservative, highlighting the sizable deviation of first-order beliefs in the experiment from Bayesianism.[41] We also find that 43.75% of the higher-order subjects are More Conservative in the public treatment, while 62.79% are More Conservative in the private treatment ($P < 0.05$ according to a Fisher's exact test). We highlight the discussion so far in the following:

Result 2 (Conservatism). *Higher-order expectations are more conservative with private than public signals.*

	Player 1, all	Player 2, public	Player 3, public	Player 2, private	Player 3, private	All
More Bayesian	55.42%	60%	52.50%	39.53%	34.88%	49.40%
More Conservative	44.58%	40%	47.50%	60.47%	65.12%	50.60%

Table 2: Information processing types.

Notice that Result 2 is distinct from Result 1 as conservatism is defined in contrast to Bayesianism in our individual-level analysis. While higher-order expectations are predicted to evolve more slowly in equilibrium when signals are private rather than public, equilibrium expectations *do* incorporate all the available information. By contrast, conservative players show a lower degree of responsiveness to new information and thus their expectations do not fully incorporate their information. Result 2 can also be seen in Figure 6, which restricts attention to subjects in the roles of Player 2 and Player 3 and plots the fraction of reported higher-order expectations in the $[0.25, 0.75]$ range, taking observations from the private and public treatments separately. As a benchmark, it also reports the fraction of equilibrium expectations in this range. In period 30, 2.5% of the equilibrium higher-order expectations in the public treatment and 6.2% in the private treatment are in the $[0.25, 0.75]$ range. In the data, 29.6% of the expectations are in this range in the public treatment and 52.3% in the private treatment. That the deviation from the equilibrium prediction is larger in the private treatment reflects a larger degree of conservatism.

[41]In Appendix C.4, we also investigate whether subject types are stable within rounds, that is, whether a subject can be consistently classified based on the first and last 15 periods of every round. We find that most subjects have stable types.

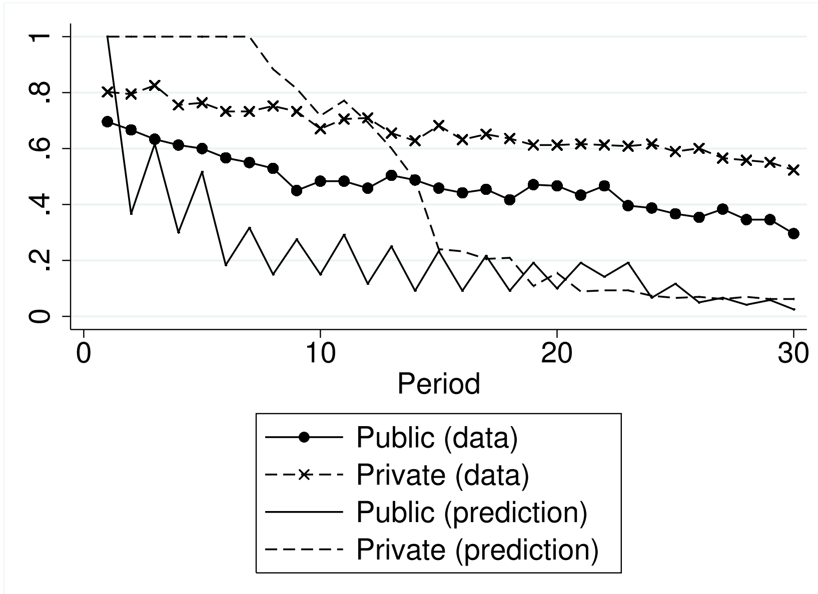


Figure 6: Fraction of higher-order expectations between 0.25 and 0.75. The predicted frequencies are constructed by replacing the actual actions in the experiment with the equilibrium actions.

4.1 Theory of mind

Part of the observed conservatism in higher-order expectations might reflect *sophisticated* behavior, where Player 2 and Player 3's actions are best-responses to the empirical distribution of actions of subjects in the lower roles. For instance, the first-order beliefs of Player 1 are biased toward the prior relative to the Bayesian prediction. The hypothesis that Player 2 is responding optimally to the average beliefs of Player 1 is in principle consistent with the results reported in Figure 3 and Figure 6. To disentangle equilibrium, conservative, and sophisticated behavior, we regress higher-order expectations on two variables: Eqm_{it} and $Soph_{it}$. Eqm_{it} is the equilibrium prediction, while the $Soph_{it}$ variable captures the optimal expectations a subject should have given the actual behavior of the subjects in lower-order roles.[42] If Player i observes a history h in period t of the public treatment, $Soph_{it}$ is computed as the average of the reported actions of Player $i-1$ given the same history.[43] In the private treatment, the average has to take into account the possibility that Player i and Player $i-1$ might observe different histories and is

[42]A subject that reports an expectation close to the optimal one is correctly anticipating the distribution of subjects' actions in the population. This is close in spirit to the "rational expectations" types in Stahl and Wilson (1995) and the "sophisticated" types of Costa-Gomes et al. (2001) and Costa-Gomes and Crawford (2006).

[43]In the case of Player 2, Player 1's reported beliefs in both the public and the private treatment are used in the average in order to obtain a more precise estimate.

therefore computed as follows:

$$P(Orange|h)\sum_{h'}a(h')P(h'|Orange)+P(Purple|h)\sum_{h'}a(h')P(h'|Purple). \quad (7)$$

Following a history h, Player i forms a posterior belief $P(Orange|h)$ about the urn being orange and a posterior belief $P(Purple|h) = 1 - P(Orange|h)$ about the urn being purple. Probabilities $P(h'|Orange)$ of Player $i-1$ observing different histories h' if the urn is orange and probabilities $P(h'|Purple)$ if the urn is purple are computed from the binomial distribution. $a(h')$ is the average action chosen by Player $i-1$ when she observes history h'. While we do not directly elicit the first-order beliefs $P(Orange|h)$ and $P(Purple|h)$ from Player 2 and Player 3, we make the natural assumption in our estimation exercise that the first-order beliefs of these players are identical to those reported by the average Player 1.[44]

	Dependent variable: Observed expectations				
	(1) Player 2, public	(2) Player 3, public	(3) Player 2, private	(4) Player 3, private	(5) Players 2 and 3, all
Eqm_{it}	0.412****	0.497****	0.396****	0.462****	0.419****
	(0.0707)	(0.0727)	(0.0966)	(0.104)	(0.0553)
$Soph_{it}$	0.229**	0.107	0.0749	-0.0693	0.194*
	(0.0972)	(0.0927)	(0.229)	(0.260)	(0.0996)
Constant	0.203****	0.186****	0.266***	0.303**	0.199****
	(0.0453)	(0.0366)	(0.0918)	(0.116)	(0.0316)
Observations	3600	3600	3870	3870	14940

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 3: Determinants of higher-order expectations of Player 2 and Player 3 in all treatments.

The results are reported in Table 3.[45] The coefficient on the $Soph_{it}$ variable is significant only in the case of Player 2 in the public treatment ($P < 0.05$), and when the data is pooled for both treatments and player types, it is only significant at a 10% level (fifth column of Table 3). Thus, the effect of the $Soph_{it}$ variable on the behavior of Player 2 and Player 3 is neither strong nor reliable. On the other hand, the coefficient on the Eqm_{it} variable is sizable in magnitude and significant in every specification in Table 3 ($P < 0.001$ in every case). I.e., higher-order

[44]More precisely, we use the empirical frequencies of beliefs of subjects in the role of Player 1, conditional on the same observed history, as a proxy. We believe this assumption to be reasonable because subjects' first-order beliefs about the color of the urn should not be affected by their role in the game.

[45]Table 7 in the appendix reports the results of a robustness check using fractional logit regression as the dependent variable is between zero and one. We find qualitatively and quantitatively similar results.

expectations can be better modeled by equilibrium than the average actions in the population.[46] The observed failure of theory of mind is further confirmed by subject-level analysis, which shows that only a marginal fraction of the subjects in our data can be classified as sophisticated.[47] This finding is summarized in the following result:

Result 3 (Failure of Theory of Mind). *Subjects fail to form correct expectations about their opponents' expectations.*

While Result 3 is consistent with previous literature (e.g., Costa-Gomes and Crawford, 2006), it is somewhat surprising in our framework given that subjects observe a significant amount of information about the underlying state of the world toward the end of each round. In the next section, we investigate whether, and if so how, higher-order expectations evolve over time.

4.2 The dynamics of higher-order expectations

To investigate whether higher-order expectations become more precise as information accumulates, we analyze the absolute distance $|Obs_{it} - Soph_{it}|$ between a subject's observed action and the optimal action given her observed history. If information were to mitigate the inaccuracy of higher-order expectations over time, we should observe that this measure decreases with the number of periods. Figure 7 suggests that this is not the case. To the contrary, we observe an overall positive trend ($P < 0.001$) which is significant in both treatments ($P < 0.05$ with public and $P < 0.001$ with private signals). In the private treatment, the trend is driven by both Player 2 and Player 3 ($P < 0.001$ in both cases), while in the public treatment it is driven by Player 2 ($P < 0.01$, while $P = 0.696$ for Player 3).[48] We highlight these results as follows:[49]

[46] Let Obs_{it} denote the observed action taken by Player i in period t. Another way to see the failure of theory of mind in the aggregate data for higher-order players is to do an OLS regression of $Obs_{it} - Soph_{it}$, which is the signed mistake made by subject i in period t, against a Player 3 dummy, a treatment dummy, and an interaction term. The regression shows that the average mistake is significantly different from zero for all higher-order players (largest $P < 0.1$) except for Player 3 in the private treatment ($P = 0.584$)./

Incorrect higher-order expectations have also been found recently by Danz et al. (2017), who test the predictions of projection equilibrium developed by Madarász (2014) and find corroborating evidence. Contrast these findings with the theoretical assumptions in Eyster and Rabin (2005), where each player correctly predicts the distribution of her opponent's actions, and Jehiel (2005) and Jehiel and Koessler (2008), where beliefs are also correct on average.

[47] See Appendix C.2 for details.

[48] All $P-$values in this paragraph are obtained from OLS regressions of $|Obs_{it} - Soph_{it}|$ against period number with standard errors clustered by subject which are reported in Table 8. In the last two columns of the table, we also report the results of a fractional logit regression which takes into account that the dependent variable lies in the $[0, 1]$ interval. The robustness checks confirm the results from the OLS regressions.

[49] In Section B of the Appendix, we perform additional analysis of the observed divergence in higher-order expectations. While we find that subjects' expectations are more accurate in later rounds, the trend is still positive within each round.

Result 4 (Divergence). *Higher-order expectations diverge over time.*

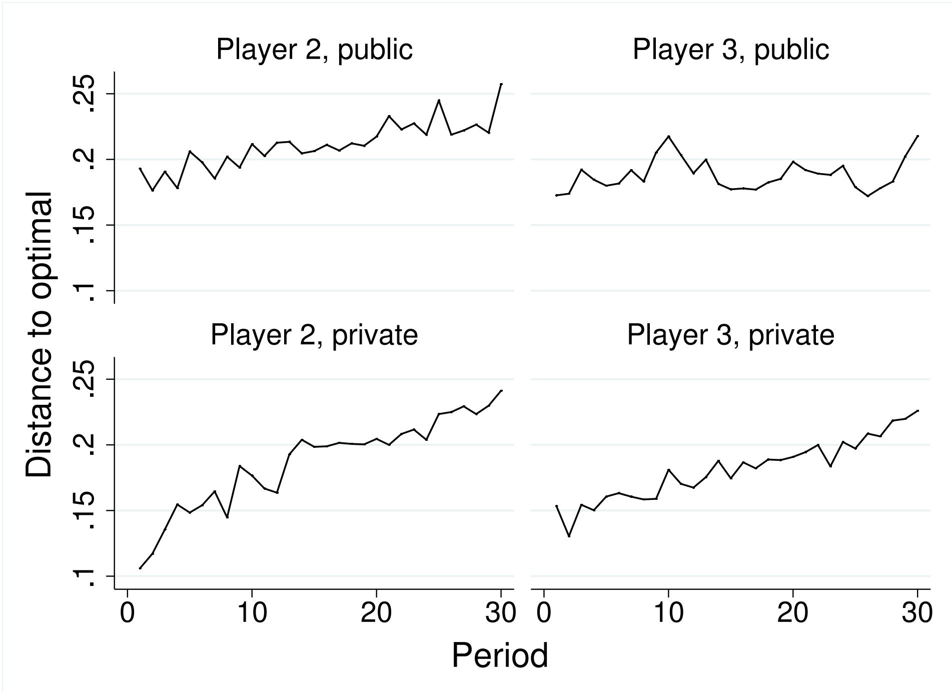


Figure 7: Divergence of higher-order expectations.

Table 4 shows the sign distribution of the coefficient associated with the period number used in the aggregate analysis when it is significant at a 1% level.[50] A positive coefficient means that the subject's expectations diverge over time from those of her intended target, while a negative coefficient implies some degree of convergence. Whenever the coefficient is statistically significantly different from zero, it is positive for most subjects. Furthermore, the rate of divergence is more likely to be positive with private than public signals ($P < 0.1$ in a Fisher's exact test)[51].

Considering Player 2 and Player 3 in each treatment separately, Figure 8 plots the relationship between the conservatism parameter β^C and the rate of divergence from the subject-level regressions. The correlation between β^C and the rate of divergence is positive and significant in the public treatment ($\rho = 0.4160$ with $P < 0.01$ for Player 2 and $\rho = 0.4958$ with $P < 0.01$ for Player 3) and negative and significant in the private treatment ($\rho = -0.4721$ with $P < 0.01$ for Player 2 and $\rho = -0.4655$ with $P < 0.01$ for Player 3). I.e., divergence of higher-order expectations is driven by More Conservative subjects in the public treatment and More Bayesian

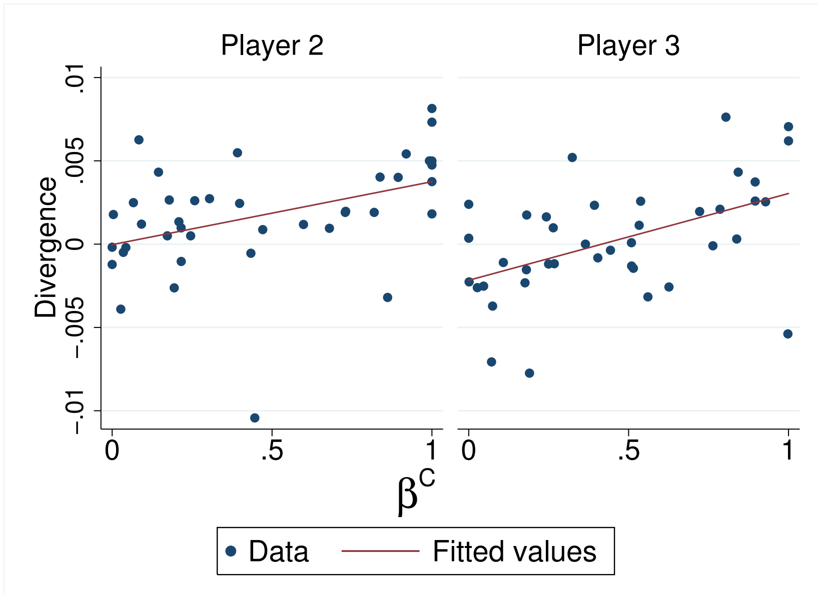
[50]That is, when the coefficient on the period number of the corresponding regression of $|Obs_{it} - Soph_{it}|$ against the period number (1-30) and a constant for a subject is significant at a 1% level.

[51]Keeping the insignificant coefficients, coded as 0, and performing a Fisher' exact test on the resulting 3×2 contingency table also leads to rejection of the null hypothesis of independence between columns ($P < 0.001$).

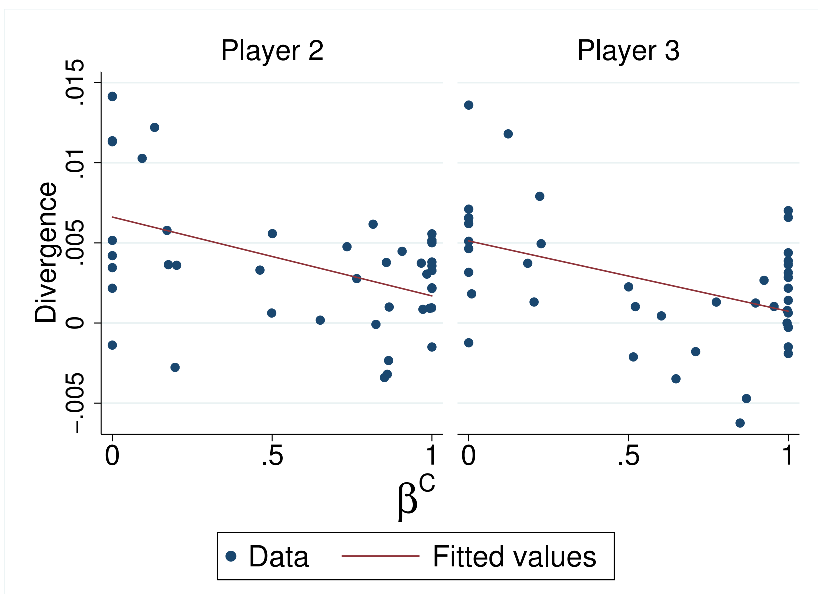
	(1) Player 2, public	(2) Player 3, public	(3) Player 2, private	(4) Player 3, private
Positive rate	22.50%	7.50%	51.16%	37.21%
Insignificant rate	72.50%	87.50%	48.84%	55.81%
Negative rate	5.00%	5.00%	0.00%	6.98%

Table 4: Sign of the rate of divergence by type and treatment when significant at a 1% level.

subjects in the private treatment. The results in Figures 7 and 8 can easily be reconciled with the failure of theory of mind highlighted in Result 3. In the public treatment, the expectations of More Conservative subjects grow worse over time because they evolve more slowly than their targets. In the private treatment, the expectations of More Bayesian subjects grow worse over time because they evolve faster than their targets.



(a) Public signals



(b) Private signals

Figure 8: Divergence of higher-order expectations by player type.

5 Discussion

One possible explanation for increased conservatism with private signals is the presence of cognitive attention costs. If a subject faces an attention cost in processing information, the evolution of expectations will be slower than predicted by equilibrium, with some subjects potentially avoiding information processing altogether. In particular, subjects in the higher-order roles face a more complex information processing task, which involves thinking about other players' processing behavior. Private signals can further increase the cognitive burden. A simple, albeit simplistic, way to check for this in our data is to look for differences in the number of updating decisions made by subjects across treatments. Table 9 in the appendix reports the results of a logit regression with the dependent variable being equal to one whenever a subject's action in the current period is different from her previous period's action against a dummy variable indicating whether the action was taken in the private treatment, dummies for each player role, and interactions. Although Player 2 in the private treatment is less likely to change her action in response to new information than Player 2 in the public treatment ($P < 0.05$), in line with the complexity story, there is no significant difference for Player 3 across treatments ($P = 0.871$). Furthermore, we find that Player 3 is more likely to change her action than Player 2 is when information is private, although the difference is only marginally significant ($P < 0.1$). This contrasts with the idea that increased complexity should lead subjects in the role of Player 3 to process information less often than Player 2 subjects given the relatively higher cognitive burden of their task.[52]

Another possibility is that some subjects treat private information as public and thus behave *as if* they were in the public treatment. The high complexity of Player 3's task might induce some subjects in that role to misinterpret the nature of information, which could explain the increased updating frequency. We formalize this idea through the concept of *information projection* proposed by Madarász (2012).[53] For a subject i in the role of Player 2, projection can only mean behaving as a Player 1 would but with i's information. For a subject i in the role of Player 3, projection can either mean that i behaves like a Player 2 but with i's information (thus, acknowledging that Player 1 has different information) or, in an even more extreme manner, behaving like a Player 1.

[52] We also find no difference in the probability of changing one's action between higher-order roles in the public treatment ($P = 0.5916$). In the private treatment, we find that higher-order players update more often than Player 1 does, although the difference is only significant between Player 1 and Player 3 ($P < 0.01$). This is due to Player 1 updating less often in the private compared to the public treatment, although only marginally so ($P < 0.1$).

[53] As in Madarász (2012), we interpret information projection as subjects exaggerating "the probability that the information content of their signals is reflected in the information available to others."

To explore whether subjects in the private treatment engage in information projection, define Eqm_{it}^{P1} as the equilibrium action of Player 1 given i's observed history in period t and Eqm_{it}^{P2} the equilibrium action of Player 2 given i's observed history. Table 5 regresses the observed actions of Player 2 in the private treatment against Eqm_{it} and Eqm_{it}^{P1} (first column) and the observed actions of Player 3 against Eqm_{it}, Eqm_{it}^{P1}, and Eqm_{it}^{P2} (fourth column). The Eqm_{it}^{P1} variable is not significant for Player 2 ($P = 0.6303$), when Eqm_{it} is controlled for. By contrast, the Eqm_{it} variable is not significant for Player 3 when Eqm_{it}^{P1} is controlled for ($P = 0.3063$), while Eqm_{it}^{P1} is significant at a 5% level. Controlling for the Eqm_{it}^{P2} variable does not change these results. This suggests that at least some of the subjects in the role of Player 3 treated private information as public while also behaving as a Player 1.[54] Appendix C.3 further shows that information projection can be seen at the level of individual subjects.

	Dependent variable: Observed expectations			
	(1)	(2)	(3)	(4)
	Player 2	Player 3	Player 3	Player 3
Eqm_{it}	0.362****	0.111	-0.389**	0.266
	(0.101)	(0.0718)	(0.172)	(0.257)
Eqm_{it}^{P1}	0.0417	0.220****		0.279**
	(0.0861)	(0.0566)		(0.114)
Eqm_{it}^{P2}			0.684****	-0.200
			(0.178)	(0.319)
Constant	0.295****	0.334****	0.353****	0.327****
	(0.0393)	(0.0408)	(0.0390)	(0.0425)
Observations	3870	3870	3870	3870

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 5: Analysis of information projection in the private treatment for subjects in higher-order roles.

Information projection can also explain the observed similarity in average expectations between Player 2 and Player 3 in the private treatment. While the increased shift toward conservatism, coupled with the rational response to shade down expectations, suggests that Player 3's beliefs should be lower than those of Player 2 on average, information projection leads to higher actions thus reducing the average gap between higher-order expectations.

[54] Table 10 in the appendix performs a robustness check using a fractional logit regression and finds very similar results.

6 Conclusion

Our results show that higher-order expectations evolve more slowly when information is private rather than public. While this is consistent with the equilibrium predictions, we also find that higher- and lower-order expectations diverge over time even as a substantial amount of information is accumulated. The failure of higher-order learning observed in our experiment has important implications for models of common learning. An event (like the realization of a state) may not become even approximate common knowledge, increasing the likelihood of coordination failure. Economic agents may fail to coordinate on an efficient coarse of action even if they individually learned the state of the world (e.g., the fundamentals of the economy). Particularly striking when information is public, this result is entirely due to the inability of agents to take into account the heterogeneity in information processing in the population.

Our results also have implications for macroeconomic models. For instance, in a Calvo model with incomplete information about nominal shocks, Angeletos and La'O (2009) show that knowledge about the evolution of first-order beliefs is insufficient to quantify the rate of price adjustment without taking into account the evolution of higher-order beliefs. In turn, higher-order beliefs affect firms' forecasts of other firms' equilibrium actions, which determine their own pricing choices. Our analysis shows that firms might have persistently incorrect beliefs about other firms' beliefs about the size of a nominal shock. More importantly, sluggish price adjustments could persist even when firms in the economy observe only publicly available information.

We conclude with a promising venue for future research. Cognitive ability affects the formation and evolution of players' higher-order expectations which ultimately drive behavior in strategic environments. While a recent literature has investigated how cognitive ability influences behavior in beauty contest games (Brañas-Garza et al., 2012; Burnham et al., 2009), or the speed of learning on how to play equilibrium (Gill and Prowse, 2016),[55] to the best of our knowledge no empirical or experimental work has yet investigated the more fundamental role of cognition in higher-order learning.[56] Our paradigm can be combined with the methodology of Gill and Prowse (2016) to address the question of how cognitive skills and other personality traits affect the evolution of players' beliefs about others.

[55] Agranov et al. (2012) find that a subject's observed cognitive level in a beauty contest depends on her expectation about the level of sophistication of her opponents. Similarly, in a study involving children between the age of 5 and 12, Fe and Gill (2018) show that older children respond to information about the cognitive ability of their opponents in a variant of the 11-20 game (Arad and Rubinstein, 2012).

[56] Oechssler et al. (2009) find that subjects with higher cognitive ability, measured by the Cognitive Reflection Test of Frederick (2005), display significantly less conservatism in (first-order) belief updating.

References

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.

Angeletos, G.-M., F. Collard, and H. Dellas (2018): "Quantifying Confidence," *Econometrica*, forthcoming.

Angeletos, G.-M. and J. La'O (2009): "Incomplete information, higher-order beliefs and price inertia," *Journal of Monetary Economics*, 56, S19–S37.

——— (2013): "Sentiments," *Econometrica*, 81, 739–779.

Arad, A. and A. Rubinstein (2012): "The 11-20 money request game: A level-k reasoning study," *American Economic Review*, 102, 3561–73.

Aumann, R. and A. Brandenburger (1995): "Epistemic conditions for Nash equilibrium," *Econometrica: Journal of the Econometric Society*, 1161–1180.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2018): "Incentives in experiments: A theoretical analysis," *Journal of Political Economy*, 126, 1472–1503.

Banerjee, S., R. Kaniel, and I. Kremer (2009): "Price drift as an outcome of differences in higher-order beliefs," *The Review of Financial Studies*, 22, 3707–3734.

Brañas-Garza, P., T. Garcia-Munoz, and R. H. González (2012): "Cognitive effort in the beauty contest game," *Journal of Economic Behavior & Organization*, 83, 254–260.

Burnham, T. C., D. Cesarini, M. Johannesson, P. Lichtenstein, and B. Wallace (2009): "Higher cognitive ability is associated with lower entries in a p-beauty contest," *Journal of Economic Behavior & Organization*, 72, 171–175.

Chen, Y.-C., A. Di Tillio, E. Faingold, and S. Xiong (2016): "Characterizing the strategic impact of misspecified beliefs," *The Review of Economic Studies*, 84, 1424–1471.

Coricelli, G. and R. Nagel (2009): "Neural correlates of depth of strategic reasoning in medial prefrontal cortex," *Proceedings of the National Academy of Sciences*, 106, 9163–9168.

Costa-Gomes, M., V. P. Crawford, and B. Broseta (2001): "Cognition and behavior in normal-form games: An experimental study," *Econometrica*, 69, 1193–1235.

Costa-Gomes, M. A. and V. P. Crawford (2006): "Cognition and behavior in two-person guessing games: An experimental study," *American Economic Review*, 96, 1737–1768.

Costa-Gomes, M. A. and G. Weizsäcker (2008): "Stated beliefs and play in normal-form games," *The Review of Economic Studies*, 75, 729–762.

Coutts, A. (2018): "Good news and bad news are still news: Experimental evidence on belief updating," *Experimental Economics*, 1–27.

Cripps, M. W., J. C. Ely, G. J. Mailath, and L. Samuelson (2008): "Common learning," *Econometrica*, 76, 909–933.

Danz, D., K. Madarász, and S. Wang (2017): "The Biases of Others: Anticipating Informational Projection in an Agency Setting," mimeo.

Edwards, W. (1968): "Conservatism in human information processing," in *Formal Representation of Human Judgement*, ed. by B. Kleinmuntz, New York, 17–52.

Eil, D. and J. M. Rao (2011): "The good news-bad news effect: asymmetric processing of objective information about yourself," *American Economic Journal: Microeconomics*, 3, 114–38.

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

Fe, E. and D. Gill (2018): "Cognitive skills and the development of strategic sophistication," mimeo.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178.

Frederick, S. (2005): "Cognitive reflection and decision making," *Journal of Economic perspectives*, 19, 25–42.

Gill, D. and V. Prowse (2016): "Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis," *Journal of Political Economy*, 124, 1619–1676.

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

——— (1992): "Testing Bayes rule and the representativeness heuristic: Some experimental evidence," *Journal of Economic Behavior & Organization*, 17, 31–57.

Harsanyi, J. C. (1967): "Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model," *Management science*, 14, 159–182.

Holt, C. A. and A. M. Smith (2009): "An update on Bayesian updating," *Journal of Economic Behavior & Organization*, 69, 125–134.

Hossain, T. and R. Okui (2013): "The binarized scoring rule," *Review of Economic Studies*, 80, 984–1001.

Huck, S. and G. Weizsäcker (2002): "Do players correctly estimate what others do? Evidence of conservatism in beliefs," *Journal of Economic Behavior & Organization*, 47, 71–85.

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.

Keynes, J. M. (1936): "The general theory of employment," *The quarterly journal of economics*, 51, 209–223.

Kneeland, T. (2015): "Identifying Higher-Order Rationality," *Econometrica*, 83, 2065–2079.

Kübler, D. and G. Weizsäcker (2004): "Limited depth of reasoning and failure of cascade formation in the laboratory," *The Review of Economic Studies*, 71, 425–441.

Lim, W. and S. Xiong (2016): "On Identifying Higher-Order Rationality," mimeo.

Madarász, K. (2012): "Information projection: Model and applications," *The Review of Economic Studies*, 79, 961–985.

——— (2014): "Projection equilibrium: Definition and applications to social investment and persuasion," CEPR D.P.

Mertens, J.-F. and S. Zamir (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 14, 1–29.

Mobius, M. M., M. Niederle, P. Niehaus, and T. S. Rosenblat (2014): "Managing self-confidence: Theory and experimental evidence," Tech. rep., National Bureau of Economic Research.

Morris, S. and H. S. Shin (2002): "Social value of public information," *american economic review*, 92, 1521–1534.

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

Nyarko, Y. and A. Schotter (2002): "An experimental study of belief learning using elicited beliefs," *Econometrica*, 70, 971–1005.

Oechssler, J., A. Roider, and P. W. Schmitz (2009): "Cognitive abilities and behavioral biases," *Journal of Economic Behavior & Organization*, 72, 147–152.

Papke, L. E. and J. M. Wooldridge (1996): "Econometric methods for fractional response variables with an application to 401 (k) plan participation rates," *Journal of applied econometrics*, 11, 619–632.

Phelps, E. (1983): "The trouble with "rational expectations" and the problem of inflation stabilization," in *Individual forecasting and aggregate outcomes*, ed. by F. R. and E. Phelps, Cambridge: Cambridge University Press.

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

——— (1995): "On players' models of other players: Theory and experimental evidence," *Games and Economic Behavior*, 10, 218–254.

Tversky, A. and D. Kahneman (1971): "Belief in the law of small numbers." *Psychological bulletin*, 76, 105.

——— (1974): "Judgment under uncertainty: Heuristics and biases," *science*, 185, 1124–1131.

Weizsäcker, G. (2003): "Ignoring the rationality of others: evidence from experimental normal-form games," *Games and Economic Behavior*, 44, 145–171.

Williams, D. (1991): *Probability with Martingales*, Cambridge University Press.

Woodford, M. (2001): "Imperfect common knowledge and the effects of monetary policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: in Honor of Edmund S. Phelps*, Princeton: Princeton University Press.

Appendix

A Omitted Tables and Figures

	Dependent variable: Normalized observed expectations			
	(1)	(2)	(3)	(4)
	All	Players 2 and 3 (private and public)	Players 1, 2 and 3 public	Players 1, 2 and 3 private
Bayesian prediction	2.2247****			
	(0.1581)			
Private		-0.2810***		
		(0.0959)		
Player 2			0.1161	-0.3024*
			(0.1624)	(0.1567)
Player 3			0.1505	-0.3211**
			(0.1603)	(0.1509)
Constant	-1.1471****	0.6910****	0.5577****	0.7218****
	(0.0916)	(0.0717)	(0.1255)	(0.1247)
Observations	22410	14940	10800	11610

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 6: Analysis of average normalized observed expectations using a fractional logit regression.

	Dependent variable: Observed expectations				
	(1)	(2)	(3)	(4)	(5)
	Player 2, public	Player 3, public	Player 2, private	Player 3, private	Players 2 and 3, all
Eqm_{it}	1.7067****	2.1113****	1.6201****	1.9012****	1.7016****
	(0.3212)	(0.3614)	(0.4118)	(0.4470)	(0.2528)
$Soph_{it}$	1.2153**	0.6328	0.4033	-0.2597	1.0844**
	(0.5092)	(0.4886)	(1.0091)	(1.1045)	(0.4979)
Constant	-1.3444****	-1.4343****	-1.0032**	-0.8255*	-1.3649****
	(0.2486)	(0.2049)	(0.4156)	(0.0.4996)	(0.1689)
Observations	3600	3600	3870	3870	14940

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 7: Determinants of expectations of Player 2 and Player 3 in all treatments using a fractional logit regression.

	OLS		Fractional Logit	
	Public Abs distance to optimal	Private Abs distance to optimal	Public Abs distance to optimal	Private Abs distance to optimal
Period	0.0018***	0.0037****	0.0104***	0.0246****
	(0.0005)	(0.0007)	(0.0031)	(0.0043)
Player 3	0.0012	0.0131	0.0023	0.0885
	(0.0209)	(0.0236)	(0.1362)	(0.1865)
Period × Player 3	-0.0015**	-0.00012	-0.0091*	-0.0077
	(0.0008)	(0.0009)	(0.0047)	(0.0059)
Constant	0.1839****	0.1296****	-1.4844****	-1.8643****
	(0.0138)	(0.0165)	(0.0890)	(0.1320)
Observations	7200	7740	7200	7740

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 8: Divergence by type. The first two columns contain the results of an OLS regression. The last two columns report the results of a fractional logit regression.

	Dependent variable: 1 if subject changed action, 0 otherwise
Player 1 (public)	0.4872****
	(0.0470)
Player 1 (private)	0.3695****
	(0.0463)
Player 2 (public)	0.5858****
	(0.0433)
Player 2 (private)	0.4395****
	(0.0495)
Player 3 (public)	0.5525****
	(0.0495)
Player 3 (private)	0.5633****
	(0.0495)
Observations	22410

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 9: Analysis of subjects' probability of changing one's action from period t to period $t + 1$ by type and treatment using a logit regression.

	Dependent variable: Observed expectations			
	(1)	(2)	(3)	(4)
	Player 2	Player 3	Player 3	Player 3
Eqm_{it}	1.5667****	0.4913	-1.5290**	1.1466
	(0.4399)	(0.3100)	(0.7091)	(1.0803)
Eqm_{it}^{P1}	0.1226	0.8845****		1.1301**
	(0.3405)	(0.2299)		(0.4651)
Eqm_{it}^{P2}			2.7551****	-0.8420
			(0.7314)	(1.3131)
Constant	-0.8569****	-0.6870****	-0.6125****	-0.7164****
	(0.1780)	(0.1777)	(0.1699)	(0.1860)
Observations	3870	3870	3870	3870

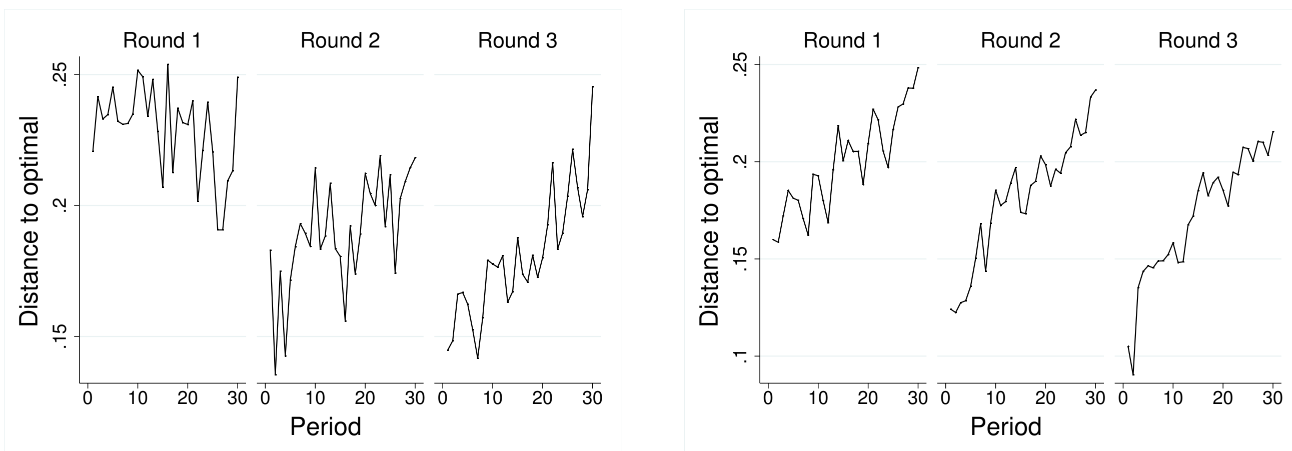
Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 10: Analysis of information projection in the private treatment for subjects in higher-order roles using a fractional logit regression.

B Additional Analysis of Divergence

Figure 9 shows that Result 4 is robust to across-round learning. While the rate of divergence does not significantly vary across rounds in the treatment with private signals, it is significantly higher in later rounds when the signals are public ($P < 0.05$ in the public treatment, smallest $P = 0.1726$ in the private treatment; first and second column of Table 11). Also, despite the fact that subject's expectations are more accurate in later rounds (largest $P < 0.001$ in the public treatment, and largest $P < 0.05$ in the private treatment), the rate of divergence is still significantly positive within each later round in both treatments (largest $P < 0.01$).[57]

Finally, Figure 10 reports the average divergence in the public treatment using the actions of matched partners rather than the empirical averages. As it can clearly be seen, the average absolute distance shows a similar pattern as observed in Figure 7.



(a) Public signals

(b) Private signals

Figure 9: Divergence of higher-order expectations by round.

[57]The third and fourth columns of Table 11 report the results of a fractional logit regression which confirms the results of the OLS analysis.

	Absolute distance to optimal			
	OLS		Fractional Logit	
	Public	Private	Public	Private
Period	-0.0008	0.0025****	-0.0047	0.0158****
	(0.0006)	(0.0007)	(0.0034)	(0.0040)
Round 2	-0.0748****	-0.0326**	-0.4603****	-0.2426**
	(0.0141)	(0.0134)	(0.0855)	(0.1013)
Round 3	-0.0967****	-0.0410***	-0.6169****	-0.3168***
	(0.0155)	(0.0133)	(0.0979)	(0.1063)
Period × Round 2	0.0023***	0.0009	0.0142***	0.0074*
	(0.0007)	(0.0007)	(0.0043)	(0.0044)
Period × Round 3	0.0031****	0.00085	0.0202****	0.0081
	(0.0008)	0.0007	(0.0047)	0.0050
Constant	0.2416****	0.1606****	-1.1427****	-1.6392****
	(0.0151)	(0.0127)	(0.0827)	(0.0882)
Observations	7200	7740	7200	7740

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 11: Divergence by round. For the last two columns, we report the average marginal effects using a fractional probit regression.

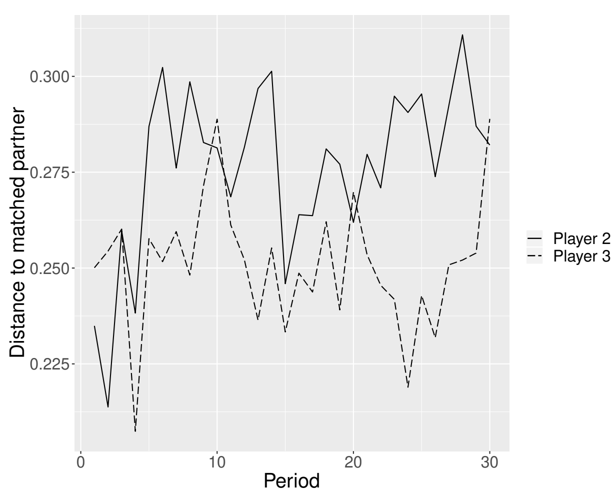


Figure 10: Divergence measured using absolute distance from the action of the matched partner in the public treatment.

C Additional Analysis of Heterogeneity

C.1 Heterogeneity with individual-specific constant

In this section, we generalize the model of equation (6) by removing the assumption that the constant is equal to 0.5 and the same for all subjects. We estimate the following model at the subject level:

$$ObsTruth_{it} = (1 - \beta_i^C) EqmTruth_{it} + \beta_i^C c_i + \epsilon_{it}. \quad (8)$$

$ObsTruth_{it}$ is i's normalized action which equals her action when the urn is orange and one minus her action when the urn is purple, where i denotes the subject ID, and t denotes the observation. Similarly, $EqmTruth_{it}$ is the normalized equilibrium prediction. The parameter $\beta_i^C \in [0, 1]$ captures subject i's degree of conservatism, and c_i is a subject-specific constant which is also estimated from the data.

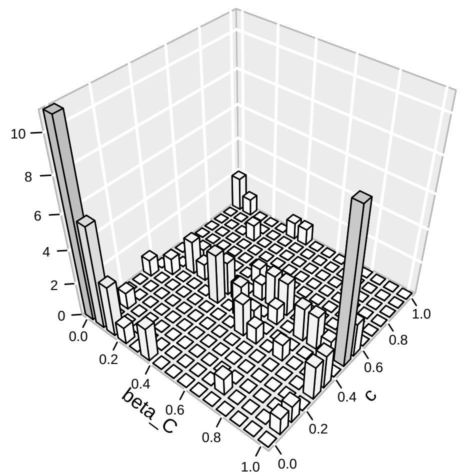
The estimated distribution of (β_i^C, c_i) is shown in Figure 11.[58] There is a large fraction of subjects with $\beta_i^C \approx 1$ and $c_i \approx 0.5$: these are subjects that update their beliefs little and often report the initial prior. There is also a large fraction of subjects with $\beta_i^C \approx 0$: these are subjects' whose beliefs are closer to the equilibrium benchmark. The mass around $c_i = 0$ and $c_i = 1$ reflects overconfident behavior where a subject starts taking an action equal to zero or one, even if the equilibrium action would in the $(0, 1)$ range. While the constant is undefined for a subject playing according to equilibrium, the mass around $c_i = 0$ shows some continuity in behavior as subjects that are closer to being equilibrium players also have a progressively lower constant as their behavior is less conservative and more responsive to information. However, 12.22% of all subjects in the public treatment and 9.3% of all subjects in the private treatment take zero-one actions in period 1 across all three rounds.

Using a Wilcoxon rank-sum test, we find no pairwise differences in β_i^C for Player 1, 2, and 3 in the public treatment (smallest $P = 0.8142$).[59] This is consistent with the fact that, overall, the actions of Player 2 and Player 3 are no closer to the initial prior belief than those of Player 1 in the public treatment (Figure 3).

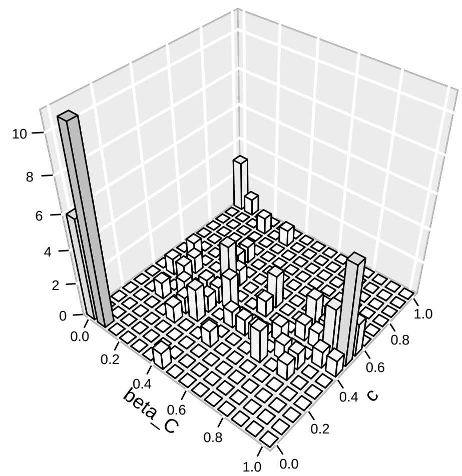
Figure 11 also shows that the estimated distributions of (β_i^C, c_i) are markedly different for higher-order subjects in the private treatment compared to all other subjects. The β_i^C parameter is significantly higher for Player 2 and Player 3 in the private treatment than for all other player

[58] Figure 12 shows the histograms for the β_i^C and c_i parameters separately.

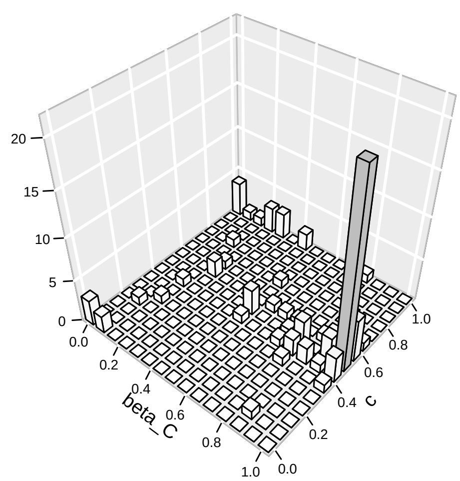
[59] The analysis uses all observations from subjects in the role of Player 1. A t-test with robust standard errors also finds no significant difference between the β^C parameter of subjects in the role of Player 1 compared to higher order players (smallest $P = 0.6676$). Using a Wilcoxon rank-sum test, we find no pairwise differences in c_i for Player 1, 2, and 3 in the public treatment (smallest $P = 0.5207$). Figure 13 shows the empirical CDFs for each parameter.



(a) Player 1, all

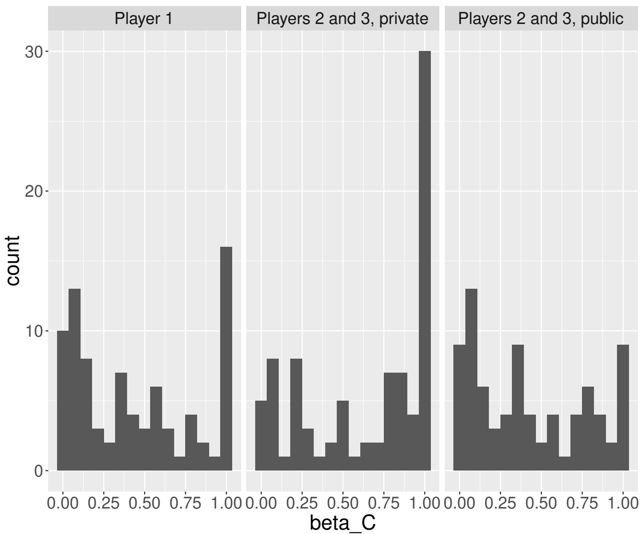


(b) Player 2 and 3, public

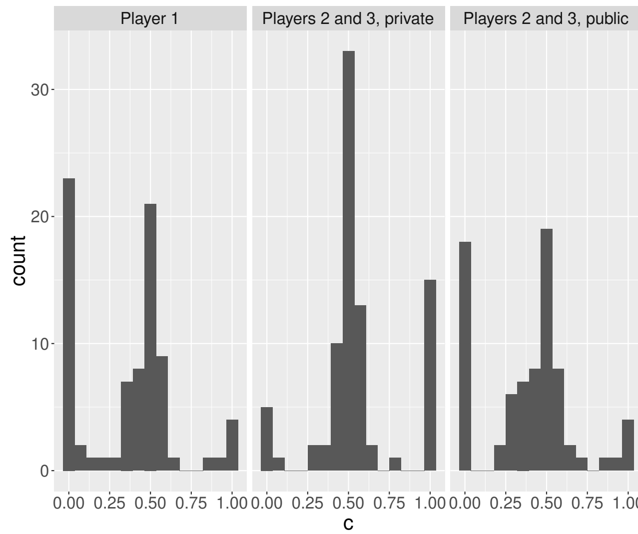


(c) Player 2 and 3, private

Figure 11: The distribution of the β_i^C and c_i parameters, estimated at the subject level using the model $ObsTruth_{it} = (1 - \beta_i^C)EqmTruth_{it} + \beta_i^C c_i + \epsilon_{it}$.

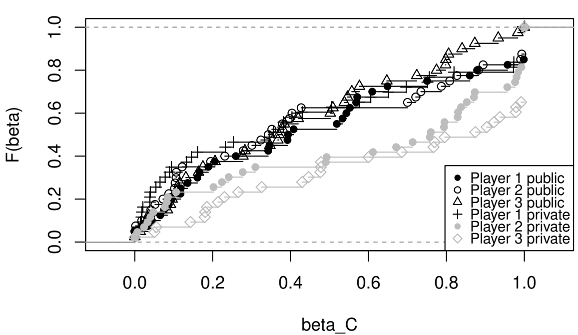


(a)

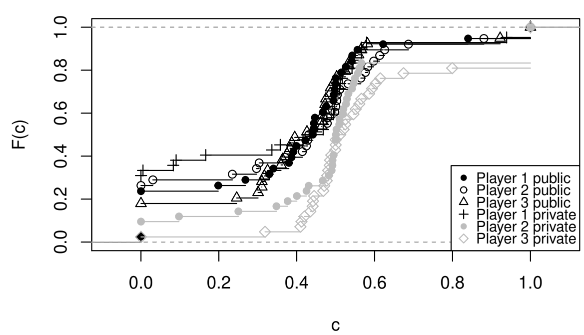


(b)

Figure 12: The distribution of the β_i^C and c_i parameters estimated at the subject level using the model $ObsTruth_{it} = (1 - \beta_i^C)EqmTruth_{it} + \beta_i^C c_i + \epsilon_{it}$.



(a)



(b)

Figure 13: The CDFs of the β_i^C and c_i parameters estimated at the subject level using the model $ObsTruth_{it} = (1 - \beta_i^C)EqmTruth_{it} + \beta_i^C c_i + \epsilon_{it}$.

	Player 1, all	Player 2, public	Player 3, public	Player 2, private	Player 3, private	All
More Bayesian	56.63%	62.50%	57.50%	39.53%	37.21%	51.41%
More Conservative	43.37%	37.50%	42.50%	60.47%	62.79%	48.59%

Table 12: Information processing types.

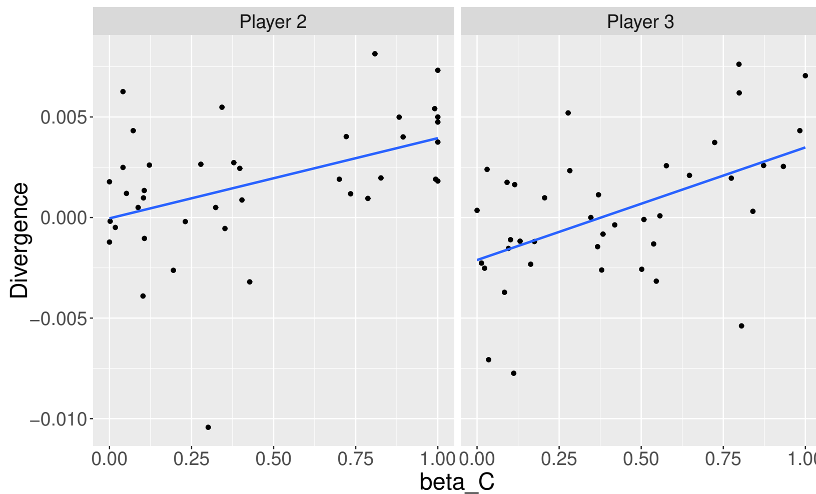
types across the two treatments ($P < 0.01$ according to a Wilcoxon rank-sum test).[60] This suggests that higher-order players with private signals were more conservative than either lower-order players or higher-order players with public signals.

We can use the estimates of the β_i^C parameter to classify subjects into information processing types as shown in Table 12. Calling a subject More Conservative if she is better described by the conservative benchmark than the equilibrium one ($\beta_i^C \geq 0.5$) and More Bayesian if she is better described by the equilibrium benchmark ($\beta_i^C < 0.5$), we find that 43.37% of subjects in the role of Player 1 are More Conservative, highlighting the sizable deviation of first-order beliefs in the experiment from Bayesianism. We also find that 40% of the higher-order subjects were More Conservative in the public treatment, while 61.63% were More Conservative in the private treatment ($P < 0.01$ according to a Fisher's exact test).

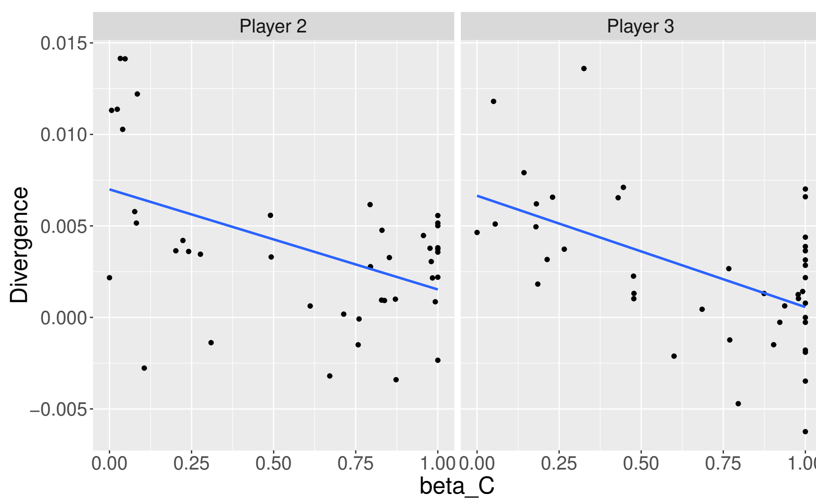
Figure 14 is the analog of Figure 8 in the main text and shows that the divergence of higher-order expectations is driven by different subject types for different treatments. Considering Player 2 and Player 3 in each treatment separately, the figure plots the relationship between the conservatism parameter β^C and a measure of divergence obtained for each subject as the coefficient on the period number in a regression of $|Obs_{it} - Soph_{it}|$ against the period number (1-30) and a constant. The correlation between β^C and the divergence measure is positive and significant in the public treatment ($\rho = 0.444$ with $P < 0.01$ for Player 2 and $\rho = 0.512$ with $P < 0.001$ for Player 3) and negative and significant in the private treatment ($\rho = -0.487$ with $P < 0.001$ for Player 2 and $\rho = -0.543$ with $P < 0.001$ for Player 3). I.e., divergence of higher-order expectationss is still driven by More Conservative subjects in the public treatment and More Bayesian subjects in the private treatment.

In summary, the results are qualitatively and quantitatively similar to those in the main text.

[60] A comparison of the distribution of the c_i parameter shows no difference between Player 2 and Player 3 ($P = 0.3744$) but a significant difference either player and Player 1 (largest $P < 0.01$).



(a) Public signals



(b) Private signals

Figure 14: Divergence of higher-order expectations by player type.

C.2 Individual-level analysis of sophisticated behavior

To check the robustness of our aggregate analysis of sophisticated behavior in Section 4.1 of the main text, we estimate the following model at the subject-round level for subjects in the roles of Player 2 and Player 3 of both treatments:

$$Obs_{irt} = (1 - \beta^C_{i,r})\big[(1 - \beta^S_{i,r})Eqm_{irt} + \beta^S_{i,r}Soph_{irt}\big] + \beta^C_{i,r}c_{i,r} + \epsilon_{irt}, \tag{9}$$

where the additional parameter $\beta^S_{i,r} \in [0, 1]$ measures the degree of sophistication of subject i in round r. The model differs from the one estimated in equation (8) by performing the estimation at the subject-round level. The reason is that normalizing optimal actions with respect to the true state of the world is tricky in the private treatment as it would implicitly assume that subjects know which urn has realized. Similarly, estimating the model without normalization would bundle together observations in which the signal generating process differs because the realized urn might differ across rounds. By contrast, estimating the model at the subject *and* round level only uses observations which are derived conditional on the same urn.[61] Calling a round More Sophisticated if it is better described by sophisticated behavior than by either conservative behavior or the equilibrium benchmark,[62] we find that only 11.45% of the rounds are More Sophisticated. Calling a subject More Sophisticated if all three of her rounds are More Sophisticated, we find that none of the subjects in the role of Player 2 and Player 3 are More Sophisticated, which provides a strong confirmation of the aggregate level analysis at the individual level. We also find that 22.29% of subjects are classified as More Bayesian, 26.51% are More Conservative, and 51.2% are inconsistent.

If we relax the requirement for classification and only require consistent behavior in two out of three rounds, we find that 2.55% (2/80) of subjects in the public treatment, 4.65% (4/86) in the private treatment, and overall only 3.62% (6/166) of the subjects can be classified as More Sophisticated. Of the remaining subjects, 40.96% are classified as More Bayesian, 46.99% as More Conservative, and 8.43% are inconsistent.

[61] While the estimated constant parameter could be different across rounds for the same subject, the estimated $\beta^C_{i,r}$ will be comparable across rounds.

[62] That is, if $\beta^S_{i,r} > 1 - \beta^S_{i,r}$ and $(1 - \beta^C_{i,r})\beta^S_{i,r} > \beta^C_{i,r}$.

C.3 Individual-level analysis of information projection

To study information projection at the individual level, we estimate the following model for each Player 2 in the private treatment:[63]

$$ObsTruth_{it} = (1-\beta_i^C)\left[(1-\beta_i^P)EqmTruth_{it} + \beta_i^P EqmTruth_{it}^{P1}\right] + \beta_i^C c_i + \epsilon_{it}. \tag{10}$$

The parameter $\beta^P \in [0,1]$ measures the degree of information projection. For Player 2, projection is equivalent to behaving like a Player 1 who possesses the same information as she does.

For each Player 3 in the private treatment, we instead estimate the following model:

$$ObsTruth_{it} = (1-\beta_i^C)\left[(1-\beta_i^P)EqmTruth_{it} + \beta_i^P\{(1-\beta_i^{ProjP2})EqmTruth_{it}^{P1} + \beta_i^{ProjP2}EqmTruth_{it}^{P2}\right] + \beta_i^C c_i + \epsilon_{it}. \tag{11}$$

In this case, we allow for two types of projection: 1) a naive type which consists of behaving like a Player 1 would; and 2) a more sophisticated kind of projection by which a Player 3 behaves like a Player 2 with her same information. Panels (a) and (b) of Figure 15 show the histograms of the (β^C, β^P) parameters pair for Player 2 and Player 3, respectively. The figures suggest a noticeable degree of information projection by both player types. However, Figure 15.c finds little evidence for sophisticated projection by subjects in the role of Player 3 in line with the aggregate analysis in the fourth column of Table 5.

	Player 2, private	Player 3, private
More Bayesian	20.93%	9.30%
More Conservative	65.12%	65.12%
More Projecting	13.95%	16.28%
More Sophisticated Projecting		9.30%

Table 13: Typology for subjects in higher-order roles in the private treatment allowing for projection.

Finally, Table 13 reports the distribution of updating types. We classify a subject as More Projecting if $(1-\beta^C)\beta^P > \max\left\{(1-\beta^C)(1-\beta^P), \beta^C\right\}$ and as More Sophisticated Projecting if $(1-\beta^C)\beta^P\beta^{ProjP2} > \max\left\{(1-\beta^C)(1-\beta^P), (1-\beta^C)\beta^P(1-\beta^{ProjP2}), \beta^C\right\}$. We find that

[63] Given the analysis from the previous section, we avoid including a parameter for sophisticated behavior as only a marginal fraction of subjects show sophistication.

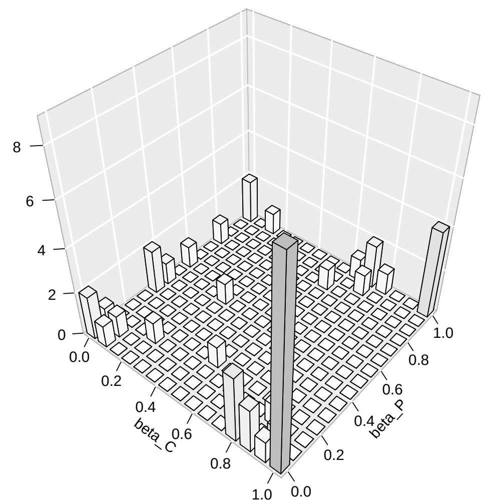
13.95% of Player 2 subjects in the private treatment are classified as projecting. By contrast, 25.58% of subjects in the role of Player 3 can be classified as projecting, but only about 9.30% seem to engage in more sophisticated projection. Thus, the subject level analysis confirms our results from the aggregate level analysis.[64]

C.4 Type stability

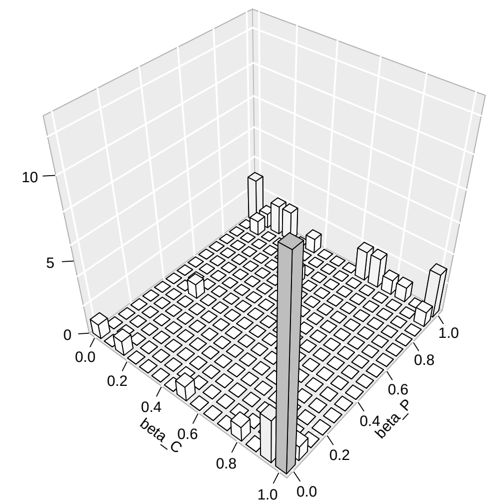
It is possible that subjects might process information differently in the early periods of each round compared to later periods. To check for this possibility, we reestimate subject types separately for the first and last 15 periods of a round. For subjects in the role of Player 1 (regardless of treatment) and Player 2 and Player 3 in the public treatment, we estimate the model in equation (8). For subjects in higher-order roles of the private treatment, we estimate the model in equation (10) for Player 2 and the model in equation (11) for Player 3.

We call a subject type *stable* if her type is the same in the first and second halves of a round, and unstable otherwise. We find that 74.3% of subjects in the public treatment (including all subjects in the role of Player 1 in both treatments) have stable types. We find that 67.9% of subjects in the role of Player 1 are stable, with proportions of 80% and 77.5% for subjects in the role of Player 2 and 3, respectively. For the higher-order roles in the private treatment, 69.77% of subjects in the role of Player 2 are stable, compared to 65.12% of subjects in the role of Player 3.

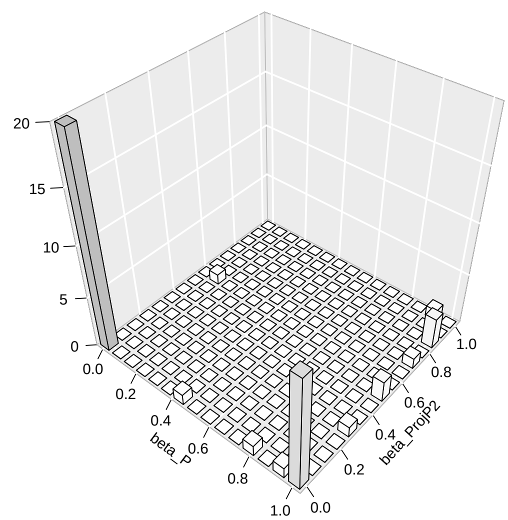
[64]The careful reader will notice that the proportion of subjects classified as More Conservative has slightly increased compared to the baseline type classification from Table 12. This is due to the fact that estimating a new parameter can change the comparison across coefficients thus leading some subjects to be classified in a different way. A robust classification should be immune to such phenomena and in fact this occurs only for a very small number of subjects.



(a) Player 2



(b) Player 3



(c) Player 3

Figure 15: The figure shows the distribution of the β^C and β^P parameters for Player 2 (panel a) and Player 3 (panel b) in the private treatment with projection, and the distribution of the β^P and β^{ProjP2} parameters for Player 3 (panel c).

Online Appendix

D Omitted Proofs

Proof of Proposition 1: For any $i \in I$ and $\mathbf{t}_i \in T_i$,

$$
\begin{aligned}
U_i(a_i, \mathbf{s}_{-i}; \mathbf{t}_i) &= \int_0^1 \sum_{\theta \in \Theta} p_i(\theta|\mathbf{t}_i)\hat{u}_i\left(a_i, s_{i-1}(\tau_{i-1}(\theta)), k\right) dk \\
&= E_k\left[E_\theta\left[u_i\left(\pi_i(a_i, s_{i-1}(\tau_{i-1}(\theta)), k)\right)|\mathbf{t}_i\right]\right] \\
&= E_k\left[E_\theta\left[u_i\left(R_0\mathbf{1}\left((a_i - s_{i-1}(\tau_{i-1}(\theta)))^2 > k\right) + R_1\mathbf{1}\left((a_i - s_{i-1}(\tau_{i-1}(\theta)))^2 \le k\right)\right)|\mathbf{t}_1\right]\right] \\
&= u_i\left(R_0\right) + \left(u_i\left(R_1\right) - u_i\left(R_0\right)\right)E_k\left[E_\theta\left[\mathbf{1}\left((a_i - s_{i-1}(\tau_{i-1}(\theta)))^2 \le k\right)|\mathbf{t}_1\right]\right] \\
&= u_i\left(R_0\right) + \left(u_i\left(R_1\right) - u_i\left(R_0\right)\right)E_\theta\left[E_k\left[\mathbf{1}\left((a_i - s_{i-1}(\tau_{i-1}(\theta)))^2 \le k\right)\right]|\mathbf{t}_1\right] \\
&= u_i\left(R_0\right) + \left(u_i\left(R_1\right) - u_i\left(R_0\right)\right) \\
&\qquad \times E_\theta\left[1 - (a_i - s_{i-1}(\tau_{i-1}(\theta)))^2|\mathbf{t}_i\right] \qquad (12)
\end{aligned}
$$

Maximizing Player i's expected utility with respect to a_i shows that the unique best-response to the profile of candidate equilibrium strategies of her opponents is $s_i^*(\mathbf{t}_i) = E_\theta\left[s_{i-1}(\tau_{i-1}(\theta))|\mathbf{t}_i\right]$.[65] Next, recall that by construction Player 1's payoff is only affected by Nature's moves, which immediately gives that $s_1^*(\mathbf{t}_1) = E_\theta[\theta|\mathbf{t}_1]$. Recursive substitution shows that $s_i^*(\mathbf{t}_i) = \bar{E}_\theta^i\left[\theta|\bar{\mathbf{t}}^i\right]$, which completes the proof.

Proof of Proposition 2: *a*) Let $\mathbf{t}^n$ denote the publicly observed signal in chain game n. By Proposition 1, $s_1^n(\mathbf{t}^n) = E_n[\theta|\mathbf{t}^n]$ and $s_i^n(\mathbf{t}^n) = E_n[s_{i-1}^n(\mathbf{t}^n)|\mathbf{t}^n]$, for any $i > 1$. Suppose by induction hypothesis that $s_{i-1}^n(\mathbf{t}^n) = E_n[\theta|\mathbf{t}^n]$, then $s_i^n(\mathbf{t}^n) = E_n[s_{i-1}^n(\mathbf{t}^n)|\mathbf{t}^n] = E_n[E_n[\theta|\mathbf{t}^n]|\mathbf{t}^n] = E_n[\theta|\mathbf{t}^n]$. Taking expectations conditional on θ gives that $E_n[s_i^n(\mathbf{t}^n)|\theta] = E_n[E_n[\theta|\mathbf{t}^n]|\theta] = E_n[s_1^n(\mathbf{t}^n)|\theta]$.

b) Let $\Theta = \{\theta, \neg\theta\}$ and $T_i = T = \{t_\theta, t_{\neg\theta}\}$. Let $p^n(\theta) = p_\theta$ denote the common prior belief that the state is θ in chain game n. Without loss of generality, we can define the signal generating technology as follows

$$
\text{Prob}(t = t_\theta|\theta) = q_\theta > \frac{1}{2} > \text{Prob}(t = t_\theta|\neg\theta) = q_{\neg\theta}. \qquad (13)
$$

Since we are assuming that players are Bayesian expected utility maximizers, a sufficient statistic

[65] The second-order condition is immediately satisfied under the assumption that $R_1 > R_0$ and that utility is monotonic in rewards.

for an arbitrary type $\mathbf{t}^n$ is the number of signals of type t_θ out of n total signals observed, which we denote by n_θ. From now on, we refer to n_θ instead of $\mathbf{t}^n$.

Let $p_i^n(\theta|n_\theta)$ denote Player i's posterior belief that the state is θ after having observed n_θ signals of type t_θ. By Bayes' rule, this probability equals

$$p_i^n(\theta|n_\theta) = \frac{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta}{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta + q_{\neg\theta}^{n_\theta}(1-q_{\neg\theta})^{n-n_\theta}p_{\neg\theta}}. \tag{14}$$

Thus, all players form the same belief about the state of the world after having observed the same set of signals. However, different players are likely to observe different signals.

We start by defining some notation. Let $r_n(n_\theta|\theta)$ denote the probability of observing n_θ signals of type t_θ out of n signals, conditional on state θ, which equals

$$r_n(n_\theta|\theta) = \begin{pmatrix} n \\ n_\theta \end{pmatrix} q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}, \tag{15}$$

from the binomial distribution. Next, define

$$R_n(n'_\theta|n_\theta) = \sum_{\theta' \in \{\theta, \neg\theta\}} p_i^n(\theta'|n_\theta) r_n(n'_\theta|\theta'), \tag{16}$$

which is the probability that Player i assigns to Player $i-1$ having observed n'_θ signals of type t_θ in n draws, conditional on Player i's own observation of n_θ such signals.

If Player $i-1$ observed a history (n, n'_θ), and Player i knew which history Player $i-1$ observed, then Player i would assign probability one to Player $i-1$ assigning probability $p_{i-1}^n(\theta|n'_\theta)$ to state θ. This is the case with public signals. However, here signals are private or, more precisely, conditionally independent. So, if Player i observes history (n, n_θ), she uses this information to update her belief about the possible histories that Player $i-1$ might have observed. For any $n'_\theta \in \{0, 1, ..., n\}$, Player i assigns probability $R_n(n'_\theta|n_\theta)$ to Player $i-1$ having observed n'_θ signals of type t_θ in n draws when she observed n_θ signals of type t_θ in n draws. From Proposition 1, we can thus write the equilibrium action of Player 2 as

$$s_2^n(\theta|n_\theta) = E_n\left[s_1^n(\theta|n'_\theta)|n_\theta\right] = E_n\left[p_1^n(\theta|n'_\theta)|n_\theta\right] = \sum_{n'_\theta=0}^{n} p_1^n(\theta|n'_\theta)R_n(n'_\theta|n_\theta), \quad \forall\, n_\theta. \tag{17}$$

The careful reader will notice an abuse of notation as the equilibrium action is written as $s_2^n(\theta|n_\theta)$ instead of $s_2^n(n_\theta)$. Given that the set of fundamentals, or states of the world, Θ is binary, we can without loss of generality write Player 1's equilibrium action as the probability that she assigns to the state being equal to θ. Thus, the notation $s_i^n(\theta|n_\theta)$ highlights that Player 1's equilibrium action is the probability that the true state is θ. This notational abuse simplifies the analysis in the following proofs.

Similarly, for any $i > 2$, Player i's equilibrium action is given by

$$s_i^n(\theta|n_\theta) = E_n\left[s_{i-1}^n(\theta|n'_\theta)|n_\theta\right] = \sum_{n'_\theta=0}^{n} s_{i-1}^n(\theta|n'_\theta)R_n(n'_\theta|n_\theta), \quad \forall\, n_\theta. \tag{18}$$

The proof of the proposition proceeds by induction through a sequence of intermediate steps. We first show two auxiliary results.

Lemma D.1. *For any $n > 0$, $E_n\left[s_1^n(\theta|n_\theta)|\theta\right] < E_{n+1}\left[s_1^{n+1}(\theta|n_\theta)|\theta\right]$.*

Proof. Let

$$A(n_\theta, n) = \frac{q_\theta^{2n_\theta}(1-q_\theta)^{2(n-n_\theta)}p_\theta}{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta + q_{\neg\theta}^{n_\theta}(1-q_{\neg\theta})^{n-n_\theta}p_{\neg\theta}}, \tag{19}$$

and notice that

$$\begin{aligned}\frac{\partial A(n_\theta, n)}{\partial n} \quad &= \quad p_\theta \ln\left(\frac{q_\theta}{1-q_\theta}\right) q_\theta^{2n_\theta}(1-q_\theta)^{2(n-n_\theta)} \\ &\times \frac{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta + \left[1+\ln\left(\frac{q_\theta}{1-q_\theta}\right) - \ln\left(\frac{q_{\neg\theta}}{1-q_{\neg\theta}}\right)\right] q_{\neg\theta}^{n_\theta}(1-q_{\neg\theta})^{n-n_\theta}p_{\neg\theta}}{\left[q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta + q_{\neg\theta}^{n_\theta}(1-q_{\neg\theta})^{n-n_\theta}p_{\neg\theta}\right]^2}, \end{aligned} \tag{20}$$

which is strictly positive because $q_\theta > \frac{1}{2} > q_{\neg\theta}$. Then,

$$\frac{\partial}{\partial n}\left[\binom{n}{n_\theta} A(n_\theta, n)\right] = A(n_\theta, n)\binom{n}{n_\theta}\sum_{k=0}^{n_\theta-1}\frac{1}{n-k} + \binom{n}{n_\theta}\frac{\partial A(n_\theta, n)}{\partial n} > 0. \tag{21}$$

This implies that

$$
\begin{aligned}
E_n\left[s_1^n(\theta|n_\theta)|\theta\right] &= \sum_{n_\theta=0}^{n} \binom{n}{n_\theta} A(n_\theta, n) \\
&\leq \sum_{n_\theta=0}^{n} \binom{n+1}{n_\theta} A(n_\theta, n+1) \\
&\leq \sum_{n_\theta=0}^{n} \binom{n+1}{n_\theta} A(n_\theta, n+1) + \binom{n+1}{n+1} A(n+1, n+1) \\
&\leq \sum_{n_\theta=0}^{n+1} \binom{n+1}{n_\theta} A(n_\theta, n+1) \\
&= E_{n+1}\left[s_1^{n+1}(\theta|n_\theta)|\theta\right],
\end{aligned}
$$

which completes the proof. □

LEMMA D.2. *$E_n\left[s_1^n(\theta|n_\theta)|\theta\right] > E_n\left[s_1^n(\theta|n_\theta)|\neg\theta\right]$, for any $n > 0$.*

Proof. Note that, for any $n > 0$,

$$
\begin{aligned}
E_n\left[s_1^n(\theta|n_\theta)|\neg\theta\right] &= E_n\left[p_1^n(\theta|n_\theta)|\neg\theta\right] \\
&= E_n\left[1 - p_1^n(\neg\theta|n_\theta)|\neg\theta\right] \\
&= 1 - E_n\left[p_1^n(\neg\theta|n_\theta)|\neg\theta\right] \\
&\geq 1 - E_{n+1}\left[p_1^{n+1}(\neg\theta|n_\theta)|\neg\theta\right] \\
&= E_{n+1}\left[s_1^{n+1}(\theta|n_\theta)|\neg\theta\right], \qquad (22)
\end{aligned}
$$

where the inequality follows from Lemma D.1. Since

$$
E_n\left[s_1^1(\theta|n_\theta)|\theta\right] - E_n\left[s_1^1(\theta|n_\theta)|\neg\theta\right] = \frac{q_{\neg\theta}(1-q_\theta)p_\theta}{(1-q_\theta)p_\theta + (1-q_{\neg\theta})p_{\neg\theta}} + \frac{q_\theta(1-q_{\neg\theta})p_\theta}{q_\theta p_\theta + q_{\neg\theta}p_{\neg\theta}} > 0, \qquad (23)
$$

the result follows from the opposite monotonicity of the two sequences $\left\{E_n\left[s_1^n(\theta|n_\theta)|\theta\right]\right\}_n$ and $\left\{E_n\left[s_1^n(\theta|n_\theta)|\neg\theta\right]\right\}_n$. □

The next lemma is the final step to prove the proposition.

LEMMA D.3. *For any $n \geq 1$ and $i \geq 2$, $E_n[s_{i-1}^n(\theta|n_\theta)|\theta] > E_n[s_i^n(\theta|n_\theta)|\theta]$.*

Proof. Notice that

$$\begin{aligned}
E_n[s_i^n(\theta|n_\theta)|\theta] &= \sum_{n_\theta=0}^{n} s_i^n(\theta|n_\theta) r_n(n_\theta|\theta) \\
&= \sum_{n_\theta=0}^{n} \left[\sum_{n'_\theta=0}^{n} s_{i-1}^n(\theta|n'_\theta) R_n(n'_\theta|n_\theta) \right] r_n(n_\theta|\theta) \\
&= \sum_{n'_\theta=0}^{n} s_{i-1}^n(\theta|n'_\theta) \left[\sum_{n_\theta=0}^{n} R_n(n'_\theta|n_\theta) r_n(n_\theta|\theta) \right] \\
&= \sum_{n'_\theta=0}^{n} s_{i-1}^n(\theta|n'_\theta) \left[\left(r_n(n'_\theta|\theta) - r_n(n'_\theta|\neg\theta) \right) E_n\left[p_i^n(\theta|n_\theta)|\theta \right] + r_n(n'_\theta|\neg\theta) \right] \\
&= E_n[s_{i-1}^n(\theta|n_\theta)|\neg\theta] \qquad (24) \\
&\quad + E_n\left[p_i^n(\theta|n_\theta)|\theta \right] \left\{ E_n[s_{i-1}^n(\theta|n_\theta)|\theta] - E_n[s_{i-1}^n(\theta|n_\theta)|\neg\theta] \right\}.
\end{aligned}$$

Then,

$$\begin{aligned}
E_n[s_{i-1}^n(\theta|n_\theta)|\theta] - E_n[s_i^n(\theta|n_\theta)|\theta] &= \{1 - E_n[p_i^n(\theta|n_\theta)|\theta]\} \qquad (25) \\
&\quad \times \left\{ E_n[s_{i-1}^n(\theta|n_\theta)|\theta] - E_n[s_{i-1}^n(\theta|n_\theta)|\neg\theta] \right\},
\end{aligned}$$

which is positive if and only if $E_n[s_{i-1}^n(\theta|n_\theta)|\theta] > E_n[s_{i-1}^n(\theta|n_\theta)|\neg\theta]$.[66] We know from Lemma D.2 that this inequality holds for $i = 1$. So, suppose that it holds for $i = k$, we will show that it also holds for $i = k + 1$. By equivalently expressing equation (24) in terms of $\neg\theta$ instead of θ gives

$$\begin{aligned}
E_n[s_k^n(\neg\theta|n_\theta)|\neg\theta] &= E_n\left[s_{k-1}^n(\neg\theta|n_\theta)|\theta \right] + E_n[p_k^n(\neg\theta|n_\theta)|\neg\theta] E_n\left[s_{k-1}^n(\neg\theta|n_\theta)|\neg\theta \right] \\
&\quad - E_n\left[p_k^n(\neg\theta|n_\theta)|\neg\theta \right] E_n\left[s_{k-1}^n(\neg\theta|n_\theta)|\theta \right]. \qquad (26)
\end{aligned}$$

Next, given that the state space is binary, it holds that

$$\begin{aligned}
E_n[s_k^n(\theta|n_\theta)|\neg\theta] &= E_n[E_n[s_{k-1}^n(\theta|n_\theta)|n_\theta]|\neg\theta] \\
&= E_n[1 - E_n[s_{k-1}^n(\neg\theta|n_\theta)|n_\theta]|\neg\theta] \\
&= 1 - E_n[E_n[s_{k-1}^n(\neg\theta|n_\theta)|n_\theta]|\neg\theta] \\
&= 1 - E_n[s_k^n(\neg\theta|n_\theta)|\neg\theta]. \qquad (27)
\end{aligned}$$

[66] Since the distributions are non-degenerate, it holds that $E_n[p_i^n(\theta|n_\theta)|\theta] < 1$.

Combining (24) and (27), and rearranging using (27), we obtain

$$\begin{aligned} E_n[s_k^n(\theta|n_\theta)|\theta] - E_n[s_k^n(\theta|n_\theta)|\neg\theta] &= \{1 + E_n\left[p_k^n(\theta|n_\theta)|\theta\right] - E_n\left[p_k^n(\neg\theta|n_\theta)|\neg\theta\right]\} \\ &\quad \times \left\{E_n[s_{k-1}^n(\theta|n_\theta)|\theta] - E_n[s_{k-1}^n(\theta|n_\theta)|\neg\theta]\right\} \\ &= \{1 + E_n\left[p_1^n(\theta|n_\theta)|\theta\right] - E_n\left[p_1^n(\theta|n_\theta)|\neg\theta\right]\} \\ &\quad \times \left\{E_n[s_{k-1}^n(\theta|n_\theta)|\theta] - E_n[s_{k-1}^n(\theta|n_\theta)|\neg\theta]\right\} \\ &= \{1 + E_n\left[s_1^n(\theta|n_\theta)|\theta\right] - E_n\left[s_1^n(\theta|n_\theta)|\neg\theta\right]\} \\ &\quad \times \left\{E_n[s_{k-1}^n(\theta|n_\theta)|\theta] - E_n[s_{k-1}^n(\theta|n_\theta)|\neg\theta]\right\} \end{aligned} \tag{28}$$

where the second equality follows from the observation that all players share the same beliefs given the same history, while the third equality follows from the equilibrium action of Player 1. Since $E_n\left[s_1^n(\theta|n_\theta)|\theta\right] - E_n\left[s_1^n(\theta|n_\theta)|\neg\theta\right] > 0$ by Lemma D.2, and $E_n[s_{k-1}^n(\theta|n_\theta)|\theta] - E_n[s_{k-1}^n(\theta|n_\theta)|\neg\theta] > 0$ by the induction hypothesis, it follows that $E_n[s_k^n(\theta|n_\theta)|\theta] - E_n[s_k^n(\theta|n_\theta)|\neg\theta] > 0$. Finally, this implies that $E_n[s_k^n(\theta|n_\theta)|\theta] > E_n[s_{k+1}^n(\theta|n_\theta)|\theta]$, for any $k \geq 1$. □

Proof of Proposition 3: We will first show that $\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\theta] = 1$. By Lemma D.1, the sequence $\{E_n[s_1^n(\theta|n_\theta)|\theta]\}_{n=1}^\infty$ is monotonically increasing. As the sequence is also uniformly bounded above by 1, it clearly converges. By Markov's inequality, for any $\epsilon \in (0,1)$, and $n > 0$,

$$\text{Prob}_n\left(s_1^n(\theta|n_\theta) \geq 1-\epsilon|\theta\right) \leq \frac{E_n[s_1^n(\theta|n_\theta)|\theta]}{1-\epsilon}. \tag{29}$$

Suppose for now that $\lim_{n\to\infty} s_1^n(\theta|n_\theta) = 1$ a.s. conditional on θ. Since almost sure convergence implies convergence in probability, it follows that

$$1 = \lim_{n\to\infty} \text{Prob}\left(s_1^n(\theta|n_\theta) \geq 1-\epsilon|\theta\right) \leq \frac{\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\theta]}{1-\epsilon} \leq \frac{1}{1-\epsilon}, \tag{30}$$

which implies that $\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\theta] \in [1-\epsilon, 1]$. Since ϵ was arbitrary, we can conclude that $\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\theta] = 1$.

The fact that $\lim_{n\to\infty} s_1^n(\theta|n_\theta) = 1$ a.s. is well-known but we prove it here for completeness. Suppose without loss of generality that the true state is θ, and define the likelihood ratio

$$l_n = \frac{\text{Prob}(\neg\theta|\mathbf{t}^n)}{\text{Prob}(\theta|\mathbf{t}^n)}, \tag{31}$$

where $\mathbf{t}_n$ denotes the set of one-dimensional signals observed until chain game n. Notice that

$$l_{n+1} = \frac{\text{Prob}(\neg\theta|\mathbf{t}^n \cup \{t\})}{\text{Prob}(\theta|\mathbf{t}^n \cup \{t\})} = \frac{\frac{\text{Prob}(\mathbf{t}|\neg\theta,\mathbf{t}^n)\text{Prob}(\neg\theta|\mathbf{t}^n)}{\text{Prob}(t|\mathbf{t}^n)}}{\frac{\text{Prob}(t|\theta,\mathbf{t}^n)\text{Prob}(\theta|\mathbf{t}^n)}{\text{Prob}(t|\mathbf{t}^n)}} = \frac{\text{Prob}(t|\neg\theta,\mathbf{t}^n)}{\text{Prob}(t|\theta,\mathbf{t}^n)} l_n. \tag{32}$$

Then, $E[l_{n+1}|\mathbf{t}^n] = l_n E\left[\left.\frac{\text{Prob}(t|\neg\theta,\mathbf{t}^n)}{\text{Prob}(t|\theta,\mathbf{t}^n)}\right|\mathbf{t}^n\right]$, since $\mathbf{t}^n$ is a sufficient statistic for l_n. This shows that if $E\left[\left.\frac{\text{Prob}(t|\neg\theta,\mathbf{t}^n)}{\text{Prob}(t|\theta,\mathbf{t}^n)}\right|\mathbf{t}^n\right] = 1$, then the sequence of random variables $\{l_n\}$ is a bounded martingale,

$$\begin{aligned}
E\left[\left.\frac{\text{Prob}(t|\neg\theta,\mathbf{t}^n)}{\text{Prob}(t|\theta,\mathbf{t}^n)}\right|\mathbf{t}^n\right] &= E\left[\left.\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)}\right|\mathbf{t}_n\right] \\
&= \text{Prob}(t_\theta|\theta)\frac{\text{Prob}(t_\theta|\neg\theta)}{\text{Prob}(t_\theta|\theta)} + \text{Prob}(t_{\neg\theta}|\theta)\frac{\text{Prob}(t_{\neg\theta}|\neg\theta)}{\text{Prob}(t_{\neg\theta}|\theta)} \\
&= \text{Prob}(t_\theta|\neg\theta) + \text{Prob}(t_{\neg\theta}|\neg\theta) \\
&= 1.
\end{aligned} \tag{33}$$

By the martingale convergence theorem (e.g., Williams (1991), Theorem 11.5), l_n converges to a random variable l_∞ almost surely. Next, fix an arbitrary $\epsilon > 0$,

$$\begin{aligned}
\text{Prob}\left(|l_{n+1} - l_n| > \epsilon\right) &= \text{Prob}\left(\left|\left(\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right) l_n\right| > \epsilon\right) \\
&\geq E\left[\mathbf{1}_{\{l_n > \sqrt{\epsilon}\}}\mathbf{1}_{\left\{\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right\}}\right] \\
&= E\left[E\left[\left.\mathbf{1}_{\{l_n > \sqrt{\epsilon}\}}\mathbf{1}_{\left\{\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right\}}\right| l_n\right]\right] \\
&= E\left[\mathbf{1}_{\{l_n > \sqrt{\epsilon}\}}\text{Prob}\left(\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right)\right] \\
&= \text{Prob}\left(l_n > \sqrt{\epsilon}\right)\text{Prob}\left(\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right).
\end{aligned} \tag{34}$$

Then,

$$\begin{aligned}
0 &\leq \text{Prob}\left(l_n > \sqrt{\epsilon}\right)\text{Prob}\left(\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right) \\
&\leq \text{Prob}\left(|l_{n+1} - l_n| > \epsilon\right) \\
&\leq \text{Prob}\left(|l_{n+1} - l_\infty| + |l_\infty - l_n| > \epsilon\right).
\end{aligned} \tag{35}$$

Since we have shown that $l_n \to l_\infty$ a.s., then

$$\text{Prob}\left(|l_{n+1} - l_\infty| + |l_\infty - l_n| > \epsilon\right) \to 0, \tag{36}$$

and $\text{Prob}\left(\left|\frac{\text{Prob}(t|\neg\theta)}{\text{Prob}(t|\theta)} - 1\right| > \sqrt{\epsilon}\right) \neq 0$ implies that $\text{Prob}\left(l_n > \sqrt{\epsilon}\right) \to 0$. Given that ϵ was arbitrary, it follows that $l_\infty = 0$ a.s. Finally, $l_t \to 0$ a.s. implies that $\text{Prob}(\neg\theta|h_n) \to 0$ a.s. or, equivalently, $s_1^n(\theta|n_\theta) \to 1$ a.s.

If $\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\theta] = 1$, then by a similar argument $\lim_{n\to\infty} E_n[s_1^n(\neg\theta|n_\theta)|\neg\theta] = 1$ which implies that $\lim_{n\to\infty} E_n[s_1^n(\theta|n_\theta)|\neg\theta] = 0$. Therefore, from (24) with $i = 2$, we can conclude that $\lim_{n\to\infty} E_n[s_2^n(\theta|n_\theta)|\theta] = 1$. Using this fact and (24) again, we can show that the claim holds for $i = 3$, etc.

Finally, the monotonicity and convergence of expectations of any order further implies that $\lim_{n\to+\infty} |E[s_i^n(\mathbf{t}_i^n)|\theta] - E[s_{i-1}^n(\mathbf{t}_{i-1}^n)|\theta]| \to 0$, for any $i > 2$. This completes the proof.