Measuring Agents' Reaction to Private and Public Information in Games with Strategic Complementarities*

Camille Cornand
GATE CNRS, Lyon University
93 Chemin des Mouilles – BP 167
69131 Ecully cedex, France

Email: cornand@gate.cnrs.fr

Frank Heinemann
Technische Universität Berlin
Chair of Macroeconomics, H 52
Straße des 17. Juni 135
10 623 Berlin, Germany
Email: frank.heinemann@ww.tu-berlin.de

February 19, 2013

Abstract: In games with strategic complementarities, public information about the state of the world has a larger impact on equilibrium actions than private information of the same precision, because public signals are more informative about the likely behavior of others. We present an experiment in which agents' optimal actions are a weighted average of the fundamental state and their expectations of other agents' actions. We measure the responses to public and private signals. We find that, on average, subjects put a larger weight on the public signal. In line with theoretical predictions, as the relative weight of the coordination component in a player's utility increases, players put more weight on the public signal when making their choices. However, the weight is smaller than in equilibrium, which indicates that subjects underestimate the information contained in public signals about other players' beliefs.

JEL Classification: C92, D82, D84.
Keywords: coordination games, strategic uncertainty, private information, public information.

1 – Introduction

Financial markets, macroeconomic environments, and network games are often characterized by strategic complementarities. In games with strategic complementarities, public and private information about an uncertain state of the world have distinct effects, because information that is shared by several agents is more informative about the likely beliefs and actions of others than private information. Since agents have an incentive to align their actions to the expected actions of others, equilibrium choices are affected more by public than by private signals of equal precision. This has been used to explain why markets 'overreact' to public signals, so that public information may raise price volatility on financial markets and cause negative welfare effects. The aim of this paper is to experimentally evaluate the weights that subjects attach to private and public signals in games of weak

* We are thankful to Gabriel Desgranges, John Duffy, Petra Geraats, Ed Hopkins, Paul Pezanis-Christou, and Alexander Meyer-Gohde for helpful comments, to Romain Zeiliger for programming Regate, to Kene Boun My for providing assistance, to the students for their participation in the experiment and to the *Conseil Scientifique de l'Université de Strasbourg* and the ANR-DFG joint grant for financial support.

strategic complementarities[1]. We show that on average agents indeed attach a higher weight to public than to private information, but the asymmetry in weights is smaller than in equilibrium.

Our experiment builds upon Morris and Shin (2002), who present a generalized beauty contest game with weak strategic complementarities for analyzing the distinct effects of public and private information. In their game, an agent's action is a weighted average of a public and a private signal. The equilibrium weight attached to the public signal is higher than its relative precision. This 'overreaction' (Morris and Shin (2002), p. 1530) is due to the higher informational content of public signals regarding the likely beliefs of other agents. In this context, Morris and Shin show that raising the precision of public signals may reduce expected welfare.[2] Welfare effects of public and private information depend on two considerations requiring empirical investigation: the relative precision of the two signals and the relative weight that agents attach to these signals. These considerations interact: the higher the weight that agents attach to public signals, the more likely it is that an imprecise public signal will reduce welfare. In equilibrium, the relative weight is determined by the signals' relative precision and agents' payoff functions.

This paper precisely aims at measuring and analyzing how much public signals are taken into account compared to private ones – in a game with weak strategic complementarities. We present an experiment on a game that is characterized by both fundamental and strategic uncertainty. As in Morris and Shin (2002), agents have to choose actions that are close to a fundamental state but also close to each other. We test predictions of this approach by implementing two-player versions of this game with varying weights on fundamental and strategic uncertainty. In line with theory, we find that: (i) subjects put larger weights on the public signal and (ii) as the relative weight of the coordination component in a player's utility rises, players put more weight on the public signal. However, overall, the weights put on the public signal are smaller than theoretically predicted.

A cognitive hierarchy model with 2 levels of reasoning helps explaining the observation that subjects attach a lower weight on public signals than predicted by the equilibrium. Stated second order beliefs, however, indicate that subjects underestimate how informative the public signal is in assessing other players' expectations. This may be viewed as an alternative explanation of behavior. However, these non-Bayesian higher-order beliefs can explain only a part of the observed systematic deviation of behavior from equilibrium.

The next section presents the basic game. Section 3 sets the experimental design and motivates hypotheses. Section 4 states the experimental results and provides measures for the relative

---

[1] Angeletos and Pavan (2004) distinguish weak and strong strategic complementarities. Under weak complementarities, the game has a unique equilibrium, while strong strategic complementarities define a coordination game with multiple equilibria.

[2] Svensson (2006) has argued that detrimental welfare effects of public information are unlikely, because they require that public information be of lower precision than private information, which conflicts with empirical findings. For a response, see Morris et al. (2006). For discussions about the welfare effects of public information in closely related frameworks, see Cornand and Heinemann (2008), Baeriswyl and Cornand (2011), Woodford (2005), Angeletos and Pavan (2007 a,b), Hellwig (2005), and Myatt and Wallace (2008) among others.

weights that subjects put on private and common signals, relating them to equilibrium predictions. Section 5 presents two behavioral theories to explain the observed deviation of behavior from equilibrium. Finally, section 6 concludes.

2 – Two-player game with weak strategic complements

We describe a game in which agents have to choose actions that are close to a fundamental state but also close to what the others believe. The framework is a two-player version of Morris and Shin (2002) adapted for conducting an experiment.[3]

The fundamental state of nature is given by θ and has a uniform distribution on the reals. Agent i chooses action $a_i \in \Re$, and we write a for the action profile over all agents. The utility function for individual i has two loss components:

$$u_i(a,\theta) \equiv U_0 - (1-r)\,(a_i - \theta)^2 - r\,(a_i - a_j)^2\,, \tag{1}$$

where j is the partner of i and U_0 is a constant that cannot be affected by player i. The first quadratic term captures losses arising from the distance between an action and the underlying state θ. The second quadratic term describes losses from miscoordination between both players' actions. Finally, $r \in [0,1]$ indicates the relative weight attributed to the coordination component. The coordination component is reminiscent of Keynes' beauty contest.

Agents face uncertainty concerning θ. They receive two kinds of signals that deviate from θ by some error terms with uniform distribution. All agents receive a public (common) signal $y \sim U[\theta - \varepsilon, \theta + \varepsilon]$. In addition, each agent receives a private signal $x_i \sim U[\theta - \varepsilon, \theta + \varepsilon]$. Noise terms $x_i - \theta$ of distinct individuals and the noise of the public signal $y - \theta$ are independent and their distribution is treated as exogenously given.

The optimal action of agent i is given by the first order condition from maximizing (1):

$$a_i = (1-r)E_i(\theta) + r\,E_i(a_j)\,, \tag{2}$$

where $E_i(\cdot)$ is the posterior expectation conditional on x_i and y and a_j is the action of the other player j. One needs to calculate $E_i(\theta)$ and $E_i(a_j)$. Because both signals have the same precision,

$$E_i(\theta) = E(\theta\,|x_i, y) = \frac{x_i + y}{2} = E(x_j\,|x_i, y).$$

The optimal strategy of any agent is a linear combination of his signals: $a_j = \gamma\, x_j + (1-\gamma)y$. So,

[3] Compared to the model of Morris and Shin (2002), we required the number of players to be finite and we also changed the distributions of the signals from normal to uniform in order to have a simple distribution with bounded support for the experiment. In Morris and Shin, the constant U_0 adds the average losses from miscoordination, so that the coordination part becomes a zero-sum game and aggregate welfare depends only on the distance between actions and fundamental state.

$$E_i(a_j) = E\left(a_j \left|x_i, y\right.\right) = \gamma\, E\left(x_j \left|x_i, y\right.\right) + (1-\gamma)y \;= (1-\gamma)y + \gamma \frac{x_i + y}{2}.$$

Inserting this in equation (2) yields

$$a_i = rE_i\left(a_i \left|x_i, y\right.\right) + (1-r)E_i\left(\theta \left|x_i, y\right.\right) = \frac{x_i\left[(1 - r(1-\gamma))\right] + y\left[1 + r(1-\gamma)\right]}{2}.$$

Thus, in equilibrium $\gamma^* = \frac{1-r(1-\gamma^*)}{2}$, which is equivalent to $\gamma^* = \frac{1-r}{2-r}$. Agent *i*'s equilibrium action is now given by $a_i^* = \frac{y+(1-r)x_i}{2-r}$. The equilibrium weight on private information γ^* falls short of its weight of ½ in the Bayesian expectation of the fundamental, because the public signal conveys more information about the likely action of the other player than the private signal.

In equilibrium, actions are distorted away from the expected θ towards *y*. This mirrors the disproportionate impact of the public signal in coordinating agents' actions and emphasizes the role of public information as a focal point for private actions. Note that the equilibrium weight on the private signal γ^* decreases in the relative importance of coordination measured by *r*. Morris and Shin (2002) concentrate on the welfare effects of public information and show that if welfare depends only on how close actions are to the fundamental, public signals of low precision may actually reduce ex-ante expected welfare, even though they improve agents' ability to predict the fundamental. Here, the negative welfare effect from distorting actions towards the public signal exceeds the positive welfare effect that arises from the additional information. In these environments, the economy is better off without a public signal. This argument has been used to question the merits of central bank transparency.

Negative welfare effects of public signals require that agents put a sufficiently large weight on public signals. In beauty contest games like the one used by Morris and Shin (2002), real players' behavior has been found consistent with limited levels of reasoning (Nagel, 1995). As will be explained in Section 5 below, such behavior leads to weights on the public signal that are lower than in equilibrium. Cornand and Heinemann (2012) show that negative welfare effects of public disclosures require that agents put weights on public signals that are at least as large as the weight associated with level-2 reasoning. Thus, it is an empirical question, whether real players put a weight on public signals that is sufficiently strong to cause negative welfare effects of public signals.[4] To

[4] Other experiments concerning the welfare effects of public information have concentrated on games with multiple equilibria: Anctil *et al.* (2004) demonstrate that private signals with high precision are not sufficient to achieve coordination on an efficient equilibrium. Heinemann et al. (2004) compare perfect public and noisy private signals in a speculative attack game and find small effects towards higher efficiency with perfect information. Cornand (2006) shows that subjects put a larger weight on the public signal if they receive both a private and a public signal about the state of the economy. For other experiments dealing with public versus private information, see Forsythe et al. (1982), Plott and Sunder (1988), McKelvey and Ordeshook (1985), McKelvey and Page (1990) and Hanson (1996). They investigate how individuals use public information to

answer this question, we propose an experiment that measures the actual weights that subjects put on private versus public information.

3 – Experimental design

The experiment aims at measuring reactions to private and public information in the game introduced in the last section. In particular, we are interested in testing whether subjects put larger weights on public than on private signals, whether the observed weights are in line with equilibrium predictions, and how changes in r affect these weights.

18 sessions were run at the BETA (*Bureau d'Economie Théorique et Appliquée*) laboratory in Strasbourg (using software Regate (Zeiliger, 2000)) between January 2008 and March 2010. Each session had 16 participants who were mainly students from Strasbourg University (most were students in economics, mathematics, biology and psychology). Subjects were seated in random order at PCs. Instructions were then read aloud and questions answered in private. Throughout the sessions, students were not allowed to communicate with one another and could not see each others' screens. Each subject could only participate in one session. Before starting the experiment, subjects were required to answer a few questions to ascertain their understanding of the rules. Instructions and questionnaires are given in Appendix A. The experiment started after all subjects had given the correct answers to these questions.

Each session consisted of four stages with a total of 50 periods. While Stages 1, 2, and 4 were the same in all sessions, Stage 3 contained three different treatments across sessions. In each period, subjects were paired up randomly. They did not know the identity of their partner and they knew that they would most likely not meet the same partner in the next period. For each pair of subjects, a fundamental state θ was drawn randomly using a uniform distribution from the interval [50, 450]. Each subject received a private signal $x_i \in [\theta - 10, \theta + 10]$. In addition, each pair of subjects received a common (public) signal $y \in [\theta - 10, \theta + 10]$. Signals were drawn independently from these intervals using a uniform distribution.[5] The random process was explained in the instructions.

In the first three stages, each subject had to decide on an action a_i conditional on her signals.[6] Signals and actions were limited to one decimal point.

---

augment their original private information and whether, in doing so, a rational expectations equilibrium is attained.

[5] Having a sufficiently large support of the prior distribution enables us to reduce the informational content conveyed by the prior mean. In addition, this reduces the set of signals, for which the conditional posterior distribution is skewed.

[6] To limit potential losses that might arise from typing mistakes, we restricted choices and stated expectations to the interval $[y - 20, y + 20]$. This restriction was not mentioned in the instructions. If a subject entered a number outside this interval, the screen displayed a message indicating that her choice is too far away from signals. Apparently, this restriction was not binding. We observed only one (out of 12,960) decision situations, where a subject actually chose an action at the border of this interval, and less than 50 decisions, where the difference between action and public signal exceeded 15. Some subjects actually reported typing mistakes.

In the first stage ("r=0"), the payoff function was given by $100-(a_i-\theta)^2$. Here, subjects should choose an action as close as possible to the fundamental state, independent from their partners' choices. This stage aimed at familiarizing subjects with the game form and controlling for their Bayesian rationality. Since both signals have the same precision, the rational choice is $a_i=(x_i+y)/2$ whenever both signals are in the interval [60, 440].[7]

In the second stage ("r=1"), the payoff function was given by $100-(a_i-a_j)^2$, where a_i denotes the partner's action. Here, subjects should coordinate their actions irrespective of the fundamental state: this is a pure coordination game from which subjects should learn that public information is more helpful in overcoming coordination problems than private information. Any strategy that maps the public signal into the reals is an equilibrium[8], provided that both subjects coordinate on the same strategy. The public signal, however, provides a focal point,[9] so that we expected subjects to coordinate on actions $a_i=y$.

Stage 3 contained three different treatments, each conducted in 6 of the sessions. In Treatment "r=0.5", the payoff function was $200-(a_i-a_j)^2-(a_i-\theta)^2$, while it was $400-(a_i-a_j)^2-3(a_i-\theta)^2$ in Treatment "r=0.25" and $400-3(a_i-a_j)^2-(a_i-\theta)^2$ in Treatment "r=0.75". Contrary to Stages 1 and 2, the three treatments of Stage 3 exhibit both a fundamental and a coordination motive. They correspond to the game by Morris and Shin (2002). Given that players choose linear combinations of both signals $a_j=\gamma x_j+(1-\gamma)y$, the equilibrium weight[10] on the private signal in Stage 3 is $\gamma^*=(1-r)/(2-r)$. Note that the limit of γ^* for $r\to 1$ selects the equilibrium with $a_i=y$ for $r=1$.

In Stage 4 we elicited higher-order beliefs. Subjects also received private and public signals. But here, each subject was asked to state two expectations: one regarding the fundamental state and one regarding her partner's expectation of the fundamental state. The payoff function was

---

[7] When signals are smaller than 60 or larger than 440, the posterior distribution of θ is skewed, because of the bounded support of θ.

[8] In the limiting case where fundamental uncertainty disappears and subjects' payoffs depend only on the distance between their actions, equilibrium theory does not yield a unique prediction. Any coordinated strategy is an equilibrium. However, the limit of equilibria in games with a decreasing weight on fundamental uncertainty uniquely selects a strategy in which all agents follow the public signal and ignore all private information.

[9] A formalization of focal points is provided by Alós-Ferrer and Kuzmics (2008).

[10] Due to the bounded support of θ, the equilibrium deviates from this linear combination into the direction of the center of the support. To see this, imagine that player i receives signals x_i = 40 and y = 50. From her private signal, she can deduce that θ = 50. The posterior distribution of the other player's signal is uniform in [40, 60]. Since the other player should never choose an action below 50, the expected action by the other player is above 50. Therefore, player i should also choose an action above 50. Obviously, it is too demanding to assume that players update their beliefs correctly at the edges of the support. In the data, we find systematic effects for signals up to 60 and above 440. We checked that restricting data analysis to situations with $70<\theta<430$ does not alter our results.

$100-\left(e_i(\theta)-\theta\right)^2-\left(e_i\left(e_j(\theta)\right)-e_j(\theta)\right)^2$, where $e_i(\cdot)$ denotes the stated expectation of subject *i*. The last term of the payoff function incentivizes the statement of higher-order beliefs. To our knowledge, this is the first experiment with a direct elicitation of higher-order beliefs.

Since individual decisions are not independent in repeated coordination games, we conducted 18 sessions with at least 6 sessions per treatment.[11] For statistical tests, we use the average weight on the private signal within a session as an independent observation. Table 1 gives an overview over sessions, stages, and treatments.

Sessions	*Stage 1*	*Stage 2*	*Stage 3*	*Stage 4*
1 to 6	*"r=0"* *(5 periods)*	*"r=1"* *(10 periods)*	*Treatment "r=0.5" (30 periods)*	*Belief Elicitation (5 periods)*
7 to 12			*Treatment "r=0.25" (30 periods)*	
13 to 18			*Treatment "r=0.75" (30 periods)*[12]	

Table 1 – Composition of the different sessions

After each period, subjects were informed about the true state, their partner's decision and their payoff. Information about past periods from the same stage (including signals and own decisions) was displayed during the decision phase on the lower part of the screen. Subjects knew from the start that they were playing the aforementioned stages in this order. At the end of each session, the points earned were summed up and converted into Euros. 100 points were converted to 25 cents in sessions 1 – 6 and 13 – 18, and to 43 cents in sessions 7 – 12. Payoffs ranged from € 12 to € 33. The average payoff was about € 22.[13] Sessions lasted for around 90 minutes.

Our main hypotheses are derived from the theory predictions: Table 2 summarizes the equilibrium weights on the private signal γ^* for interior states. The natural hypothesis arising from theory is that subjects attach weights to private signals that are distributed around the equilibrium weights. As explained above, we intend to test this hypothesis against the alternative arising from a cognitive hierarchy model which predicts higher weights on the private signal for $r > 0$. Two other testable hypotheses are consistent with equilibrium *and* with limited levels of reasoning: the weights put on the private signal are smaller than 0.5 (for $r > 0$) and decreasing in *r*.

[11] We also conducted 6 control sessions to include treatments with only one signal in order to test the welfare reducing effect of private signals for *r*=1 directly by comparing payoffs. Control sessions elicited beliefs at the second stage before strategic games and with a slightly different representation of the payoff function. We only present standard sessions in the main text. The results of control sessions are reported in Appendix B.

[12] In Session 13, Treatment "*r*=0.75" unintentionally stopped after 26 periods.

[13] In all stages, it was possible to earn negative points. This actually occurred in about 3.3% of all decision situations in standard sessions. Realized losses were of a size that could be counterbalanced by positive payoffs within a few periods. In general, losses were covered by earnings from the next three periods or balanced by earnings of previous periods.

r	0	0.25	0.5	0.75	$\to 1$
γ^*	0.500	0.429	0.333	0.200	$\to 0$

Table 2 – Equilibrium weight γ^* as a function of r

4 – Experimental results

We use the following structure in analyzing data. First, we analyze whether subjects' choices are between their two signals as predicted by theory. Here we also look at the proportion of choices that are closer to the public or private signal contingent on the stage or treatment. Then, we estimate the average weights that subjects attach to the private signal for each session, stage, and treatment. We investigate whether average weights on the private signal are positive, smaller than 0.5, and whether they deviate systematically from the equilibrium value. If subjects best respond to the observed actions of their respective partners, their strategies should converge over time towards the equilibrium. We test whether there is a time trend by comparing the data from the first and second halves of a stage. Comparing observed weights from different treatments, we test comparative static predictions arising from theory.

4.1. Some considerations about rationality

Theory predicts that subjects choose actions that are weighted averages of their two signals. Hence, actions should be contained in $I_1 = [\min(y, x_i), \max(y, x_i)]$. However, we observe that about 10% of all decisions are outside this interval, in particular if both signals are very close to each other. To get a better impression, how far these choices go outside, we also look whether choices are consistent with the contingent support of the state's distribution. From her signals, subject *i* can deduce that the true state of the world is contained in the interval $I_2 = [\max(y, x_i) - 10, \min(y, x_i) + 10]$. We observe that the proportion of choices outside this interval also becomes large when the interval is small. Furthermore, the smaller is I_1, the larger is I_2. Since both of these intervals have some appeal for reasonable choices, we check, how many choices were contained in the union of both sets, $I_3 = I_1 \cup I_2$. Note that the size of interval I_3 is less sensitive to the randomly drawn signals and never gets smaller than 10, while the size of I_1 and I_2 may shrink to zero for appropriate signals.

Result 1: 98% of choices are contained in I_3. The proportion of choices inside I_1 is 85%.

Value of r	*0*	*0.25*	*0.5*	*0.75*	*1*
Inside $I_1 = [\min(y,x_i), \max(y,x_i)]$	*75%*	*85%*	*86%*	*86%*	*86%*
Choice = y	*1%*	*1%*	*7%*	*21%*	*33%*
Closer to y	*29%*	*34%*	*44%*	*35%*	*26%*
Middle (+/-0.05)	*14%*	*18%*	*11%*	*10%*	*9%*
Closer to x_i	*31%*	*31%*	*24%*	*20%*	*18%*
Choice = x_i	*1%*	*1%*	*1%*	*0.4%*	*1%*
Outside I_1 *beyond y*	*12%*	*8%*	*8%*	*8%*	*8%*
Outside I_1 *beyond* x_i	*12%*	*8%*	*6%*	*7%*	*6%*
Inside $I_2 = [\max(Y,X)-10; \min(y,X)+10]$	*94%*	*97%*	*96%*	*91%*	*88%*
Inside $I_3 = I_1 \cup I_2$	*97%*	*99%*	*99%*	*99%*	*98%*

Table 3 – Crude classification of individual choices

Table 3 displays the percentage of all choices within these intervals. Counting the number of choices that are closer to the common signal, closer to the private, or in the middle provides a crude first impression of whether subjects put a larger weight on the common signal and how dispersed the distribution of relative weights is.[14] As *r* is increased, the proportion of choices closer or equal to the common signal *y* increases.

4.2. Estimated weights on the private signal

We estimate the relative weights of the two signals on subjects' decisions by fitting linear regressions to data. Because the payoff function rewards coordination when *r* > 0, we may expect that the weights applied by different subjects within a session converge over time. Hence, individual weights are not independent. To be on the conservative side, our main statistical analyses are based on the average weights per session. We estimate these weights by joining the data from all subjects within a session. The regression equation is

$$a_{it} - y_t = c + \gamma (x_{it} - y_t) + u_{it},$$

where γ is the session-specific relative weight on private information. The constant c stands for a bias towards higher or lower numbers. Using the Bonferroni correction for multiple hypotheses, this constant turns out to be insignificant in all stages and treatments. Implicitly, our regression imposes a restriction that the absolute coefficients on private and common signals add up to 1. In regressions without this restriction, we checked that violations of this restriction are not significant either.[15]

Table 4 displays average weights on the private signal for all sessions and stages and compares them with theoretical weights. For games that were played at least 10 rounds, we split the data from the first and second halves of the stage so that we could analyze whether there was some systematic change of behavior over time. Figure 1 presents the evolution of these weights over periods; this representation captures the differences across stages within each group of sessions and between

[14] To break ties in counting: if both signals and the action coincide, we count the choice as "Middle". If $|y - x_i| = 0.1$, and $a_i = x_i$ $[a_i = y]$, we count the choice as "equal to x_i" ["equal to *y*"].

[15] Data and programs are available at http://www.macroeconomics.tu-berlin.de/menue/working_papers/.

different treatments across sessions. Note that average weights on private signals stay above the equilibrium values (for *r* > 0) in all periods.

Session	*r* = 0	*r* = 1, 1st half	*r* = 1, 2nd half	*r* = 0.5, 1st half	*r* = 0.5, 2nd half	*r*=0.25, 1st half	*r*=0.25, 2nd half	*r*=0.75, 1st half	*r*=0.75, 2nd half
1	0.522	0.266	0.158	0.419	0.408				
2	0.506	0.369	0.219	0.452	0.463				
3	0.570	0.318	0.390	0.453	0.447				
4	0.475	0.380	0.344	0.416	0.475				
5	0.453	0.328	0.098	0.340	0.393				
6	0.516	0.349	0.200	0.409	0.453				
7	0.435	0.353	0.185			0.485	0.492		
8	0.472	0.346	0.209			0.498	0.466		
9	0.579	0.398	0.363			0.521	0.500		
10	0.496	0.339	0.135			0.480	0.460		
11	0.559	0.286	0.195			0.454	0.475		
12	0.494	0.257	0.210			0.477	0.473		
13	0.497	0.376	0.205					0.417	0.391
14	0.504	0.190	0.208					0.309	0.294
15	0.492	0.468	0.387					0.405	0.457
16	0.489	0.390	0.350					0.408	0.377
17	0.501	0.352	0.162					0.289	0.187
18	0.555	0.389	0.331					0.288	0.338
Average	0.506	0.342	0.242	0.415	0.440	0.486	0.478	0.353	0.341
St.dev.	0.039	0.063	0.092	0.041	0.032	0.022	0.015	0.063	0.093
Equilibrium weight	0.5	0	0	0.333	0.333	0.429	0.429	0.2	0.2

Table 4 – Group specific weights on the private signal

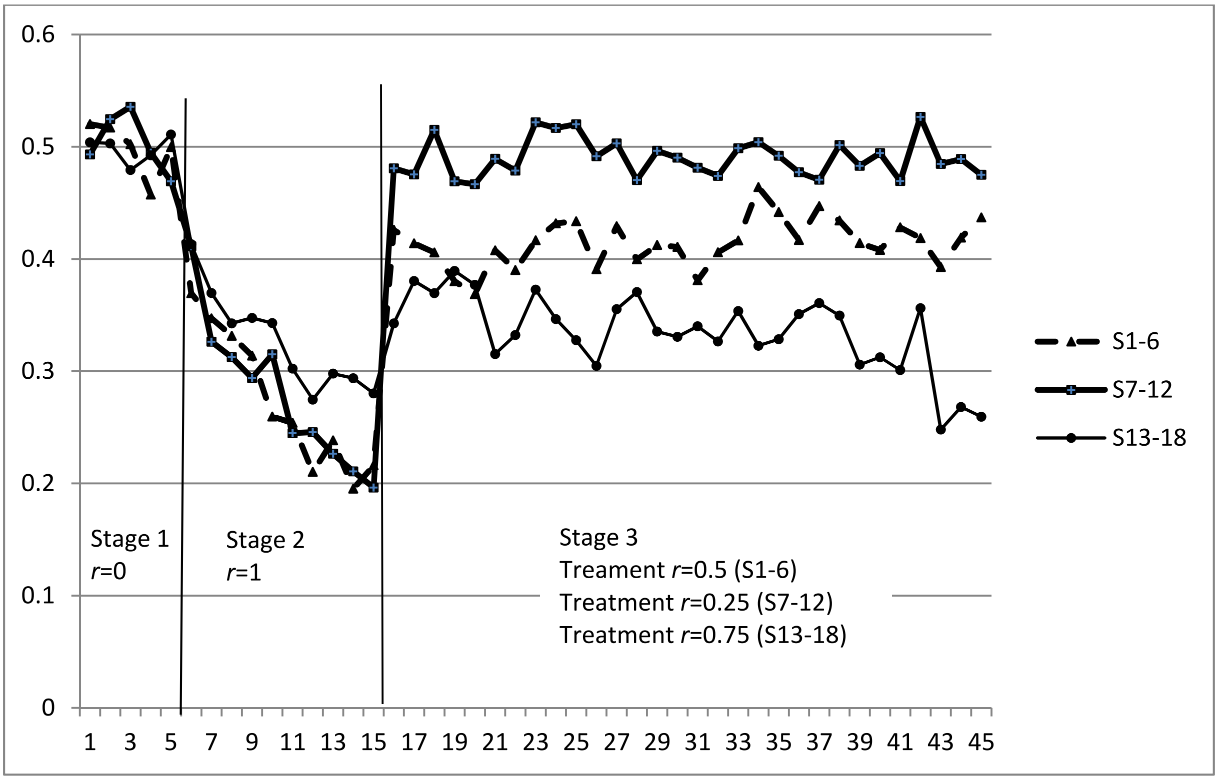


Figure 1 - Evolution of average weights on the private signal

Data indicate that the relative weight on the private signal is around 0.5 for “*r*=0” and tends to be decreasing in *r*. For *r*=1, *r*=0.5, and *r*=0.25, all session-specific weights are larger than in equilibrium.

For r=0.75, there is one exception (Session 17, 2nd half). Comparing first and second halves, there seems to be some trend towards equilibrium in stage "r=1", but not in treatments with $0 < r < 1$.

These impressions are widely confirmed by nonparametric tests, reported below. Specifically, we test whether estimated weights differ from 0.5, from the respective equilibrium values, and between first and second halves of a stage/treatment. We also test monotony of the weights in r. Unless otherwise noted, all tests are based on counting session-specific averages as independent observations with a significance level of 5%. When we split data from the first and second halves of a stage, we apply tests to both halves separately.

Result 2: For r = 0, subjects put an equal weight on both signals consistent with Bayesian rationality.

In Stage "r=0", there is no significant difference between group specific estimated weights and 0.5. Two-tailed Wilcoxon matched pairs signed rank tests yield p-values of 0.77. Since individual decisions are independent in this first stage, we also performed tests on estimated individual weights.[16] The hypothesis that individual weights are distributed around 0.5 cannot be rejected either (the p-value is 0.70).

Our main result is that for $r > 0$, subjects put a lower weight on public signals than in equilibrium.

Result 3: For 0 < r < 1, subjects tend to put larger weights on public than on private signals, but the difference is smaller than theoretically predicted. There is no trend towards equilibrium.

Our prior expectation (arising from the widespread success of cognitive hierarchy models in beauty contest games) was that estimated group-specific weights on the private signal are smaller than 0.5 and larger than equilibrium weights. Therefore, we use one-sided tests on the null hypotheses that the median weight is equal to the respective equilibrium value against the alternative hypotheses that it is smaller than 0.5 or exceeds the equilibrium weight, respectively. One-tailed Wilcoxon matched pairs signed rank tests yield the p-values, reported in Table 5. While most tests reject the null hypotheses at 5%, we cannot reject that the median weight equals 0.5 for the first half of Treatment "r=0.25".

Null hypothesis	Alternative hypothesis	r=0.5, 1st half	r=0.5, 2nd half	r=0.25, 1st half	r=0.25, 2nd half	r=0.75, 1st half	r=0.75, 2nd half
$\gamma=0.5$	$\gamma < 0.5$	0.015	0.015	0.106	0.030	0.015	0.015
$\gamma=0.429$	$\gamma > 0.429$			0.015	0.015		
$\gamma=0.333$	$\gamma > 0.333$	0.015	0.015				
$\gamma=0.2$	$\gamma > 0.2$					0.015	0.031

Table 5 – P-values of one-tailed Wilcoxon matched pairs signed rank tests.

[16] Average weights per individual are estimated by a similar procedure as average weights per session, with a separate regression for each subject.

Comparing the weights assigned to private signals in the first and second halves of each treatment, we find no significant differences. Using the two-tailed Wilcoxon matched pairs test, p-values range between 15% and 85%. In particular, for Treatment "*r*=0.75", the p-value is 85%. Thus, we conclude that subjects tend to put larger weights on public than on private information[17], but the difference is smaller than theoretically predicted and there is no systematic trend over time.

Result 4: For r = 1, subjects assign larger weights to public than to private information: they lean towards coordinating on the public signal, but do not achieve full coordination during the course of the stage.

In Stage "*r*=1", all estimated weights are positive and below 0.5. Hence, the Wilcoxon matched pairs signed rank test yields p-values below 1% for the null hypotheses $\gamma = 0$ and $\gamma = 0.5$. Testing whether the weights assigned to private signals in the last 5 rounds were equal to those in the first half against the alternative of a systematic trend towards equilibrium (i.e. lower weights in the last rounds), the one-tailed Wilcoxon matched pairs test yields p-values of 0.02%. Thus, we can reject the hypothesis of no trend towards equilibrium for this stage.

Result 5: Over all sessions, the weight assigned to the private signal tends to decrease in r as predicted by theory.

A one-tailed Wilcoxon matched pairs signed rank test finds that γ is getting significantly smaller with increasing *r*, if we compare "*r*=0" with "*r*=0.5" or "*r*=0.5" with "*r*=1". P-values are always below 4%. Note that here the difference in *r* is 0.5. Comparing treatments where the difference in *r* is only 0.25, test results are mixed: comparing "*r*=0.5" with "*r*=0.25" or "*r*=0.5" with "*r*=0.75", one-tailed Mann-Whitney tests yields p-values below 3% except for "*r*=0.5" (first half) compared to "*r*=0.75" (second half), where p=6.6%. However, the one-tailed Wilcoxon matched pairs fails to reject the hypothesis of equal weights for Treatment "*r*=0.25" versus "*r*=0" and for Treatment "*r*=0.75" versus "*r*=1" (p>9%) except for Treatment "*r*=0.75" (second half) compared to "*r*=1" (second half), where p=3.1%. Figure 2 relates the average weight on the private signal for each treatment to theoretical predictions and also displays how the weight depends on *r*.

[17] This result seems to be corroborated by subjects' written comments in the post-experimental questionnaire (this questionnaire is presented in Appendix A.3). 43% of subjects explicitly wrote (without being directly asked) that y is more informative than x_i on the other participant's decision. The percentage of subjects reporting such statements is positively correlated to the value of *r* in stage 3.

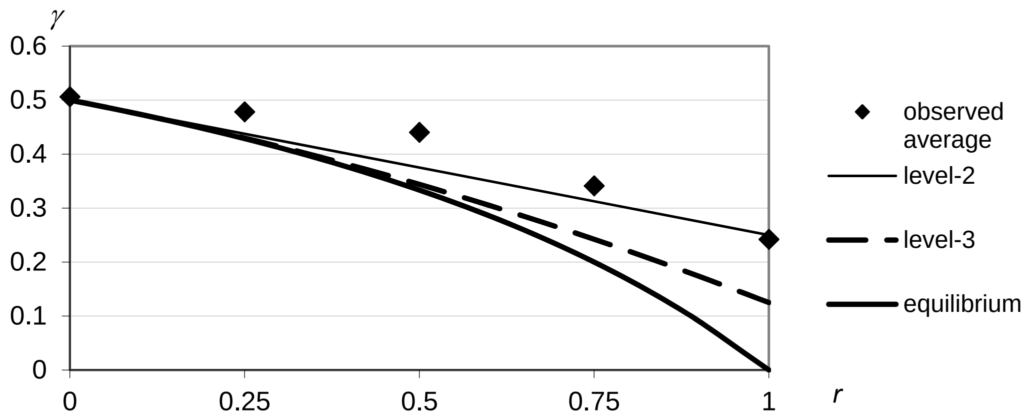


Figure 2 – Weight on the private signal for different values of *r*

5. Alternative explanations for observed behavior

We know from Nagel (1995), Kübler and Weizsäcker (2004) and many other experimental papers that the equilibrium may be a poor predictor for actual behavior when it requires infinite levels of reasoning. In their environments, cognitive-hierarchy models with limited levels of reasoning yield better predictions.[18] In this section, we study two models of bounded rationality: one is a model of a cognitive hierarchy (limited levels of reasoning) and the other one is based on non-Bayesian higher-order beliefs. Both models are able to explain weights on the private signal that exceed the equilibrium values.[19]

5.1. Cognitive hierarchy

The model of cognitive hierarchies assumes that agents process information rationally, but fail to consider that others decide with the same degree of rationality. Nagel (1995) and Stahl and Wilson (1994) define level-0 types as subjects who choose an action randomly from a uniform distribution over all possible actions. For $k > 0$, a level-k type is playing best response to level-$k - 1$. In our game, if the partner's action has a uniform distribution on the reals, expected losses from the coordination part of the payoff function are infinite for any action, so that any action is a best reply as long as $r > 0$. However, expected losses resulting from deviations from the fundamental are minimized by choosing the expected state. Hence, we define the best response to a uniform distribution over all reals as

[18] Shapiro, Shi, and Zillante (2009) analyze the predictive power of level-*k* reasoning in a game that combines features of Morris and Shin (2002) with Nagel (1995). They try to identify whether individual strategies are consistent with level-*k* reasoning. They argue that the predictive power of level-*k* reasoning is positively related to the strength of the coordination motive and to the symmetry of information. Their experiment however treats public and private signals in an asymmetric way by normalizing public signals.

[19] Hypotheses that arise from combining non-Bayesian higher-order beliefs with limited levels of reasoning are derived and tested in Appendix D.

$a_i^1 = E_i(\theta)$, so that level-1 players place weight 0.5 on both signals. Note that $a_i^1 = E_i(\theta)$ is the best response of a player who ignores the strategic part of the payoff function. Starting from this definition, actions for higher levels of reasoning can be calculated as follows: suppose that a player *j* attaches weight γ_k to her private signal. The best response to such behavior is

$$a_i^{k+1} = (1-r)E_i(\theta) + rE_i(a_j) = (1-r)E_i(\theta) + r\gamma_k E_i(x_j) + r(1-\gamma_k)y.$$

Since the expected private signal of the other player equals the expected state,

$$a_i^{k+1} = \left[(1-r) + r\gamma_k\right]E_i(\theta) + r(1-\gamma_k)y = \frac{(1-r)+r\gamma_k}{2}x_i + \left[\frac{(1-r)+r\gamma_k}{2} + r(1-\gamma_k)\right]y.$$

Hence, the next level of reasoning is $\gamma_{k+1} = \frac{(1-r)+r\gamma_k}{2}$. With higher levels, the weight on the private signal decreases from 0.5 towards the equilibrium value.

Table 6 summarizes the resulting weights on the private signal for increasing levels of reasoning and the equilibrium weight on the private signal γ^* for different values of *r*. Comparing observed weights with these theoretical values allows testing the cognitive hierarchy model.

Value of r	*Level-1*	*Level-2* γ_2	*Level-3* γ_3	*Level-4* γ_4	...	*Equilibrium weight (infinite levels of reasoning)* γ^*
0	0.50	0.50	0.50	0.50	…	0.50
0.25	0.50	0.437	0.429	0.429	…	0.429
0.5	0.50	0.375	0.344	0.336	…	0.333
0.75	0.50	0.312	0.242	0.216	…	0.2
1	0.50	0.25	0.125	0.062	…	0

Table 6 – Theoretical weight on x_i depending on *r* and on the level of reasoning

Result 6: For interior values of r, most estimated group weights on private signals are close to or above the weights given by level-2 reasoning.

As can be seen from Figure 2, the observed average weights are close to the weights of level-2 reasoning for all values of *r*. Comparing session-specific weights presented in Table 4 with weights from limited levels of reasoning displayed in Table 6, shows that in Treatments "*r*=0.5" and "*r*=0.25", average group weights are higher than the weights from level-2 reasoning with only one exception for *r*=0.5 (Session 5, first half). The hypothesis that average weights are equal to or below those from level-2 reasoning is rejected by one-tailed Wilcoxon matched pairs signed rank test at p-values below 5%. In Treatment "*r*=0.75", session-specific weights are distributed around the weight from level-2, and we cannot reject that the weights are equal to or below those from level-2 reasoning. In this treatment, one-sided tests reject level-3 reasoning (observed weights are not lower than γ_3), but two-sided tests do not (equality of observed weights with γ_3 cannot be rejected).

For r=1, second half, the hypothesis $\gamma \leq \gamma_2$ cannot be rejected either. These inhomogeneous test results indicate that subjects might behave in accordance to higher levels of reasoning with increasing r.

As Cornand and Heinemann (2012) show, the welfare-detrimental effects of public information require that agents put a weight on the private signal that is smaller than the one associated with level-2 reasoning. Since we cannot reject this for $r \geq 0.75$, we cannot rule out that for a strong coordination motive (high value of r), public information may be detrimental to welfare, if welfare depends only on how close actions are to the fundamental as in Morris and Shin (2002).

5.2. Non-Bayesian higher-order beliefs

Weights on the private signal above equilibrium values can also be explained by systematic mistakes in the formation of higher-order beliefs. For using information rationally, a subject has to form conditional expectations about his partner's beliefs based on his own information. Theoretically, subjects should put a weight of 0.5 on the private signal when estimating θ and a weight of 0.25 in estimating the other's estimation of θ. This does not require any assumptions about others' behavior apart from Bayesian rationality. If subject j has rational first-order beliefs, then $E_j(\theta) = 0.5 x_j + 0.5 y$ and rational second-order beliefs are $E_i(E_j(\theta)) = 0.5 E_i(\theta) + 0.5 y = 0.25 x_i + 0.75 y$. More generally, if $E_j(\theta) = \lambda x_j + (1-\lambda) y$, then $E_i(E_j(\theta)) = \lambda E_i(\theta) + (1-\lambda) y = 0.5\lambda x_i + (1-\lambda/2) y$.

Previous studies have explored whether subjects update beliefs according to Bayes' rule, when information arrives sequentially. Various studies from the 1960s found that "*[u]pon receipt of new information, subjects revise their posterior probability estimates in the same direction as the optimal model [Bayes' rule], but the revision is typically too small*" (Slovic and Lichtenstein, 1971, 693). Recent studies explore whether subjects are Bayesian in updating self assessments. While some find significant deviations from Bayes' rule (Eil and Rao, 2011), others report that subjects update overconfident prior beliefs correctly (Grossman and Owens, 2010). To our knowledge there is no previous experimental study exploring whether subjects use Bayes' rule in forming higher-order beliefs. Dale and Morgan (2012) test some predictions of Morris and Shin (2002) comparing treatments with private and public signals of different precisions. The weights that subjects put on signals of low precision and on the prior (which also has a low precision for indicating the fundamental) are much higher than under Bayesian belief formation. Since in their experiment public signals always have a low precision, they find higher weights on the public signal than we do. They find weights that are distributed around the equilibrium values and negative welfare effects of public signals.[20] In our experiment, all signals have the same precision, so that we can detect mistakes in the

[20] In the light of our previous impression that the level of reasoning may be correlated with r, the high parameter of r=0.8, chosen by Dale and Morgan (2012), may contribute to higher weights on public signals. They also inform subjects about the best response after each round which may enforce convergence to equilibrium.

formation of higher-order beliefs without this potentially confounding bias towards low quality signals.

Result 7: Subjects attach too high a weight to the private signal when predicting other subjects' beliefs about fundamentals.

Data from belief elicitation reveal that subjects attach (on average) a weight of 0.463 on the private signal when estimating θ and an average weight of 0.291 when estimating their partner's estimation of θ. Two-tailed Wilcoxon tests on estimated group-specific weights reject the hypotheses that the median weight in estimating θ equals 0.5 (p=0.005) in favor of lower weights on private signals, and reject that the median weight in estimating the other player's estimation equals 0.25 (p=0.038) in favor of higher weights on private signals. 229 out of 288 subjects attribute a weight higher than 0.25 to the private signal in this task. Although Bayesian rationality requires that the weight on the private signal in second-order beliefs be half of the weight in estimating θ, it is in fact 62.2% and thereby too high to be consistent with Bayesian rationality. This can explain a part of the differences between observed weights and equilibrium predictions in the other stages. As we elicited beliefs directly after Stage 3 ($0 < r < 1$), we may expect that higher-order beliefs in the later periods of Stage 3 are formed in about the same way as elicited beliefs. However, as we show in Appendix C, the deviations from Bayesian rationality in forming higher-order beliefs are too small to explain the whole difference between observations and equilibrium predictions in the second half of Stage 3.

It is surprising, though, that subjects underused private signals in forming their expectation of θ, especially in light of the results from Stage 1 ("r=0") that is equivalent to asking first-order expectations. We attribute this to an order effect: after 40 periods in Stages 2 and 3, in which the public signal was more important than the private, some subjects may have developed a routine in putting higher weights on public signals that was applied to both tasks in Stage 4. In Appendix C, we present the results of 6 control sessions, in which we elicited beliefs directly after Stage 1 and before all stages, in which common signals are theoretically more important than private ones. Here, we find that the weights on the private signal in first-order expectations are distributed around 0.5 (equality cannot be rejected) while the weights in second-order expectations are all above 0.3 and on average 83% of the weight in first-order expectations. Here, the violation of Bayesian rationality in forming higher-order beliefs is even more pronounced. These results indicate that subjects underestimate how informative public signals are for estimating other players' beliefs.

6 – Conclusion

Our experiment provides convincing data for the hypothesis that subjects attach larger weights to public than to private signals if they have incentives to coordinate their actions. The experiment also investigates whether players put higher weights on the public signal if the relative importance of the coordination component is increased. We find clear evidence in line with this qualitative prediction of equilibrium theory. The observed weights are, however, lower than equilibrium values. They are closer to the weights arising from level-2 reasoning of a cognitive hierarchy model, for which Cornand and Heinemann (2012) show that public information cannot be welfare detrimental.

References

Alós-Ferrer, C., and C. Kuzmics (2008), "Hidden Symmetries and Focal Points," University of Konstanz, TWI Research Paper no. 35.

Anctil, R., J. Dickhaut, C. Kanodia, and B. Shapiro (2004), "Information Transparency and Coordination Failure: Theory and Experiment," *Journal of Accounting Research*, 42, 159-195.

Angeletos, G.-M., and A. Pavan (2004), "Transparency of Information and Coordination in Economies with Investment Complementarities," *American Economic Review* (Papers and Proceedings), 94, 91-98.

Angeletos, G.-M., and A. Pavan (2007a), "Socially optimal coordination: Characterization and policy implications," *Journal of the European Economic Association* (Papers and Proceedings), 5, 585-593.

Angeletos, G.-M., and A. Pavan (2007b), "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103-1142.

Baeriswyl, R., and C. Cornand (2011), "Reducing overreaction to central banks' disclosures : theory and experiment," *Working Paper du GATE Lyon Saint-Etienne*, 41.

Cornand, C. (2006), "Speculative Attack and Informational Structure: An Experimental Study," *Review of International Economics*, 14, 797-817

Cornand, C., and F. Heinemann (2008), "Optimal Degree of Public Information Dissemination," *The Economic Journal*, 118, 718-742.

Cornand, C., and F. Heinemann (2012), "Higher-Order Beliefs and the Social Value of Public Information," mimeo.

Dale, Donald J., and John Morgan (2012), "Experiments on the Social Value of Public Information," mimeo.

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, 114-138.

Forsythe, R., T.R. Palfrey, and C.R. Plott (1982), "Asset Valuation in an Experimental Market," *Econometrica*, 50, 537-567.

Grossman, Z., and D. Owens (2010), "An Unlucky Feeling: Persistent Overestimation of Absolute Performance with Noisy Feedback," University of California at Santa Barbara, *Economics Working Paper* No. 1937484.

Hanson, R. (1996), "Correction to MacKelvey and Page, "Public and Private Information: An Experimental Study of Information Pooling"," *Econometrica*, 64, 1223-1224.

Heinemann, F., R. Nagel, and P. Ockenfels (2004), "The Theory of Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information," *Econometrica*, 72, 1583-1599.

Hellwig, C. (2005), "Heterogeneous Information and the Welfare Effects of Public Information Disclosures," UCLA, mimeo.

Kübler, D., and G. Weizsäcker (2004), "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory," *Review of Economic Studies*, 71, 425-442.

Lorenzoni, G. (2010), "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77, 305-338.

McKelvey, R. D., and P. C. Ordeshook (1985), "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources," *Journal of Economic Theory,* 35, 55-85.

McKelvey, R. D., and T. Page (1990), "Public and Private Information: An Experimental Study of Information Pooling," *Econometrica*, 58, 1321-1339.

Morris, S., and H.S. Shin (2002), "Social Value of Public Information," *American Economic Review*, 92, 1522-1534.

Morris, S., and H.S. Shin (2004), "Coordination Risk and the Price of Debt," *European Economic Review*, 48, 133-153.

Morris, S., H. S. Shin, and H. Tong (2006), "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply," *American Economic Review*, 96, 453-55.

Myatt, D.P., and C. Wallace (2008), "On the sources and value of information: public announcements and macroeconomic performance," University of Oxford, Dept. of economics discussion paper no. 411.

Nagel, R. (1995), "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313-1326.

Plott, C.R., and S. Sunder (1988), "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085-1118.

Roca, M. (2010), "Transparency and Monetary Policy with Imperfect Common Knowledge," IMF Working Paper 10/91.

Romer, C., and D. Romer (2000), "Federal Reserve Private Information and the Behavior of Interest Rates," *American Economic Review*, 90, 429-457.

Shapiro, D., X. Shi, and A. Zillante (2009), "Robustness of Level-*k* Reasoning in Modified Beauty Contest Games," mimeo.

Slovic, P., and S. Lichtenstein (1971), "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," *Organizational Behavior and Human Performance*, 6, 649-744.

Stahl, D. O., and P.W. Wilson (1994), "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization*, 25, 309-327.

Svensson, L.E.O. (2006), "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con," *American Economic Review*, 96, 448-452.

Woodford, M. (2005), "Central Bank Communication and Policy Effectiveness," in: *The Greenspan Era: Lessons for the Future*, Kansas City: Federal Reserve Bank of Kansas City, 399–474.

Zeiliger, R. (2000), "A Presentation of Regate, Internet based Software for Experimental Economics," http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt, GATE.

Appendix

A. Instructions and questionnaires

A.1. Instructions

Instructions to participants varied between sessions according to the different treatments in Stage 3. We present the instructions for a session with treatment r=0.25. For the other sessions, instructions were adapted accordingly and are available upon request.[1]

Instructions

General information

Thank you for participating in an experiment in which you can earn money. These earnings will be paid to you in cash at the end of the experiment. We ask you not to communicate from now on. If you have a question, then raise your hand and the instructor will come to you.

Framework of the experiment

You are 16 persons participating in this experiment. The experiment consists of 4 stages, the first one including 5 situations, the second one 10, the third one 30 and the fourth one 5. In each situation you will be randomly matched with one of the other 15 participants. You will not get to know with whom you are matched. The rules are the same for all participants. Situations are independent and in each of them, you will have to take a decision.

RULES COMMON TO ALL STAGES

Decision situation

In each situation you will be randomly matched with one of the other participants.

For each situation a number called Z is drawn randomly from the interval 50 to 450. This number is the same for both of you. All numbers in the interval [50, 450] have the same probability to be drawn. When you make your decision, you will **not** know the drawn number Z.

However, you will be receiving two hints (numbers) on Z:

- You and the person with whom you are matched, both receive a common hint number Y for the unknown number Z. This common hint number is randomly selected from the interval [Z-10, Z+10]. All numbers in this interval are equally likely. This common hint number Y is the same for both of you.

---

[1] What follows is a translation (from French to English) of the instructions and the questionnaire given to the participants.

- In addition to the common hint number, each participant receives a private hint number *X* for the unknown number *Z*. The private hint numbers are also randomly selected from the interval [*Z*-10, *Z*+10]. All numbers in this interval have the same probability to be drawn. Your private hint number and the private hint number of the person whom you are matched are drawn independently from this interval, so that (in general) you will not get the same private numbers.

RULES OF THE 1ST STAGE (5 situations)

You will be asked to make a decision by choosing some number.

Your payoff positively depends on the proximity between your decision and the true value of the unknown number *Z*:

$$\textit{Payoff in ECU} = 100 - \left(\textit{your decision} - Z \right)^2 .$$

This means that your payoff only depends on how close is your decision to the true value *Z* and not on your partner's decision.

Once you have made a decision, click on the OK-button. Once all participants made their decision for the game, a situation is terminated.

RULES OF THE 2ND STAGE (10 situations)

Again, you will be asked to make a decision by choosing some number.

The rules are the same as in the first stage, but here your payoffs are given by:

$$\textit{Payoff in ECU} = 100 - \left(\textit{your decision} - \textit{the other participant's decision} \right)^2 .$$

This means that your payoff only depends on how close is your action to the action of the other participant and not on the unknown number *Z*.

RULES OF THE 3RD STAGE (30 situations)

Again, you will be asked to make a decision by choosing some number.

The rules are unchanged, but here your payoffs depend positively on the one hand on the proximity between your decision and the unknown number *Z* and on the other hand on the proximity between your decision and the choice of your partner.

$$\textit{Payoff in ECU} =$$

$$400 - 3.\left(\textit{your decision} - Z\right)^2 - 1.\left(\textit{your decision} - \textit{the decision of the other participant}\right)^2 .$$

This formula says that your payoff in each situation is at most 400 ECU. It is reduced for deviations of your decisions from the unknown number *Z*, and it is also reduced for deviations between your and your partner's decision. The closer is your decision to both the other participant's choice and *Z*, the higher will be your payoff.

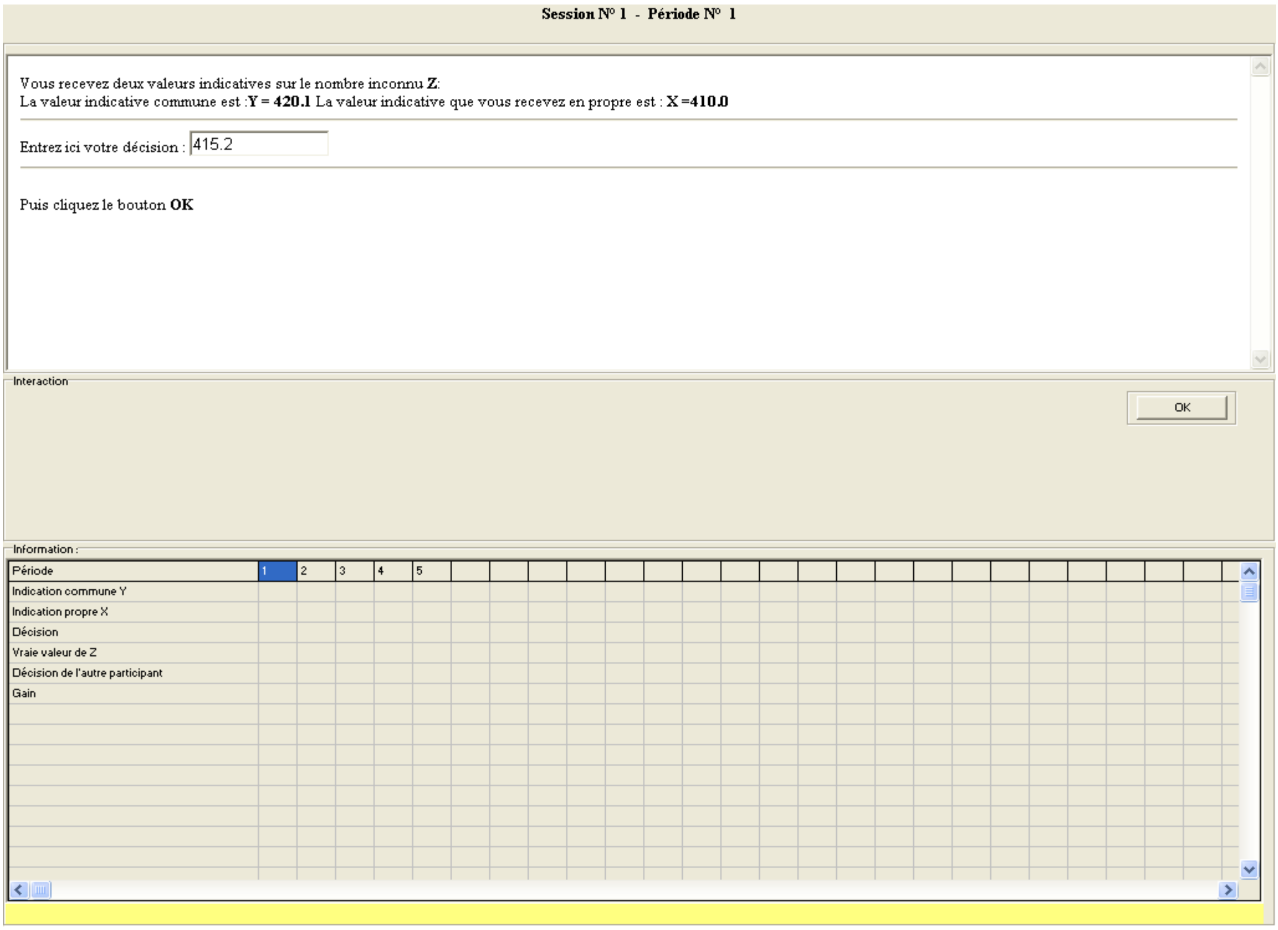
Note that your payoff depends more on the spread between your decision and *Z* than on the spread between your decision and the decision of the other participant.

Example of decision phase:

You receive two hint numbers: One hint is common for the two participants, the other is your private hint. Both hint numbers are drawn with uniform distribution from [*Z*-10, *Z*+10].

The common hint number *Y* is: 420.1.

Your private hint number *X* is: 410.0.



You enter your decision, for example here: 415.2.

Information after each situation

Each participant will be informed after each situation on

(1) the common hint *Y* and his private hint *X*,

(2) the true number *Z*, the true choice of his pair-mate,

(3) his own choice and his related payoff.

Example of information phase:

In the previous example, your own choice was 415.2.

Suppose the true value of *Z* is 419.4. Your pair-mate's decision was 413.7.

Then, your payoff will be:

Payoff = $400 – 3.(415.2 – 419.4)^2 – 1.(415.2 – 413.7)^2 = 344.83$ ECU.

RULES OF THE 4^{TH} STAGE (5 situations)

In this stage, rules are a bit different.

Again, *Z* is unknown and you receive two hint numbers, *X* (your private hint) and *Y* (the common hint).

Now, you are asked to choose two numbers:

(1) give what you think is *Z*,

(2) give what you think the other participant think about *Z*.

Your payoff positively depends on the one hand on the proximity between your estimation on *Z* and the true value of *Z* and on the other hand between your estimation of the estimation of the other participant on *Z* and the true estimation of the other participant on *Z*.

Your payoff is given by:

$$100 - (your\ estimation\ sur\ Z - Z)^2$$
$$-(your\ estimation\ of\ the\ estimation\ of\ the\ other\ participant\ on\ Z - the\ estimation\ of\ the\ other\ participant\ on\ Z)^2$$

The closer your estimations are to true values, the higher your payoff.

You will be told about each change in stage.

Questionnaires:

At the beginning of the experiment, you will be asked to fill in an understanding questionnaire on a paper. Afterward, the experiment will begin. At the end of the experiment you will fill in a "personal" questionnaire on the computer. All information will remain secret.

Payoffs:

Also at the end of the experiment the ECUs you have obtained are converted into Euros and paid in cash. 1 ECU corresponds to 0.25 Cents.

If you have any questions, please ask them at this time.

Thanks for your participation!

A.2. Understanding and training questionnaire

Fill in

- In each situation, you interact with ________ other participant(s).
- You receive in each situation ________ hints.
- The difference between the unknown number *Z* and any hint is at most _______.

Yes or no

- At stage 3, when a participant makes a decision, does his payoff depend on the decision of his pair-mate? ____
- Do pair-mates receive the same hints? _____
- Is there a hint that is more precise than another? ______
- Do you play with the same participant during the whole length of the experiment? _____

Practice

You are at the third stage of the experiment.

You receive *Y*=135 and *X*=141.

Among the next statements, choose the right one(s):

- o The true value of *Z* is between 125 and 151.
- o The true value of *Z* is 135.
- o The true value of *Z* is between 131 and 145.

Suppose that the true value of *Z* is 143 and that the true decision of your pare-mate is 133. What is your payoff (in ECU) if you choose 134?

Now, suppose again that the true value of *Z* is 143 and the true decision of your paire-mate is 133. What is your payoff (in ECU) if you choose 138?

A.3. Post-experimental questionnaire

1. How did you make a decision? On which criteria?

2. During the first 3 stages, have you tried to guess the value of *Z*? And the value of the decision of the other participant?
3. Do you think that one of the two indicative hints (private *versus* common) was more informative than the other on *Z*? And on the decision of the other participant?
4. Did you take into account the two indicative hints in the same manner? Or more your private hint? Or more the common hint?

B. Control sessions

B.1. Design of control sessions

We conducted 6 control sessions. Table B1 gives an overview over control sessions.

Sessions	*Stage 1*	*Stage 2*	*Stage 3*	*Stage 4*	*Stage 5*
19 to 21	"*r*=0" *(5 periods)*	*belief elicitation (5 periods)*	*Treatment "r=1 with 1 common signal" (10 periods)*	"*r*=0.5" (*20 periods*)	*"r=1 with 2 signals" (10 periods)*
22 to 24			*Treatment "r=1 with 1 private signal" (10 periods)*		

Table B1 – Composition of the different control sessions

Control sessions are motivated by the results from standard sessions presented in the main text. We especially wanted to elicit higher order beliefs at an earlier stage before subjects were playing strategic games. We also wanted to test directly how public or private signals affect payoffs in a pure coordination game. In control sessions, we started with a first stage where "*r*=0" as in standard sessions. In Stage 2, we proceeded with directly eliciting beliefs (as we did in the last stage of standard sessions). Here, we used two separate payoff functions, $100-(e_i(\theta)-\theta)^2$ and $100-(e_i(e_j(\theta)-e_j(\theta))^2$ instead of one unified as in standard sessions. We did these changes, because we were concerned that results from the last stage in standard sessions were distorted by spillovers from the previous stages and because the game might be easier to understand with separate payoff functions.[2] In Stage 3, we had two distinct treatments: "*r*=1 with 1 common signal" and "*r*=1 with 1 private signal"; both treatments had the same payoff function as in Stage 5, where subjects got a private *and* a common signal. In Treatment "*r*=1 with 1 common signal" they received just a common and no private signal, in Treatment "*r*=1 with 1 private signal" they received just private signals. Here, they could choose any number in [x_i – 20, x_i + 20], but the largest observed difference between choice and signal was 14.2. We used these treatments to analyze subjects' ability to coordinate when they receive only one signal and to compare the welfare effects of adding a second

[2] In addition, Stage 1 ("*r*=0") and Stage 2 (belief elicitation) are both tests of Bayesian rationality, though with a different framing. Proceeding with these stages one directly after another allows for a better comparison of their results.

signal in Stage 5. Stage 4 ("*r*=0.5") contained the same game as Treatment "*r*=0.5" in standard sessions and was meant to check robustness of the results. In the control sessions, we divided the 16 participants into 2 matching groups of 8, thereby gathering two independent observations per session.[3] 100 points were converted to 40 cents in sessions 19 – 24.[4]

B.2. Analyzing data from control sessions

As for standard sessions, we use the following structure in analyzing data. First, we analyze whether subjects' choices are between their two signals as predicted by the equilibrium theory. Then we give summary statistics for the weights that subjects attach to the private signal. We investigate whether observed weights on the private signal are positive, smaller than 0.5, and whether they deviate systematically from the theoretical prediction. We also test whether there is a time trend and we test comparative static predictions arising from theory.

B.2.1. Some considerations about rationality

As for standard sessions, we have a look at subjects' rationality. The next table displays the percentage of all choices within the intervals I_1 to I_3 as defined in the main text. Counting the number of choices that are closer to the common signal, closer to the private, or in the middle provides us a crude first impression of whether subjects put a larger weight on the common signal and how dispersed the distribution of relative weights is.

Value of r	*0*	*0.5*	*1*
Inside $I_1 = [\min(y,x_i), \max(y,x_i)]$	*75%*	*86%*	*94%*
Choice = y	*1%*	*4%*	*61%*
Closer to y	*28%*	*35%*	*15%*
Middle (+/-0.05)	*15%*	*18%*	*6%*
Closer to x_i	*31%*	*28%*	*12%*
Choice = x_i	*0%*	*1%*	*1%*
Outside I_1 *beyond y*	*11%*	*7%*	*3%*
Outside I_1 *beyond* x_i	*14%*	*7%*	*3%*
Inside $I_2 = [\max(Y,X)-10; \min(y,X)+10]$	*93%*	*97%*	*84%*
Inside $I_3 = I_1 \cup I_2$	*96%*	*99%*	*99.6%*

Table B2.1 – Crude classification of choices, control sessions 19-24

Comparing control sessions with standard sessions, the only remarkable difference is that in Stage 5 ("*r*=1") of control sessions there are many more choices equal to the public signal than in Stage 2 ("*r*=1") of standard sessions. We attribute this to the different order of stages. Experience from other stages seems to make it easier for subjects to coordinate on the public signal when coordination is the only motive for action. Result 1 still holds for control sessions. In Appendix E below we provide some more details on order effects.

[3] Subjects were randomly matched with other subjects from the same matching group only. Subjects were not informed about the size of matching groups.

[4] In sessions 22-24, 15 of 48 subjects earned negative payoffs in Treatment "*r*=1 with 1 private signal". Here, it took up to 8 periods to compensate these losses.

B.2.2. Estimated weights on the private signal

Session, group	"*r* = 0"	"*r* = 1 with 2 signals", 1st half	"*r* = 1 with 2 signals", 2nd half	"*r* = 0.5", 1st half	"*r* = 0.5", 2nd half
19, group 1	0.455	0.181	0.107	0.502	0.484
19, group 2	0.583	0.149	0.201	0.367	0.338
20, group 1	0.524	0.008	0.000	0.431	0.473
20, group 2	0.583	0.255	0.278	0.551	0.536
21, group 1	0.446	0.234	0.229	0.476	0.422
21, group 2	0.534	0.134	0.039	0.495	0.540
22, group 1	0.527	0.130	0.105	0.439	0.454
22, group 2	0.500	0.482	0.421	0.530	0.485
23, group 1	0.506	0.098	0.150	0.459	0.455
23, group 2	0.570	0.184	0.067	0.486	0.459
24, group 1	0.582	0.132	0.160	0.489	0.440
24, group 2	0.489	0.385	0.359	0.519	0.487
Average (19-24)	0.525	0.198	0.176	0.479	0.464
St.dev. (19-24)	0.048	0.129	0.128	0.050	0.053
Equilibrium weight	0.5	0	0	0.333	0.333

Table B2.2 – Group specific weights on the private signal

Table B2.2 displays the estimated group-specific weights on the private signal for control sessions. Estimates follow the same procedure as laid out in Section 4.2 for standard sessions. We tested standard sessions and control sessions separately, because they are not entirely comparable. Nonparametric tests for control sessions are reported below. We show that most results of standard sessions also hold for control sessions. We comment on the similarities and differences below.

Result 2 for control sessions also holds: for $r = 0$, subjects put an equal weight on both signals consistent with Bayesian rationality. In Stage 1 ("*r*=0"), there is no significant difference between group specific estimated weights and 0.5. Two-tailed Wilcoxon matched pairs signed rank tests yield p-values of 0.12 for control sessions. A Mann-Whitney test does not reject the hypothesis that weights in standard sessions have the same distribution as those in control sessions (p-value 0.17). Since individual decisions are independent in this stage, we also performed tests on individual weights from Regression (i). The hypothesis that individual weights are distributed around 0.5 cannot be rejected (the p-value is 0.45 for control sessions). Neither can the hypothesis be rejected that individual weights in the two groups of sessions come from the same distribution (p-value 0.65).

Result 3 also holds for control sessions: for $r = 0.5$, subjects tend to put larger weights on public than on private signals, but the difference is smaller than theoretically predicted. There is no trend towards equilibrium. One-tailed Wilcoxon matched pairs signed rank tests reject that the weight on the private signal equals the equilibrium value in favor of a higher weight ($p < 0.01$). For data from the second half of Stage 4 ("*r*=0.5") we can reject that weights equal 0.5 in favor of smaller weights ($p = 0.013$). For the first half of Stage 4 ("*r*=0.5"), however, this hypothesis cannot be rejected ($p = 0.100$). For Stage 4, estimated weights are higher than for Treatment *r*=0.5 in standard sessions (significant at 1% for the first half, but insignificant (p=7.7%) for the second half, using two-tailed Mann-Whitney tests). We attribute this to an order effect. We provide more details about order effects in Appendix E below.

Result 4 from the main text has to be slightly adjusted for control sessions. Indeed, for "r = 1, two signals", subjects assign larger weights to public than to private information (p<1%) as in standard sessions. But contrary to standard sessions, in control sessions, there seems to be no trend towards improved coordination. Testing whether the weights assigned to private signals in the last 5 rounds were equal to those in the first half against the alternative of a systematic trend towards equilibrium (i.e. lower weights in the last rounds), the one-tailed Wilcoxon matched pairs test yields p-value of 12.8% for control sessions. Thus, we cannot reject the hypothesis of no trend towards equilibrium for control sessions. The latter becomes even more pronounced if we use a sign test instead (p-value 38%). Comparing behavior between the two groups of sessions, we find that in the first 5 rounds, weights in control sessions are significantly lower than in standard sessions (two-tailed Mann-Whitney, p<1%). In the second half, there is no significant difference (p=12%). We attribute this to an order effect explained below in Appendix E. We conclude that in the extreme case of a pure coordination game, subjects condition their choices on their private signals, which prevents full coordination. Private signals matter in a pure coordination game and may be welfare reducing. One of the reasons for control sessions was to include games with only one signal, in order to test the welfare reducing effect of private signals directly by comparing payoffs. These results are reported in Appendix B.3. below.

Finally, Result 5 holds also for control sessions: over all sessions, the weight assigned to the private signal tends to decrease in r as predicted by theory. A one-tailed Wilcoxon matched pairs signed rank finds that γ is significantly smaller in treatments with higher r, if we compare "r=0" with "r=0.5" or "r=0.5" with "r=1 with 2 signals". P-values are always below 4%.

B.3. Payoff effects of providing additional information in a pure coordination game

In a pure coordination game as with r=1, private signals should be neglected if actions can be conditioned on public signals. In equilibrium, additional private signals should not affect welfare. One purpose of control sessions was to provide a direct answer to the question of whether providing additional private signals reduces welfare. Table B3 compares average payoffs from games with "r=1".

In sessions 19-21, four groups achieved a lower payoff in with two signals than with one common signal. Hence, adding a private signal to the information structure reduced their average payoffs. Two groups (Session 20, Group 1, and Session 21, Group 2) improved their payoffs. The average over all groups is lower when both signals are provided, but the difference is not significant (p=0.56). This indicates that adding private information does not reduce average payoffs in pure coordination games. On the other hand, average payoffs for Stage "r=1" in standard sessions are significantly lower than in the comparable Stage "r=1 with 2 signals" and in Treatment "r=1 with 1 common signal" of Sessions 19-21 (Mann-Whitney, p<1%).

We attribute these diverse findings to the order of stages: "*r*=1" (with both signals) was the second stage in standard sessions, while it was the last in control sessions. In control sessions, it is possible that the high payoff when "*r*=1 with 2 signals" is due to learning from previous stages. Comparing data from Stage "*r*=1" of standard sessions with data from "*r*=1 with 1 common signal" from Sessions 19-21 might give a better impression of how additional private information affects behavior for a given state of experience with this kind of games. However, we have to admit that we cannot provide a definite answer to the question of whether theoretically irrelevant private information impedes coordination. An experiment by Fehr et al. (2011) is better suited to answer this question and finds convincing evidence for welfare reducing effects of intrinsically irrelevant private signals in a pure coordination game.[5]

Session, Group	*"r=1, with 1 common signal"*	*"r=1, with 1 private signal"*	*"r=1, with 2 signals"*	*Session*	*"r=1" (two signals)*
19, Group 1	97.53		89.53	1	86.22
Group 2	99.60		85.08	2	78.84
20, Group 1	99.82		99.90	3	71.26
Group 2	93.06		75.46	4	73.01
21, Group 1	92.49		88.99	5	82.81
Group 2	72.25		89.98	6	76.59
Average (19-21)	92.46		88.16	7	65.47
22, Group 1		- 3.56	90.06	8	81.68
Group 2		25.19	82.09	9	70.07
23, Group 1		24.23	95.06	10	68.59
Group 2		- 1.60	94.53	11	78.22
24, Group 1		8.73	89.33	12	88.29
Group 2		7.47	78.94	13	80.16
				14	70.74
				15	68.69
				16	63.30
				17	83.33
				18	82.16
Average (22-24)		10.08	88.33	Average (1-18)	76.08

Table B3 – Average payoffs in games with"*r*=1"

From Table B3, it is obvious for Sessions 22-24 that adding a public signal increased average payoffs compared to treatment "*r*=1 with 1 private signal".[6]

[5] In a context of asymmetric information, Camerer et al. (1989) show that more information is not always better because agents are unable to ignore private information even when it is in their interest to do so.

[6] When receiving only a private signal in Treatment "*r*=1 with 1 private signal", some subjects apparently tried to find a focal point that the experimental design did not allow for. About one participant per session asked why he could not enter either 50 or 0. He or she was told that these numbers were outside the range of admissible choices. In Sessions 19 to 21 this question never occurred, probably because subjects could coordinate on the public signal. Having allowed participants to choose from a fixed range of numbers might have helped them to coordinate on a focal point, disregard the private signal, and increase their payoffs. In a related experiment, Fehr et al. (2011) show that subjects tend to neglect imprecise private signals in a pure coordination game with a fixed choice set that includes prominent numbers. However, they also find that adding private signals impedes coordination.

B.4. Cognitive hierarchy in control sessions

As explained in the main text, we define levels of reasoning by assuming that level-0 players decide randomly and level-1 players place weight $\gamma_1 = 0.5$ on the private signal, so that they ignore the strategic part of the payoff function. Higher levels of reasoning are then given by

$$\gamma_{k+1} = \frac{(1-r) + r\gamma_k}{2}.$$

Comparing weights from limited levels of reasoning displayed in Table 6 in the main text with the detected weights in Table B2.2, shows that for "*r*=0.5", average group weights are higher than the weights from level-2 reasoning with only one exception (Session 19, Group 2).

For "*r*=1 with 2 signals", most groups achieve weights that indicate less than level-2 reasoning, 5 groups go beyond Level 3, and one of them fully coordinates on the public signal (Session 20, Group 1). Here, the level-2 hypothesis cannot be rejected (p-value 0.09). Level-3 reasoning cannot be rejected either (p=0.26), and we need to go to level 4 to find a significant difference.

C. Non-Bayesian higher-order beliefs

As explained in the main text, weights on the private signal above equilibrium values can also be explained by systematic mistakes in the formation of higher-order beliefs. Here, we provide a formal model of non-Bayesian higher-order beliefs and test it with the data from the experiment.

In the experiment we elicit higher order beliefs directly and measure the weight that subjects attach to the private signal when estimating their partner's beliefs. Denote this weight by λ. Based on this measure, we test a model of boundedly rational behavior that assumes infinite levels of reasoning but a systematic error in forming higher-order beliefs. More precisely, denote subject *i*'s expectation about a variable *x* by $e_i(x)$ and suppose:

1. Subjects respond optimally to their expectations about the state and about their partner's action, i.e.,

   $$a_i = (1-r)\, e_i(\theta) + r\, e_i(a_j), \quad j \neq i.$$

   Furthermore, this behavioral assumption is common knowledge among players, which amounts to assuming infinite levels of reasoning.

2. Subjects use private and common signals correctly to forecast the state of the world, *i.e.*

   $$e_i(\theta) = E_i(\theta).$$

3. Subjects make a systematic error in forming higher-order beliefs, such that

   $$e_i(e_j(\theta)) = \lambda\, x_i + (1-\lambda)\, y,$$

   $$e_i(e_j(e_i(\theta))) = \lambda^{3/2}\, x_i + (1-\lambda^{3/2})\, y,$$

$$e_i(e_j(e_i(e_j(\theta)))) = \lambda^2 x_i + (1-\lambda^2) y,$$

and so on, where λ is the weight measured in the experiment.

To understand the justification of these formulas, note that a rational first-order expectations are $E_i(\theta) = \alpha x_i + (1-\alpha) y$ with $\alpha = 0.5$. Rational higher-order expectations are then given by $E_i(E_j(\theta)) = \alpha^2 x_i + (1-\alpha^2) y$, $E_i(E_j(E_i(\theta))) = \alpha^3 x_i + (1-\alpha^3) y$, and so on. The estimated λ for second-order expectations replaces α^2 in the formula, so that the weight in third-order beliefs should be $\lambda^{3/2}$ and so on.

Combining these assumptions, we get

$$\begin{aligned}
a_i &= (1-r) e_i(\theta) + r e_i(a_j) \\
&= (1-r) E_i(\theta) + r e_i((1-r) e_j(\theta) + r e_j(a_i)) \\
&= (1-r) E_i(\theta) + (1-r)\left[r\left(\lambda x_i + (1-\lambda) y\right) + r^2\left(\lambda^{3/2} x_i + (1-\lambda^{3/2}) y\right) + r^3\left(\lambda^2 x_i + (1-\lambda^2) y\right) + r^4\left(\lambda^{5/2} x_i + (1-\lambda^{5/2}) y\right) + \ldots \right] \\
&= (1-r)\left(\frac{1}{2} + r\lambda \sum_{i=0}^{\infty} (r\sqrt{\lambda})^i\right) x_i + \left[1 - (1-r)\left(\frac{1}{2} + r\lambda \sum_{i=0}^{\infty} (r\sqrt{\lambda})^i\right)\right] y \\
&= (1-r)\left(\frac{1}{2} + \frac{r\lambda}{1 - r\sqrt{\lambda}}\right) x_i + \left[1 - (1-r)\left(\frac{1}{2} + \frac{r\lambda}{1 - r\sqrt{\lambda}}\right)\right] y. \qquad (1)
\end{aligned}$$

With λ decreasing from 0.5 to its rational value of 0.25, the weight that an agent attaches to the private signal in her decision decreases from a value below 0.5 towards the equilibrium value.

In Stage 2 of control sessions and Stage 4 of standard sessions, subjects were provided with public and private signals, but instead of being asked for an action, they had to state their beliefs (i) about the true state of the world and (ii) about their partner's stated belief about the state of the world. Thus, we directly elicit first-order and second-order beliefs. As explained in the main text, subjects should put a weight of 0.5 on x_i in estimating θ and a weight of 0.25 in estimating the other's estimation of θ.

If a subject attributes a higher [lower] weight to her private signals in predicting her partner's stated expectation, her weight on the private signals in Treatments "r=0.5", "r=0.25", and "r=0.75" should also be higher [lower] than the respective equilibrium weights. Suppose that a subject puts weight 0.312 on private signals when "r=0.75". Such behavior is consistent with level-2 reasoning, but could also be explained by a weight of λ=0.48 on private signals in higher-order beliefs according to equation (1). Direct elicitation of beliefs allows discriminating between the two theories.

Data from belief elicitation in standard sessions, summarized in Table C1, reveal that subjects attached (on average) a weight lower than 0.5 on the private signal when estimating θ and a weight higher than 0.25 when estimating their partner's estimation of θ.

For the rational second-order beliefs we assumed that first-order beliefs put weight 0.5 on either signal. In fact, data reveal that on average subjects attribute a weight of only 0.46 on their private signals. If subjects actually guessed this weight correctly, the optimal weight on the private signal in higher-order beliefs would be 0.23. Thus, the higher observed weights in second-order beliefs cannot be explained by subjects' responding to distorted first-order beliefs.

Since estimating θ does not depend on others' choices, we also test the distribution of individual weights. When we estimate weights separately for each subject, the average weight on x_i (over all subjects in standard sessions) in estimating θ is 0.471, while the average weight on x_i in forming higher-order beliefs is 0.292. We can reject the hypotheses that individual weights in estimating θ are distributed around 0.5 (p<1%) and that the weights in estimating the other player's estimation of the state are around 0.25 (p<1%). 229 out of 288 subjects attribute a weight higher than 0.25 to the private signal in this task. Individuals' higher-order expectations are not independent, but since expectations of θ are generally biased towards the public signal, any adjustment in higher-order beliefs should go in the same direction, which conflicts with our observations. Although Bayesian rationality requires that the second weight be half of the first, it is in fact 62.2% of the first. Hence, all evidence indicates that subjects underestimate the importance of public information in forming higher-order beliefs.

Sessions	*Weight in estimating θ*	*Weight in estimating the partner's estimation of θ*
1	0.394	0.228
2	0.520	0.342
3	0.481	0.376
4	0.418	0.170
5	0.424	0.210
6	0.482	0.159
7	0.493	0.335
8	0.429	0.301
9	0.532	0.379
10	0.466	0.272
11	0.500	0.375
12	0.501	0.217
13	0.439	0.271
14	0.432	0.263
15	0.520	0.420
16	0.425	0.345
17	0.437	0.277
18	0.446	0.292
Average	0.463	0.291
St.dev.	0.041	0.075

Table C1 – Group specific weights on the private signal in stated expectations of Sessions 1-18

It is surprising, though, that subjects underused private signals in forming their expectation of θ, especially in light of the results from Treatment "r=0", where subjects used (on average) a weight of 0.506. The difference in coefficients betweem Stage 1 "r=0" and stated first-order expectations in Stage 4 is significant at 0.1%. This may be caused by an order effect, such that after 40 periods in stages 2 and 3, in which the public signal was more important than the private, subjects underestimate the importance of the private signal for estimating θ during Stage 4. This, however, should also hold for forming expectations about others' expectations. Thus, without an order effect, we should see larger weights on the private signal in both belief-elicitation tasks.

Testing this hypothesis was the motivation for asking participants to form expectations in the second stage of the control sessions. Here, we elicited beliefs directly after Stage "*r*=0" and before all those stages, in which common signals are theoretically more important than private ones. Table C2 summarizes the results from Stage 2 of control sessions.

The result is striking: the bias in estimating the state is almost absent now (p=0.12), and the average weight on the private signal in estimating the other subject's estimation of the state is higher than 0.25 in all groups (p<1%). Average *individual* weights were 0.52 in estimating θ and 0.437 in second-order beliefs. Individual weights are distributed around 0.5 (p=0.37), while the difference in the weight for higher-order beliefs from 0.25 is significant at the 0.1% level.

Session, group	*Weight in estimating θ*	*Weight in estimating the partner's estimation of θ*
19, Group 1	0.524	0.475
Group 2	0.417	0.389
20, Group 1	0.463	0.358
Group 2	0.538	0.550
21, Group 1	0.543	0.459
Group 2	0.447	0.317
22, Group 1	0.459	0.388
Group 2	0.555	0.499
23, Group1	0.522	0.397
Group 2	0.603	0.399
24, Group1	0.511	0.317
Group 2	0.566	0.584
Average	0.512	0.428
St.dev	0.055	0.086

Table C2 – Group specific weights on the private signal in stated expectations of Sessions 19-24

Having ruled out order effects and best responses to others' deviations from rational first-order beliefs as possible explanations leaves us with the impression that subjects underestimate how informative the public signal is for predicting others' expectations. This can be viewed as a systematic error in forming second-order beliefs and provides an alternative explanation for results from the other stages. To our knowledge, we are the first to elicit higher-order beliefs directly and relate them to controlled information. Non-Bayesian higher-order beliefs may also be responsible for observed deviations from equilibrium in other experiments. Möbius et al. (2011) found gender differences in Bayesian updating of self assessments; we could not find a significant gender difference in stated beliefs.

Next, we turn to testing the aforementioned model of non-Bayesian higher-order beliefs. We argue that systematic errors in higher-order beliefs are too low to explain the observed deviations from equilibrium in games with an interior *r*.

For each standard session, we use the observed average weight for second-order beliefs to calculate the weight that this group should attach on the private signal in Treatments "*r*=0.5", "*r*=0.25", or "*r*=0.75" according to the model presented above. As we elicited beliefs directly after these treatments, we may expect that higher-order beliefs in their later periods are formed in about the same way as elicited beliefs. Table C3 compares the results of this calculation with the estimated weights on x_i in the second half of Treatments "*r*=0.5", "*r*=0.25", or "*r*=0.75". With one exception (Session 17), the estimated weights on the private signal are higher than those that would be explained by our model of Non-Bayesian higher-order beliefs. This result stands if we use individual data instead of group averages, where the differences between estimated weights in Treatments "*r*=0.5", "*r*=0.25", or "*r*=0.75" and weights calculated from our model with estimated weights for higher-order beliefs as input is significant at the 0.1% level.

Session	Calculated weight on private signal using estimated error in higher-order beliefs	Estimated weight on private signal (*r*=0.5/ *r*=0.25/ *r*=0.75, 2nd half)
1	0.325	0.408
2	0.371	0.463
3	0.386	0.447
4	0.303	0.475
5	0.318	0.393
6	0.300	0.453
7	0.448	0.492
8	0.440	0.466
9	0.459	0.500
10	0.434	0.460
11	0.458	0.475
12	0.421	0.473
13	0.208	0.391
14	0.205	0.294
15	0.278	0.457
16	0.241	0.377
17	0.211	0.187
18	0.217	0.338

Table C3 – Comparing weights from a model of Non-Bayesian higher-order beliefs with observations

Thus, we conclude that errors in forming higher-order beliefs are in general not sufficiently strong to explain observed behavior in the game. There must be another form of irrationality in addition to non-Bayesian beliefs.[7] Hence, we cannot reject the hypothesis of limited levels of reasoning. Combining levels of reasoning with non-Bayesian beliefs yields higher estimated levels of reasoning.[8] However, for interior values of *r*, aggregate behavior is in most cases consistent with levels of reasoning not exceeding degree 2, even if we account for the observed systematic mistakes in forming higher-order beliefs. This is laid out in Appendix D.

[7] A third possible explanation that cannot be dealt with in this paper was pointed out by a discussant: subjects might pay less attention to the actions of others, because they have less information about them. Strategic uncertainty turns beliefs about others' behavior into an ambiguous guess as opposed to estimating the fundamental state for which probabilistic information is available.

[8] We are grateful to Gabriel Desgranges for asking us how non-Bayesian beliefs affect the result on levels of reasoning.

D. Combining non-Bayesian higher-order beliefs with limited levels of reasoning

What is the level of reasoning, if we account for systematic mistakes in forming higher-order beliefs? Combining the model of Appendix C with limited levels of reasoning yields the following weights on private signals denoted by $\hat{\gamma}_k$:

Level 1: $a_i^1 = e_i(\theta) = 0.5(x_i + y)$ => $\hat{\gamma}_1 = 0.5$.

Level 2: $a_i^2 = (1-r)e_i(\theta) + re_i(a_j^1) = (1-r)E_i(\theta) + re_i(e_j(\theta))$

$$= (1-r)E_i(\theta) + r[\lambda_i x_i + (1-\lambda_i)y]$$

$$= [0.5(1-r) + r\lambda_i]x_i + [0.5(1+r) - r\lambda_i]y \quad => \quad \hat{\gamma}_2 = 0.5(1-r) + r\lambda.$$

Level 3: $a_i^3 = (1-r)e_i(\theta) + re_i(a_j^2) = (1-r)E_i(\theta) + re_i\left[(1-r)e_j(\theta) + re_j(a_i^1)\right]$

$$= (1-r)E_i(\theta) + (1-r)re_i(e_j(\theta)) + r^2 e_i(e_j(e_i(\theta)))$$

$$= (1-r)E_i(\theta) + (1-r)r[\lambda_i x_i + (1-\lambda_i)y] + r^2\left[\lambda_i^{3/2} x_i + (1-\lambda_i^{3/2})y\right]$$

$$= \left[(1-r)[0.5 + r\lambda_i] + r^2\lambda_i^{3/2}\right]x_i + \left(1 - \left[(1-r)[0.5 + r\lambda_i] + r^2\lambda_i^{3/2}\right]\right)y$$

$$=> \hat{\gamma}_3 = (1-r)[0.5 + r\lambda_i] + r^2\lambda_i^{3/2}.$$

Level 4: $a_i^4 = (1-r)e_i(\theta) + re_i(a_j^3) = (1-r)E_i(\theta) + re_i\left[(1-r)e_j(\theta) + re_j(a_i^2)\right]$

$$= (1-r)E_i(\theta) + (1-r)re_i(e_j(\theta)) + r^2 e_i(e_j(\left[1-r)e_i(\theta) + re_i(a_j^1)\right]))$$

$$= (1-r)E_i(\theta) + (1-r)\left[r[\lambda_i x_i + (1-\lambda_i)y] + r^2\left[\lambda_i^{3/2} x_i + (1-\lambda_i^{3/2})y\right]\right] + r^3\left[\lambda_i^2 x_i + (1-\lambda_i^2)\right]$$

$$= \left[(1-r)\left[0.5 + r\lambda_i + r^2\lambda_i^{3/2}\right] + r^3\lambda_i^2\right]x_i + \left(1 - \left[(1-r)\left[0.5 + r\lambda_i + r^2\lambda_i^{3/2}\right] + r^3\lambda_i^2\right]\right)y$$

$$=> \hat{\gamma}_4 = (1-r)\left[0.5 + r\lambda_i + r^2\lambda_i^{3/2}\right] + r^3\lambda_i^2.$$

Level $k \geq 3$: $\hat{\lambda}_k = (1-r)\left[0.5 + r\lambda\sum_{j=0}^{k-3}(r\sqrt{\lambda})^j\right] + r^{k-1}\lambda_i^{k/2}$.

Table D1 compares the estimated weights in the second half of treatments with $0 < r < 1$ with those that subjects should put on the private signal according to the model combining limited levels of reasoning with non-Bayesian higher-order beliefs. In Treatments "*r*=0.5" and "*r*=0.25" (Sessions 1-12), the estimated weights are all higher than those arising from non-Bayesian beliefs and level-2 reasoning. When "*r*=0.75" (Sessions 13-18), however, in 3 out of 6 cases estimated weights are between those of level 2 and level 3, in 2 cases estimated weights are between those of level 1 and 2, and for Session 17, the estimated weight on the private signal is even smaller than the equilibrium weight.

Session, group	Estimated weight on private signal (r=0.5/ r=0.25/ r=0.75, 2nd half)	Calculated weights in the model combining estimated errors in higher-order beliefs with limited levels of reasoning				
		Level 1	Level 2	Level 3	Level 4	Equilibrium
1	0.408	0.5	0.364	0.334	0.327	0.325
2	0.463	0.5	0.421	0.386	0.375	0.371
3	0.447	0.5	0.438	0.402	0.391	0.386
4	0.475	0.5	0.335	0.310	0.305	0.303
5	0.393	0.5	0.355	0.326	0.320	0.318
6	0.453	0.5	0.330	0.306	0.301	0.300
7	0.492	0.5	0.459	0.450	0.449	0.448
8	0.466	0.5	0.450	0.442	0.441	0.440
9	0.500	0.5	0.470	0.461	0.459	0.459
10	0.460	0.5	0.443	0.435	0.434	0.434
11	0.475	0.5	0.469	0.460	0.458	0.458
12	0.473	0.5	0.429	0.422	0.421	0.421
13	0.391	0.5	0.328	0.255	0.227	0.208
14	0.294	0.5	<u>0.322</u>	0.250	0.222	0.205
15	0.457	0.5	0.440	0.357	0.316	0.278
16	0.377	0.5	<u>0.384</u>	0.304	0.268	0.241
17	0.187	0.5	<u>0.333</u>	<u>0.259</u>	<u>0.230</u>	<u>0.211</u>
18	0.338	0.5	<u>0.344</u>	0.269	0.238	0.217

Table D1 – Comparing weights from a model of Non-Bayesian higher-order beliefs and limited levels of reasoning with observations. Underlined numbers indicate cases where data are consistent with an application of higher levels of reasoning.

E. Order effects

The data indicate order effects. During the course of the experiment, subjects seem to learn that public signals are more important than private ones for estimating the likely action of other participants. In particular:

1. For "r=0.5", the estimated weights for the private signal are higher in Sessions 19-24 than in Sessions 1-6. The difference is significant (p=1%) for the fist half of the stage where "r=0.5", but insignificant (p=7.7%) for the second half, using two-tailed Mann-Whitney tests. The difference could result from the different stages that subjects went through before they reached "r=0.5".

2. When "r=1" is conducted in an early stage (as in Sessions 1-18), the weight on public information significantly increases in the second half of the treatment (Result 4), which is not the case in Sessions 19-24 with "r=1" in the last stage. Furthermore, when "r=1" is conducted in as last stage, subjects assign a larger weight to public signals from the start compared to Stage "r=1" in the standard sessions (first half: p<1%, second half: p=0.12, two-tailed Mann-Whitney). In consequence of the larger weight on public signals, average payoffs are higher for "r=1 with 2 signals" in control sessions than in standard sessions (p<1%). These results indicate that learning has not settled within the 10 periods of this treatment. This is in line with

Fehr et al. (2011), who also find that the convergence process to an equilibrium in pure coordination games with extrinsic private and public signals takes surprisingly long.

3. Belief elicitation: When asked for their expectation of the fundamental state θ directly after "*r*=0" (Sessions 19-24), subjects assign equal weights to both signals (as in Stage "*r*=0"). When asked after all other stages in Sessions 1-18, subjects assign a larger weight to the public signal. The difference between stating expectations at the end of the experiment in standard sessions and at Stage 2 of the experiment in control sessions is significant with p=1% for group data and 0.5% for individual data (two-tailed Mann-Whitney). In Sessions 1-18, the difference between the weights in stated expectations in the last 5 periods and that put in Treatment "*r*=0" is significant (p=0.1%, two-tailed Wilcoxon matched pairs) for group data as well as for individual data.
4. In the formation of higher-order beliefs, the weight on the public signal is significantly larger when beliefs are elicited after all other stages (Sessions 1-18) as compared to when they are elicited directly after "*r*=0" (Sessions 19-24). The difference is significant for both, group and individual data (p<1%, two-tailed Mann-Whitney).
5. Treatments "*r*=1 with 1 common signal" and "*r*=1 with 1 private signal" do not seem to have different effects on subsequent behavior. Two-tailed Mann-Whitney tests cannot reject the hypotheses that weights in subsequent stages of Sessions 19-21 come from the same distribution as in Sessions 22-24 (p-values are above 0.5).

The first four points mentioned here indicate that individual weights are adjusted gradually when a new stage starts. Subjects seem to start a stage with weights closer to the final weights of the previous stage, which is not surprising. Our tests indicate, however, that 30 periods are sufficient to undo these order effects. Since our analysis focuses on treatments with $0 < r < 1$ that all had 30 periods, we have no reasons to believe that tests on data from the second half of these treatments are affected by order effects.

References for the appendix

Camerer, C. F., G. Loewenstein, and M. Weber (1989), "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 97, 1232-54.

Fehr, D., F. Heinemann, and A. Llorente-Saguer (2011), "The Power of Sunspots", *SFB 649 Discussion Paper* 2011-070.

Möbius, M. M., M. Niederle, P. Niehaus, and T. S. Rosenblat (2011), "Managing Self-Confidence: Theory and Experimental Evidence," *NBER Working Paper* 17014.