Supplementary material for

Excessive Herding in the Laboratory: The Role of Intuitive Judgments

Christoph March and Anthony Ziegelmeyer

[For Online Publication]

Appendix A contains the instructions for Experiment 1 (translated from German) and the description of the Bayes-rational strategy adopted by *observed* in Experiment 3. Instructions for Experiments 2-4 were adapted accordingly and they are available from the authors upon request. Appendix B provides complementary results for Experiment 1, and the detailed results for Experiments 2-4 are to be found in Appendix C. We describe in Appendix D the procedure to assign *unobserved* into decision rules. Appendix E characterizes the representativeness of public guesses for signals, and it contains proofs related to the latter as well as graphical illustrations of our model of intuitive observational learning. Appendix F details our estimation and prediction procedures, and it complements the estimation and prediction results reported in the main text. Appendices G and H investigate the robustness of our prediction results with respect to the modelling specifications of intuitive observational learning. Finally, Appendix I evaluates the increase in the predictive power of intuitive observational learning due to the inclusion of efficiency concerns.

Appendix A. Instructions for Experiment 1

A.1. General Instructions

Welcome to the experiment!

Please do not touch the mouse and do not open the envelope until you are instructed to do so.

This is an experiment in decision-making and all your decisions will be treated in an anonymous way. From now on, we ask you to remain seated quietly at your computer desk. Please do not talk, exclaim, or try to communicate with other participants during the experiment. Participants who intentionally violate this rule will be asked to leave the experiment without being financially compensated. If you have any questions during the experiment, please raise your hand and wait for an experimenter to come to you.

Your earnings will depend partly on your decisions and partly on chance. In addition to the earnings from your decisions, you will receive 3 Euros. This payment is to compensate you for showing up on time. At the end of the experiment the total amount of money that you have earned will be paid to you privately in cash.

Setting of the experiment

In the experiment, there are two roles: *observed* and *unobserved.*

7 participants have been assigned randomly to the role of observed. All 8 remaining participants have been assigned to the role of unobserved. Each participant remains in the same role for the entire duration of the experiment.

The experiment consists of 4 parts. Instructions for the first part of the experiment will be distributed in a few moments. We ask you to read the instructions for the first part of the experiment carefully, and once each participant has done so an experimenter will read them aloud. After the instructions for the first part of the experiment have been read aloud, you will be informed about the role you have been assigned to, *observed* or *unobserved.* Instructions for the second, third, and fourth part of the experiment will be made available before each of the respective parts begins.

A.2. Instructions for Part 1

Part 1 of the experiment consists of 3 independent rounds and each round is conducted in the same way.

I. How a round progresses

I.1. The *assistant* picks either BLUE or ORANGE at random.

Each round begins with the *assistant* picking either the color **BLUE** or the color **ORANGE** at random. You and all other participants have just been instructed about the picking procedure which is as follows:

1. An experimenter shuffles a deck of 20 cards and lays them down on a table with the back of the cards facing the *assistant*. **11** cards have a **blue front** and **9** cards have an **orange front**.
2. The *assistant* picks 1 card out of the 20 cards.
   - If the picked card has a **blue front** then the color picked at random is **BLUE**.
   - If the picked card has an **orange front** then the color picked at random is **ORANGE**.

In each round your task, which is also the task of each of the other participants, is to guess which color has been picked at random by the *assistant*.

I.2. The *assistant* selects the "OBSERVED" and "UNOBSERVED" urns

Once a color has been picked at random, the *assistant* selects an urn labeled "OBSERVED" and an urn labeled "UNOBSERVED" from a collection of urns containing **blue** and **orange** balls.

The composition of the urn labeled "OBSERVED" depends only on the color which has been picked at random by the *assistant*. The composition of the urn labeled "OBSERVED" is as follows

In case the color **BLUE** has been picked, the "OBSERVED" urn contains **14 blue** and **7 orange** balls.	In case the color **ORANGE** has been picked, the "OBSERVED" urn contains **7 blue** and **14 orange** balls.

The composition of the urn labeled "UNOBSERVED" also depends only on the color picked at random by the *assistant*. The composition of the urn labeled "UNOBSERVED" is as follows

In case the color **BLUE** has been picked, the "UNOBSERVED" urn contains **14 blue** balls and **7 orange** balls.	In case the color **ORANGE** has been picked, the "UNOBSERVED" urn contains **7 blue** balls and **14 orange** balls.

I.3. Each participant learns the color of 1 ball

Once the "OBSERVED" and "UNOBSERVED" urns have been selected by the *assistant*, each *observed* is informed about the color of a ball drawn from the "OBSERVED" urn whereas each *unobserved* is informed about the color of a ball drawn from the "UNOBSERVED" urn. Concretely,

- one of the experimenters approaches each *observed*, one at a time, to draw a ball from the "OBSERVED" urn. Each *observed* draws a ball without being able to see the composition of the "OBSERVED" urn. After each draw, the ball is returned to the urn before making the next draw. Apart from the participant who draws the ball, no other participant sees its color. Thus, each *observed* is informed about the color of 1 and only 1 ball drawn from the "OBSERVED" urn.

- another experimenter approaches each *unobserved*, one at a time, to draw a ball from the "UNOBSERVED" urn. Each *unobserved* draws a ball without being able to see the composition of the "UNOBSERVED" urn. After each draw, the ball is returned to the urn before making the next draw. Apart from the participant who draws the ball, no other participant sees its color. Thus, each *unobserved* is informed about the color of 1 and only 1 ball drawn from the "UNOBSERVED" urn.

I.4. Each participant makes a guess

Each round consists of 8 guessing periods with one *observed* making a guess in each of the first seven periods and one *unobserved* making a guess in each of the eight periods. Thus, each participant makes one and exactly one guess in each round.

In each round the order in which *observed* make their guesses is randomly determined. If you have been assigned to the role of an *observed* then, in a given round, you might be the first *observed* to make a guess, or you might guess in any period from period 2 to period 6, or you might be the last *observed* to make a guess.

Similarly, in each round the order in which *unobserved* make their guesses is randomly determined. If you have been assigned to the role of an *unobserved* then, in a given round, you might be the first *unobserved* to make a guess, or you might guess in any period from period 2 to period 7, or you might be the last *unobserved* to make a guess.

First guessing period. In period 1, 1 *observed* and 1 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Once both guesses have been made, period 2 starts. The *observed* and the *unobserved* who made a guess in period 1 do not make any further guess in the current round.

Guessing period 2 to 7. In period 2 to 7, the guess made by the *observed* in the previous period is made public meaning that all other *observed* as well as all *unobserved* are informed of that guess. After that, 1 *observed* and 1 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Both participants do not make any further guess in the current round. Once both guesses have been made, the next period starts.

Last guessing period. In period 8, the guess made by the *observed* in period 7 is made public and only the *unobserved* who did not make a guess yet is asked to guess which color has been picked at random by the *assistant*.

Please note that the guess made by each of the *unobserved* is kept private meaning that no other *unobserved* and no *observed* is informed of the guess made by any of the *unobserved*.

Once each participant has made a guess, you and each of the other participants are informed of the color that was actually picked at random by the *assistant* at the beginning of the round. Once all participants have been informed, the round is over.

II. Earnings

In each of the 3 independent rounds, participants get paid for the guess they make. If the participant's guess matches the color picked at random by the *assistant*, the participant earns 1 Euro. If the participant's guess does not match the color picked at random by the *assistant*, the participant earns nothing.

Once the 3 independent rounds have been completed, participants are informed of the total amount of euros they earned in the first part of the experiment.

A.3. Instructions for Part 2

The second part of the experiment shares many similarities with the first part of the experiment. Still, the two parts of the experiment differ in some respects.

Hereafter, we explain thoroughly the aspects of the second part of the experiment which were not present in the first part of the experiment. On the other hand, the aspects of the second part of the experiment which were already present in the first part of the experiment are merely mentioned without much detail.

Part 2 of the experiment consists of 6 independent rounds and each round is conducted in the same way.

I. How a round progresses

I.1. The *assistant* picks either BLUE or ORANGE at random

Each round begins with the *assistant* picking either the color **BLUE** or the color **ORANGE** at random. The picking procedure used in part 2 of the experiment is identical to the picking procedure used in part 1 of the experiment.

In each round your task, which is also the task of each of the other participants, is to guess which color has been picked at random by the *assistant*.

I.2. Each participant learns the color of 1 ball

Once a color has been picked at random by the *assistant*, each *observed* is informed about the color of a ball drawn from the “OBSERVED” *virtual* urn whereas each *unobserved* is informed about the color of a ball drawn from the “UNOBSERVED” *virtual* urn.

Detailed explanations about the drawing of balls from *virtual* urns will be displayed on the screen of your computer after all participants have finished reading these two pages.[1]

I.3. Each participant makes guesses

In each round, after having learned the color of 1 ball, each participant has to guess which color has been picked at random by the assistant. Each round consists of 8 guessing periods.

- In each round, an *observed* makes between 1 and 7 guesses.
- In each round, each *unobserved* makes 8 guesses.

First guessing period. In period 1, all 7 *observed* and all 8 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Once all 15 guesses have been made, period 2 starts.

Guessing period 2. At the beginning of period 2, the guess made by 1 of the 7 *observed* in period 1 is selected at random and this guess is shown to all 15 participants. The *observed* whose guess is randomly selected does not make any further guess in the current round. Therefore, only 6 *observed* remain who can guess in period 2. Afterwards, all 6 remaining *observed* and all 8 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Once all 14 guesses have been made, period 3 starts.

Guessing periods 3, 4, 5, and 6. At the beginning of the period, the guess made by 1 of the *observed* in the previous period is selected at random and this guess is shown to all 15 participants. The *observed* whose guess is randomly selected does not make any further guess in the current round. Afterwards, all remaining *observed* and all 8 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Once all guesses have been made, the next period starts.

[1] The short on-screen demonstration of the draws from the virtual urns is available from the authors upon request.

Guessing period 7. At the beginning of period 7, the guess made by 1 of the 2 *observed* in period 6 is selected at random and this guess is shown to all 15 participants. The *observed* whose guess is randomly selected does not make any further guess in the current round. Therefore, only 1 *observed* remains who can guess in period 7. Afterwards, the *observed* and all 8 *unobserved* are asked to guess which color has been picked at random by the *assistant*. Once all 9 guesses have been made, period 8 starts.

Last guessing period. At the beginning of period 8, the guess made by the *observed* in period 7 is shown to all 15 participants. The *observed* who guessed in period 7 does not make a guess in period 8. Therefore, only the 8 *unobserved* are asked to guess which color has been picked at random by the *assistant*.

Please note that the guesses made by each of the *unobserved* are kept private meaning that no other *unobserved* and no *observed* is informed of the guesses made by any of the *unobserved*.

Once all participants have made all their guesses, you and each of the other participants are informed of the color that was actually picked at random by the *assistant* at the beginning of the round. Once all participants have been informed, the round is over.

II. Earnings

In each of the 6 independent rounds, each participant gets paid for 1 and only 1 of the guesses made. If the participant's guess matches the color picked at random by the *assistant*, the participant earns 1 Euro. If the participant's guess does not match the color picked at random by the *assistant*, the participant earns nothing.

1. **For each *observed*, only the last guess is paid.**
   Each *observed* gets paid only for the last guess he/she made in the round. Said differently, the guess of an observed is paid only in case the guess is made public meaning that it is observed by all 15 participants. Obviously, at the time a guess is made, an *observed* does not know whether the guess is going to be made public or not. So, for each guess that an *observed* makes, there is a chance that this guess is the one which is going to be paid.

2. **For each *unobserved*, only the guess of the assigned period is paid.**
   In each of the 6 independent rounds, each *unobserved* makes a guess in each period for a total of 8 guesses. Once each of the *unobserved* has made 8 guesses, the round is over. As soon as the round is over, each of the 8 *unobserved* is assigned a period number from 1 to 8. Concretely, one of the *unobserved* is assigned to period 1, another *unobserved* is assigned to period 2, ..., and another *unobserved* is assigned to period 8. The assignment is completely random meaning that the guesses made by the *unobserved* do not influence the period numbers assigned to them. An *unobserved* gets paid only for the guess made in the assigned period. Obviously, before having made all 8 guesses, an *unobserved* does not know which period number is assigned to her/him. So, each guess that an *unobserved* makes has an equal chance of being paid.

Once the 6 independent rounds have been completed, participants are informed of the total amount of euros they earned in the second part of the experiment.

A.4. Instructions for Part 3

Part 3 of the experiment consists of 6 independent rounds. Each round proceeds the same way as in part 2 except that

In case the color **BLUE** has been picked, the "UNOBSERVED" urn contains **18 blue** balls and **3 orange** balls.	In case the color **ORANGE** has been picked, the "UNOBSERVED" urn contains **3 blue** balls and **18 orange** balls.

Once the 6 independent rounds have been completed, participants are informed of the total amount of euros they earned in the third part of the experiment.

A.5. Demographic Questionnaire

1. What is your field of study?
2. When were you born? (Month/Year)
3. Your gender: ☐ Female ☐ Male

To know our subject pool better, it would be helpful to learn about your cultural background. We thus ask you to also answer the following questions.

4. What is your first language?
   (By first language we mean the language you have mainly spoken during your childhood or at your family home.)
5. What is your nationality?

A.6. Instructions for Part 4

Part 4 of the experiment consists of 6 independent rounds. Each round proceeds the same way as in part 3 except that

In case the color **BLUE** has been picked, the "UNOBSERVED" urn contains **12 blue** balls and **9 orange** balls.	In case the color **ORANGE** has been picked, the "UNOBSERVED" urn contains **9 blue** balls and **12 orange** balls.

Once the 6 independent rounds have been completed, participants are informed of the total amount of euros they earned in the course of the experiment.

A.7. Guessing Rule Adopted by *Observed* in Experiment 3

The role of *observed* is taken by computer algorithms. Each *observed* adopts the same rule to guess which color the *assistant* randomly picked at the beginning of the round. As described below, the rule prescribes that in each period the guess depends on the ball of the *observed* and the sequence of previous guesses.

Period	Ball of the *observed*	Sequence of previous guesses				Current guess of the *observed*
1	🔵	There is no previous guess in period 1				BLUE
	🟠	There is no previous guess in period 1				ORANGE
2	🔵	BLUE				BLUE
	🔵	ORANGE				BLUE
	🟠	BLUE				BLUE
	🟠	ORANGE				ORANGE
3	🔵	BLUE	BLUE			BLUE
	🔵	ORANGE	BLUE			BLUE
	🔵	ORANGE	ORANGE			ORANGE
	🟠	BLUE	BLUE			BLUE
	🟠	ORANGE	BLUE			ORANGE
	🟠	ORANGE	ORANGE			ORANGE
4	🔵	BLUE	BLUE	BLUE		BLUE
	🔵	ORANGE	BLUE	BLUE		BLUE
	🔵	ORANGE	BLUE	ORANGE		BLUE
	🔵	ORANGE	ORANGE	ORANGE		ORANGE
	🟠	BLUE	BLUE	BLUE		BLUE
	🟠	ORANGE	BLUE	BLUE		BLUE
	🟠	ORANGE	BLUE	ORANGE		ORANGE
	🟠	ORANGE	ORANGE	ORANGE		ORANGE
5	🔵	BLUE	BLUE	BLUE	BLUE	BLUE
	🔵	ORANGE	BLUE	BLUE	BLUE	BLUE
	🔵	ORANGE	BLUE	ORANGE	BLUE	BLUE
	🔵	ORANGE	BLUE	ORANGE	ORANGE	ORANGE
	🔵	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE
	🟠	BLUE	BLUE	BLUE	BLUE	BLUE
	🟠	ORANGE	BLUE	BLUE	BLUE	BLUE
	🟠	ORANGE	BLUE	ORANGE	BLUE	ORANGE
	🟠	ORANGE	BLUE	ORANGE	ORANGE	ORANGE
	🟠	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE

Period	Ball of the *observed*	Sequence of previous guesses						Current guess of the *observed*
6	●	BLUE	BLUE	BLUE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	BLUE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE		BLUE
	●	ORANGE	BLUE	ORANGE	ORANGE	ORANGE		ORANGE
	●	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE		ORANGE
	●	BLUE	BLUE	BLUE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	BLUE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	BLUE		BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE		ORANGE
	●	ORANGE	BLUE	ORANGE	ORANGE	ORANGE		ORANGE
	●	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE		ORANGE
7	●	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE	BLUE	BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE	ORANGE	ORANGE
	●	ORANGE	BLUE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE
	●	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE
	●	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	BLUE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	BLUE	BLUE	BLUE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE	BLUE	ORANGE
	●	ORANGE	BLUE	ORANGE	BLUE	ORANGE	ORANGE	ORANGE
	●	ORANGE	BLUE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE
	●	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE	ORANGE

Appendix B. Complementary Results for Experiment 1

In this appendix, we first report the fraction of (ex-post) correct guesses. Then, we explain how we implemented the split-sample instrumental variable (IV) method described in Weizsäcker (2010) and Ziegelmeyer, March, and Kruegel (2013), and we present figures on the responses to *value_contra_PI* that complement the figure shown in the main text. We also report the regression results underlying all these figures. Finally, we discuss the dynamics of *observed* guesses over the three non-practice parts.

B.1. Fractions of Correct Guesses

Table B1 reports the fraction of (ex-post) correct guesses by role and by signal quality for the *unobserved*, and for the same majorities of public guesses as in Table 2 of the main text. In each case, we distinguish between guesses which follow private information (FPI) and guesses which contradict private information (CPI), and we average the fraction of correct guesses across signals.

		Observed		***Unobserved***					
History of		Medium quality		Low quality		Medium quality		High quality	
public guesses		FPI	CPI	FPI	CPI	FPI	CPI	FPI	CPI
Favoring majority		.846	.097	.722	.231	.850	.089	.935	.048
		(1,603)	(031)	(1,484)	(026)	(1,531)	(045)	(1,738)	(021)
No majority		.688	.320	.559	.439	.641	.353	.842	.167
		(1,502)	(075)	(655)	(057)	(629)	(051)	(716)	(012)
Contrary majority of size	1	.529	.510	.471	.621	.426	.589	.720	.214
		(490)	(098)	(172)	(195)	(209)	(090)	(293)	(014)
	2	.369	.546	.381	.699	.179	.599	.659	.310
		(103)	(174)	(042)	(186)	(056)	(142)	(126)	(029)
	3	.270	.534	.280	.663	.222	.595	.606	.283
		(037)	(176)	(025)	(184)	(036)	(173)	(094)	(053)
	4	.294	.562	.143	.671	.158	.582	.563	.274
		(017)	(121)	(014)	(140)	(019)	(158)	(064)	(062)
	$\geqslant 5$	.444	.580	.059	.672	.276	.604	.600	.234
		(009)	(100)	(017)	(259)	(029)	(288)	(110)	(124)

Note: In each cell, the 1^{st} (resp. 2^{nd}) row reports the fraction of correct guesses (resp. number of guesses).

Table B1: Fractions of Correct Guesses

Whatever the size of the contrary majority, the more profitable guess consists in following private information for *unobserved* with high quality signals. In fact, FPI guesses are at least twice more likely to be correct than CPI guesses which indicates that herding at large contrary majorities with high quality signals is severely harmful. When receiving low or medium quality signals, it is more profitable for *unobserved* to contradict private information when the contrary majority is of size $\geqslant 1$. By contrast, herding is more profitable for *observed* only when the contrary majority is of size $\geqslant 2$. There is therefore a discrepancy in the relative profitability of CPI guesses between *unobserved* with medium quality signals and *observed* at contrary majorities of size 1. This discrepancy reflects the fact that fractions of correct guesses inaccurately capture how successfully subjects learn from others.

B.2. The Split-sample IV Method

The fact that *value_contra_PI* imperfectly measures the true expected value of contradicting private information could invalidate the inferences on observational learning behavior. We address this inference problem in two (non-exclusive) ways.

First, the statistical analysis is conducted on different subsets of data with more and more stringent minimum thresholds for *sitcount*. In the next subsection of this appendix we check that the analysis reported in the main text is robust to variations in the minimum threshold for *sitcount*.

The second approach uses an instrumental variable (IV) to correct for measurement error in statistical analyses where *value_contra_PI* is an explanatory variable (see, e.g., Cameron and Trivedi, 2005, Chapter 4 and 6). A valid instrument is obtained by randomly splitting the dataset in two subsets of approximately equal size, deriving *value_contra_PI* separately on each subset, and using one of the estimates as an instrument for the other. However, a considerable efficiency loss could occur because only half of the sample is used to derive the empirical value of contradicting private information.

The efficiency loss takes two forms. First, *value_contra_PI* can often not be derived in both subsets though it can be derived in the entire dataset. This results in a smaller number of observations that can be used in IV regressions. Second, the split-sample method increases the measurement error in monetary incentives as the control variable included in IV regressions is $value\_contra\_PI_1$, the empirical value of contradicting private information in the first subset. Note also that the instrument relevance, measured by the relation between $value\_contra\_PI_1$ and $value\_contra\_PI_2$ (the empirical value of contradicting private information in the second subset), may be low.

We perform 100 random splits of the dataset, and we select the split that minimizes the efficiency loss and reaches a high instrument relevance. To do so, we calculate for each split 12 penalty scores where a score captures the ranking of the considered split relative to the other splits. There are three penalty scores for *observed* and three penalty scores for each signal quality of *unobserved*. In each case, the first score relates to the number of observations that can be used in the IV regression, and the two other scores relate to the increase in measurement error and instrument relevance. Penalty scores for the relation between $value\_contra\_PI_1$ and $value\_contra\_PI$, and between $value\_contra\_PI_1$ and $value\_contra\_PI_2$, respectively, are based on the R^2 in the corresponding OLS regression. We select the split which minimizes the average penalty score among all splits which satisfy $R^2 > 0.9$ in each of the regressions. This procedure ensures reasonable results for each role and each signal quality.

Out of the 10,039 observations with *sitcount* $\geqslant 10$, the selected split enables us to estimate the value of contradicting private information in both subsets for 9,714 observations (across the 100 randomly generated splits the number of available observations ranges from 7,866 to 9,861), and it achieves an R^2 larger than 0.93 in each regression.

Clustering

Since subjects' behavior is likely to be influenced by individual characteristics and session dynamics, we have a nested hierarchy of potential levels of clustering (across which residuals are likely to be dependent). As is recommended in this case we rely on cluster-robust standard errors computed at the most aggregated level of clustering, i.e., at the session level (see e.g. Cameron and Miller, 2010). And to correct for the small number of clusters, we apply a finite-cluster correction to the cluster-robust estimate of the variance matrix.

B.3. Responses to *value_contra_PI*

Here we report the regression results discussed in subsection 2.2.2 of the main text along with robustness checks. We regress the proportion to contradict private information against a cubic polynomial in *value_contra_PI* fully interacted with indicator variables for the role and the different parts (or signal qualities) for the *unobserved*. Table B2 reports the regression results based on the IV specification for *sitcount* $\geqslant$ 10, the threshold considered in the main text, as well as for *sitcount* $\geqslant$ 1 and 20. Table B3 reports the regression results based on the OLS specification for *sitcount* $\geqslant$ 10, 20, and 30.

	sitcount $\geqslant 1$	*sitcount* $\geqslant 10$	*sitcount* $\geqslant 20$
Constant	-0.123	-0.124	-0.050
	(0.100)	(0.100)	(0.107)
$value\_contra\_PI_1$	1.962*	1.968*	1.167
	(1.107)	(1.109)	(1.187)
$(value\_contra\_PI_1)^2$	-8.612**	-8.626**	-6.163*
	(3.462)	(3.466)	(3.700)
$(value\_contra\_PI_1)^3$	12.433***	12.443***	10.293***
	(3.128)	(3.131)	(3.330)
Unobserved in part 2	0.304**	0.305**	0.232*
	(0.124)	(0.124)	(0.133)
Unobserved in part 2 × $value\_contra\_PI_1$	-3.564***	-3.570***	-2.769*
	(1.320)	(1.321)	(1.420)
Unobserved in part 2 × $(value\_contra\_PI_1)^2$	12.005***	12.019***	9.556**
	(4.048)	(4.049)	(4.345)
Unobserved in part 2 × $(value\_contra\_PI_1)^3$	-11.250***	-11.260***	-9.109**
	(3.609)	(3.610)	(3.841)
Unobserved in part 3	0.088	0.089	0.017
	(0.136)	(0.136)	(0.145)
Unobserved in part 3 × $value\_contra\_PI_1$	-0.642	-0.648	0.132
	(2.188)	(2.188)	(2.283)
Unobserved in part 3 × $(value\_contra\_PI_1)^2$	-2.714	-2.700	-5.124
	(11.815)	(11.816)	(11.976)
Unobserved in part 3 × $(value\_contra\_PI_1)^3$	18.482	18.472	20.665
	(20.692)	(20.691)	(20.537)
Unobserved in part 4	0.456**	0.459**	0.329
	(0.216)	(0.216)	(0.215)
Unobserved in part 4 × $value\_contra\_PI_1$	-4.708**	-4.707**	-3.363
	(2.146)	(2.147)	(2.155)
Unobserved in part 4 × $(value\_contra\_PI_1)^2$	14.076**	13.986**	10.053
	(6.089)	(6.085)	(6.177)
Unobserved in part 4 × $(value\_contra\_PI_1)^3$	-13.167***	-13.049***	-9.760*
	(5.027)	(5.023)	(5.139)
Observations	9,738	9,714	9,644
R^2	0.468	0.469	0.495

Notes: i) Robust standard errors in parentheses, clustered at the session level.
ii) * (10%); ** (5%); and *** (1%) significance level.

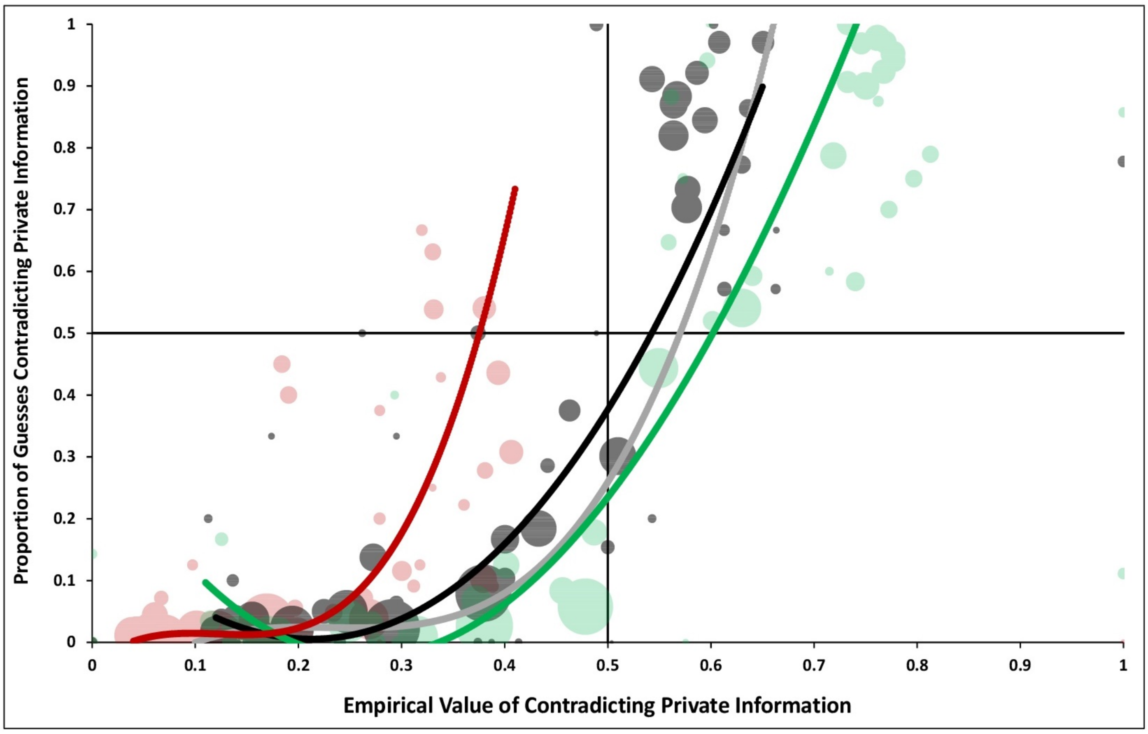
Table B2: Propensity to Contradict Private Information (IV regressions)

Figure B1 (resp. B2) plots *value_contra_PI* against the proportion of contradictions collected in identical guessing situations with *sitcount* $\geqslant$ 1 (resp. *sitcount* $\geqslant$ 20), and it superimposes fitted curves from the IV regressions. There is basically no difference between the average responses to *value_contra_PI* predicted by IV regressions for samples of guessing situations with *sitcount* $\geqslant$ 1, 10 or 20.

	$sitcount \geqslant 10$	$sitcount \geqslant 20$	$sitcount \geqslant 30$
Constant	0.124*	0.072	0.033
	(0.054)	(0.071)	(0.057)
$value\_contra\_PI$	-0.901	-0.274	0.181
	(0.599)	(0.791)	(0.599)
$(value\_contra\_PI)^2$	0.935	-1.003	-2.304
	(1.834)	(2.385)	(1.758)
$(value\_contra\_PI)^3$	2.944*	4.466*	5.261**
	(1.493)	(1.967)	(1.615)
Unobserved in part 2	0.280	0.223	0.166
	(0.153)	(0.124)	(0.109)
Unobserved in part 2 × $value\_contra\_PI$	-3.026	-2.345	-1.642
	(1.724)	(1.372)	(1.192)
Unobserved in part 2 × $(value\_contra\_PI)^2$	9.281	6.563	3.908
	(5.609)	(4.222)	(3.917)
Unobserved in part 2 × $(value\_contra\_PI)^3$	-6.950	-3.806	-0.593
	(5.215)	(3.696)	(3.903)
Unobserved in part 3	-0.088	-0.061	-0.037
	(0.058)	(0.078)	(0.059)
Unobserved in part 3 × $value\_contra\_PI$	0.085	0.109	0.351
	(1.001)	(1.108)	(0.585)
Unobserved in part 3 × $(value\_contra\_PI)^2$	4.990	2.618	-2.919
	(5.289)	(4.941)	(2.225)
Unobserved in part 3 × $(value\_contra\_PI)^3$	-7.731	-1.983	10.180*
	(8.625)	(7.592)	(5.050)
Unobserved in part 4	0.358**	0.248**	0.222**
	(0.111)	(0.103)	(0.081)
Unobserved in part 4 × $value\_contra\_PI$	-3.660**	-2.630**	-2.378**
	(1.163)	(1.098)	(0.846)
Unobserved in part 4 × $(value\_contra\_PI)^2$	10.258**	7.530**	6.719**
	(3.249)	(3.145)	(2.287)
Unobserved in part 4 × $(value\_contra\_PI)^3$	-8.821***	-6.675**	-5.657**
	(2.479)	(2.479)	(1.717)
Observations	11,217	10,391	9,155
R^2	0.517	0.527	0.546

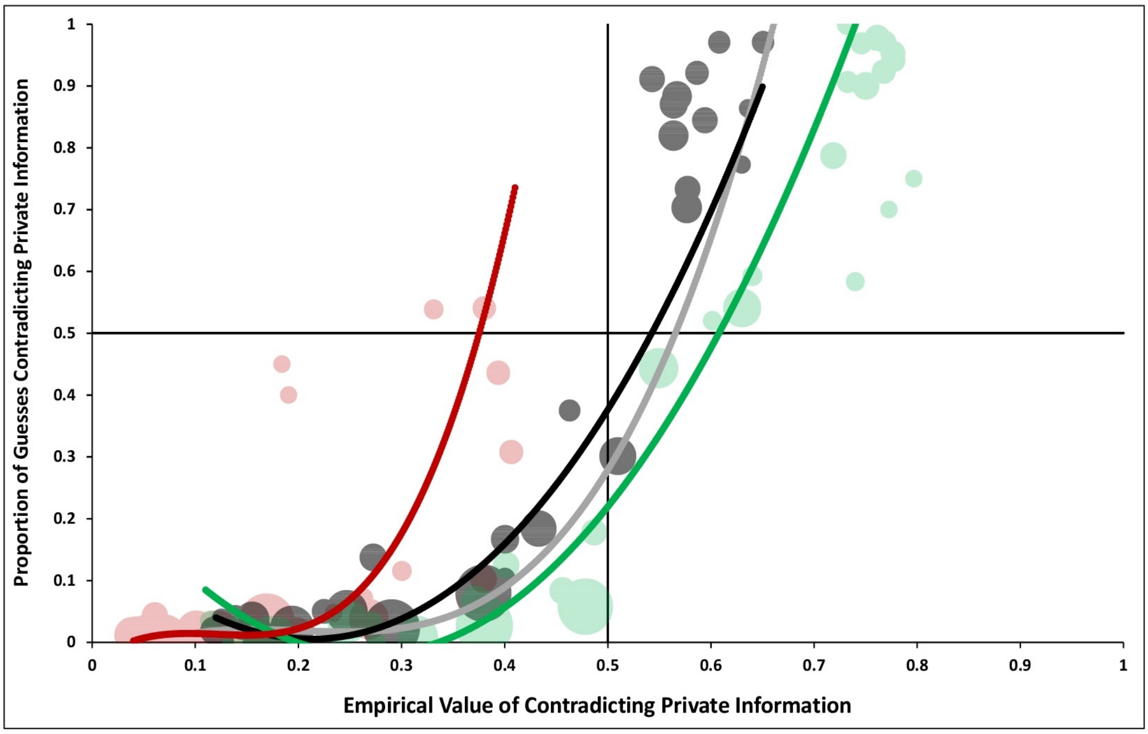
Notes: i) Robust standard errors in parentheses, clustered at the session level.
ii) * (10%); ** (5%); and *** (1%) significance level.

Table B3: Propensity to Contradict Private Information (OLS regressions)



Notes: i) ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively;
●: *Observed* guesses.
ii) The four colored curves are the fitted curves from the IV regression reported in the second column of Table B2.

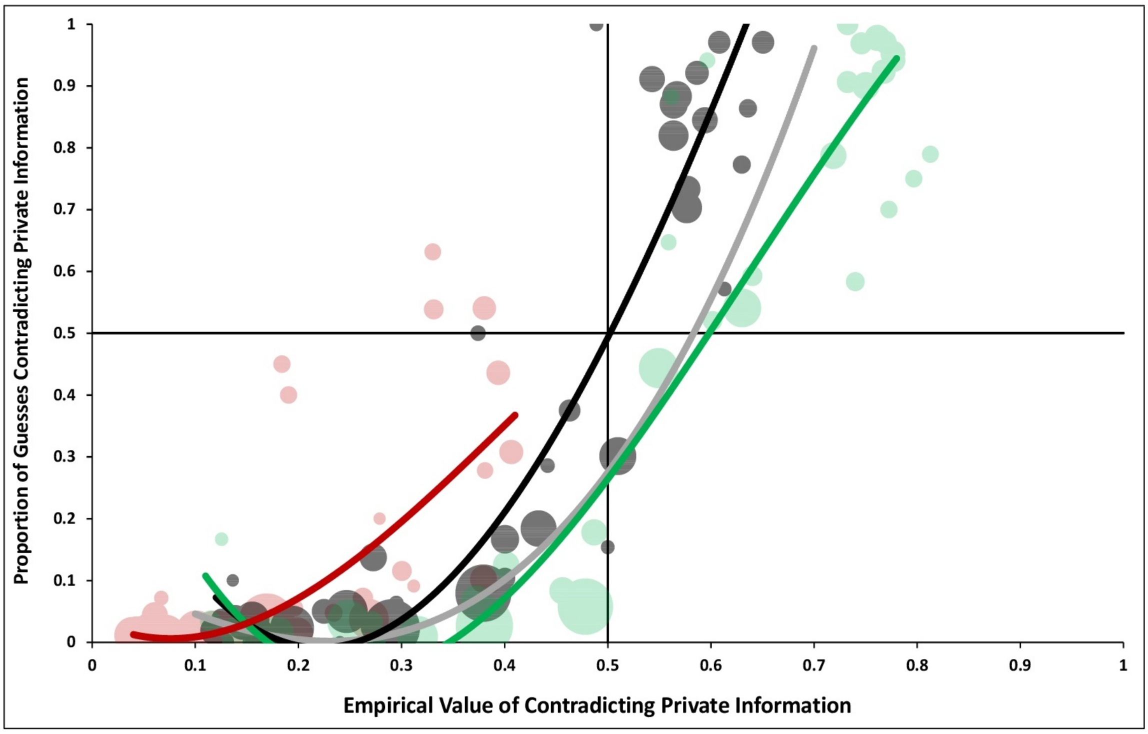
Figure B1: Responses to the Empirical Value of Contradicting Private Information (*sitcount* $\geqslant 1$)



Notes: i) ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively;
●: *Observed* guesses.
ii) The four colored curves are the fitted curves from the IV regression reported in the fourth column of Table B2.

Figure B2: Responses to the Empirical Value of Contradicting Private Information (*sitcount* $\geqslant 20$)

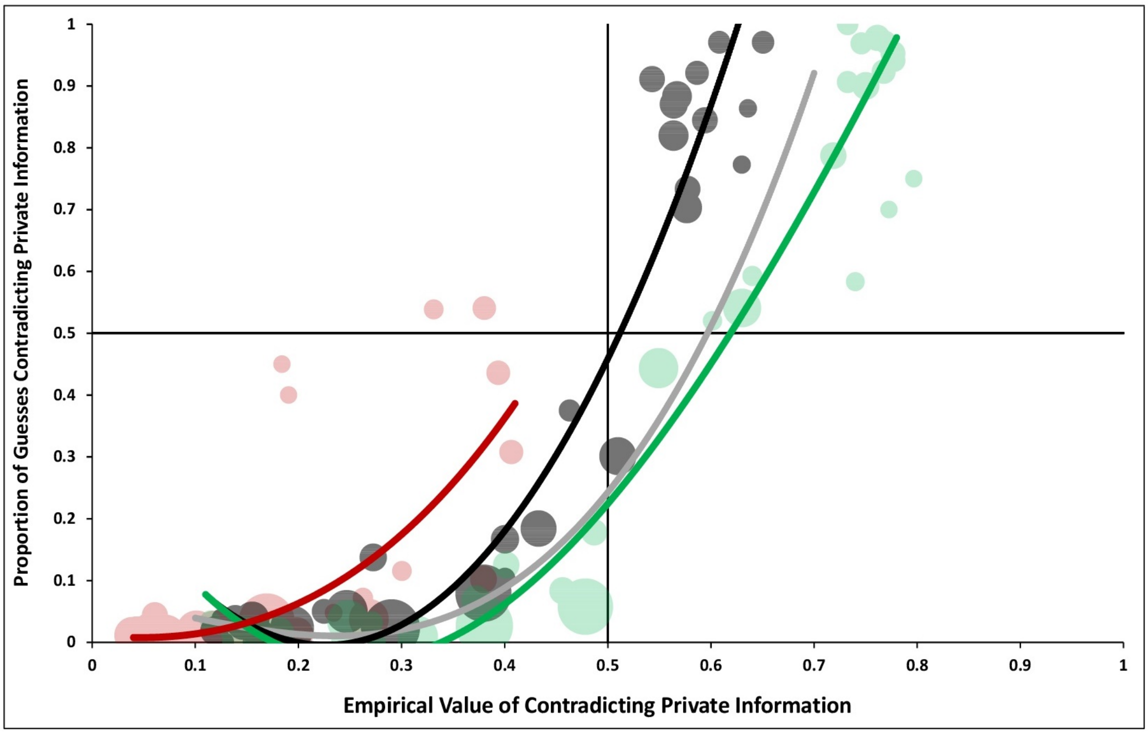
Figure B3 (resp. B4 and B5) plots *value_contra_PI* against the proportion of contradictions collected in identical guessing situations with *sitcount* $\geqslant 10$ (resp. *sitcount* $\geqslant 20$ and *sitcount* $\geqslant 30$), and it superimposes fitted curves from the OLS regressions.



Notes: i) ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.
ii) The four colored curves are the fitted curves from the OLS regression reported in the second column of Table B3.

Figure B3: Responses to the Empirical Value of Contradicting Private Information (*sitcount* $\geqslant 10$)

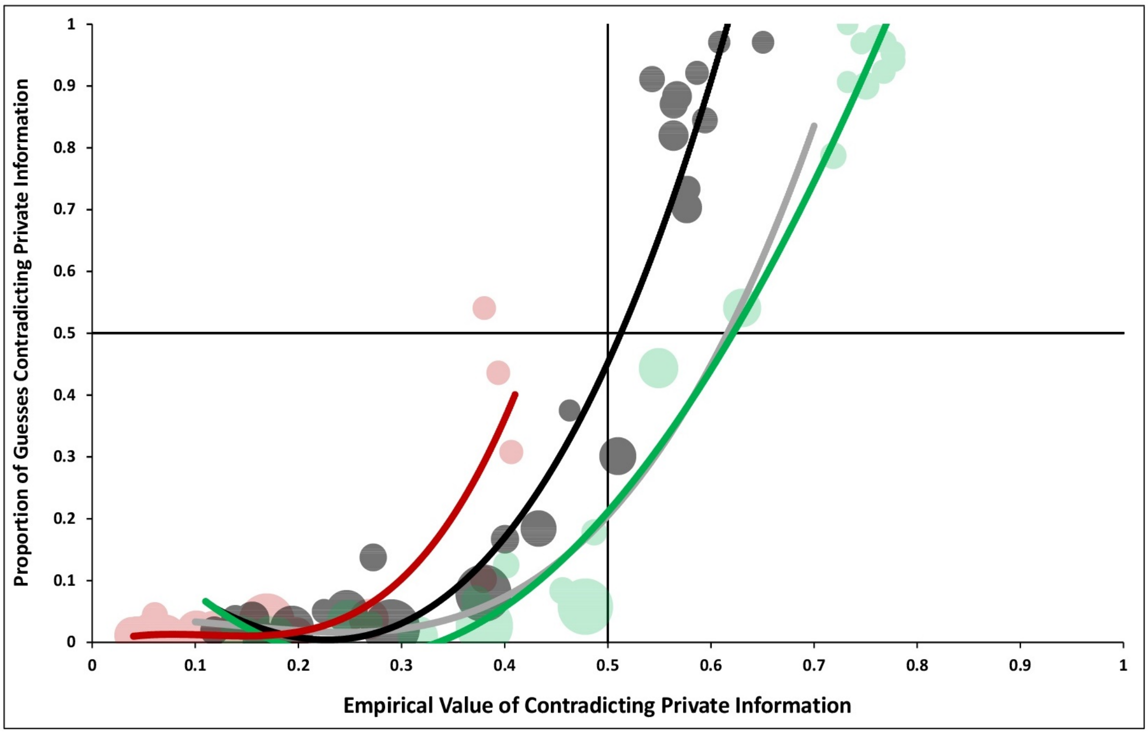
In the case of *unobserved* with low quality signals and *observed*, the average response to *value_contra_PI* predicted by the OLS regression, whether the minimum threshold for *sitount* is 10, 20 or 30, is similar to the one predicted by IV regressions. In the case of *unobserved* with high quality signals, the OLS regression predicts an excessive herding as pronounced as the one predicted by IV regressions only when the minimum threshold for *sitcount* is 30. As expected, OLS regressions may require a more precise measure of the value of contradicting private information to deliver similar statistical results as IV regressions. On the other hand, based on the OLS regression results, we fail to reject the hypothesis that *unobserved* with medium quality signals probabilistically best respond to the value of their available information. Indeed, the vertical distance between the OLS fitted line of *unobserved* with medium quality signals and (0.5, 0.5) is insignificant for all *sitcount* minimum thresholds (two-tailed p-values > 0.1). This contrasts with IV regression results according to which the null that *unobserved* with medium quality signals are better responders to *value_contra_PI* is rejected. As emphasized in subsection 2.3 of the main text, our contention is that *unobserved* with medium quality signals are quite successful in learning from others.



Notes: i) ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

ii) The four colored curves are the fitted curves from the OLS regression reported in the third column of Table B3.

Figure B4: Responses to the Empirical Value of Contradicting Private Information (*sitcount* $\geqslant 20$)



Notes: i) ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

ii) The four colored curves are the fitted curves from the OLS regression reported in the fourth column of Table B3.

Figure B5: Responses to the Empirical Value of Contradicting Private Information (*sitcount* $\geqslant 30$)

B.4. Dynamics of *Observed* Guesses

Table B4 reports the percentage of *observed* guesses that contradict private information by the signal and for different majorities of public guesses in the three non-practice parts (and averaged across them).

History of public guesses		All parts b	All parts o	Part 1 b	Part 1 o	Part 2 b	Part 2 o	Part 3 b	Part 3 o
Favoring majority		02% (889)	02% (745)	03% (301)	02% (273)	02% (316)	04% (216)	00% (272)	01% (256)
No majority		03% (796)	06% (781)	03% (257)	08% (254)	04% (271)	06% (264)	02% (268)	05% (263)
Contrary majority of size	1	14% (292)	20% (296)	13% (091)	24% (086)	16% (091)	17% (109)	12% (110)	19% (101)
	2	63% (136)	63% (141)	61% (041)	72% (047)	73% (044)	65% (048)	55% (051)	52% (046)
	3	81% (106)	84% (107)	76% (034)	87% (038)	86% (036)	91% (032)	81% (036)	76% (037)
	4	89% (062)	87% (076)	86% (022)	89% (027)	91% (023)	88% (026)	88% (017)	83% (023)
	$\geqslant 5$	91% (043)	92% (066)	89% (018)	96% (023)	82% (011)	92% (025)	100% (014)	89% (018)

Note: In each cell, the first row reports the percentage of guesses that contradict private information and the second row reports the number of guesses.

Table B4: Percentages of *Observed* Guesses that Contradict Private Information

Except in the few situations where the contrary majority is very large (of size $\geqslant 5$), *observed* guesses are more informative in part 3 than in the first two parts where they are about equally informative. This is particularly striking when *observed* face contrary majorities of size 2: the relative frequency of contradicting private information averaged over the two signal realizations is only 54% in part 3 while it is 67% in part 1 and 68% in part 2. While our dynamics are weaker than in March and Ziegelmeyer (2016), we confirm that *observed* guesses become more informative over time.

Appendix C. Detailed Results for Experiments 2-4

This appendix details the data analysis in Experiments 2-4. First, we examine the nature of the histories of public guesses in the different decision periods, i.e., the *observed* guesses that have been publicly revealed up to the (beginning of the) relevant period. Second, we measure the empirical success of observational learning. Finally, we discuss the dynamics of *unobserved* guesses over the three non-practice parts.

C.1. Histories of Public Guesses

Table C1 (resp. C2) shows the distributions of majority sizes of public signals (resp. Bayes-rational guesses) in each period derived from the 72 repetitions of the 2S3Q game in Experiment 2 (resp. 3).

	Difference between the number of blue and orange public signals														
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
Period 2							45.83% (49.29%)		54.17% (50.71%)						
Period 3						23.61% (24.80%)		48.61% (48.98%)		27.78% (26.22%)					
Period 4					9.72% (12.73%)		37.50% (36.21%)		36.11% (37.26%)		16.67% (13.80%)				
Period 5				4.17% (6.65%)		22.22% (24.29%)		45.83% (35.99%)		20.83% (25.69%)		6.94% (7.38%)			
Period 6			1.39% (3.54%)		13.89% (15.58%)		36.11% (29.56%)		27.78% (30.42%)		18.06% (16.90%)		2.78% (4.00%)		
Period 7		0% (1.91%)		8.33% (9.78%)		22.22% (22.31%)		26.39% (29.38%)		30.56% (23.59%)		11.11% (10.85%)		1.39% (2.19%)	
Period 8	0% (1.04%)		6.94% (6.06%)		15.28% (16.03%)		25.00% (25.34%)		20.83% (26.07%)		23.61% (17.39%)		6.94% (6.86%)		1.39% (1.21%)

Note: In gray and between parentheses we report the expected percentage according to the state-dependent Bernoulli distribution with parameter value 12/21.

Table C1: Distributions of Majority Sizes of Public Signals

	Difference between the number of blue and orange Bayes-rational guesses														
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
Period 2							34.72% (48.33%)		65.28% (51.67%)						
Period 3						15.28% (26.11%)		19.44% (22.22%)		65.28% (51.67%)					
Period 4					15.28% (26.11%)		9.72% (10.74%)		9.72% (11.48%)		65.28% (51.67%)				
Period 5				15.28% (26.11%)		6.94% (5.80%)		2.78% (4.94%)		9.72% (11.48%)		65.28% (51.67%)			
Period 6			15.28% (26.11%)		6.94% (5.80%)		1.39% (2.39%)		1.39% (2.55%)		9.72% (11.48%)		65.28% (51.67%)		
Period 7		15.28% (26.11%)		6.94% (5.80%)		0% (1.29%)		1.39% (1.10%)		1.39% (2.55%)		9.72% (11.48%)		65.28% (51.67%)	
Period 8	15.28% (26.11%)		6.94% (5.80%)		0% (1.29%)		0% (0.53%)		1.39% (0.57%)		1.39% (2.55%)		9.72% (11.48%)		65.28% (51.67%)

Note: In gray and between parentheses we report the expected percentage according to Bayesian rationality.

Table C2: Distributions of Majority Sizes of Bayes-rational guesses

As expected, *unobserved* mostly face short majorities of public signals in Experiment 2. For example, (almost) 85% of the majorities of public signals are of size less than or equal to 3 in period 8. On the other hand, most histories are information cascades in the later periods of Experiment 3 which implies that *unobserved* mostly face long majorities of Bayes-rational guesses (from period 3 on, 80% of the histories are cascades). The only marked difference between the empirical and expected distributions of histories concerns the ratio of B and O-cascades: though only twice more B than O-cascades were expected, there are four times more B than O-cascades in Experiment 3.

Table C3 shows the distributions of histories of public guesses in each period derived from the 72 repetitions of the $2S3Q$ game in Experiment 4. For the sake of space, we shorten the notation of histories—for example histories $BBOBB$ and $OOOO$ are shortened to $2BO2B$ and $4O$—and from period 5 on we only report histories which occur at least 4 times.

Period 2	B 63%	O 37%						
Period 3	$2B$ 39%	BO 24%	OB 18%	$2O$ 19%				
Period 4	$3B$ 33%	$2BO$ 06%	BOB 14%	$B2O$ 10%	$O2B$ 08%	OBO 10%	$2OB$ 04%	$3O$ 15%
Period 5	$4B$ 31%	$2BOB$ 06%	$BO2B$ 11%	$B3O$ 06%	$O3B$ 08%	$OB2O$ 07%	$4O$ 13%	
Period 6	$5B$ 31%	$2BO2B$ 06%	$BO3B$ 07%	$B4O$ 06%	$O4B$ 08%	$5O$ 11%		
Period 7	$6B$ 28%	$BO4B$ 06%	$B5O$ 06%	$O5B$ 08%	$6O$ 10%			
Period 8	$7B$ 28%	$BO5B$ 06%	$B6O$ 06%	$O6B$ 07%	$7O$ 10%			

Table C3: Distributions of Public Histories in Experiment 4

As in Experiment 1, many empirical histories are such that guess O follows guess B (almost a quarter of histories that occur in period 3 or later) and some empirical histories are such that either guess O follows guesses BB or guess B follows guesses OO (together, one tenth of histories that occur in period 4 or later). These non-Bayes rational guesses imply that empirical histories are more diverse than predicted. As in Experiment 1, the predicted distribution of final histories differs significantly from the empirical one (Chi-square test; $p-$value < 0.01) where only 38% of the final histories are full laboratory cascades. The only noticeable difference between histories of *observed* guesses in the two experiments is that the ratio of B to O-cascades in Experiment 1 is about half the ratio in Experiment 4.

C.2. Empirical Success of Observational Learning

C.2.1 Proportions and Values of Contradicting Private Information

Table C4 reports the percentage of (human) guesses that contradict private information in each experiment by the signal of each role and for different majorities of public guesses. We group together large contrary and favoring majorities as fewer data are available for majorities of size 4 or more, especially in Experiment 2. The table also reports in each cell the average $tvcPI$ (resp. $vcPI$) in Experiments 2-3 (resp. Experiment 4) and the number of guesses. As discussed extensively in the main text, several of the regularities found in Experiment 1 are also present in Experiments 2-4: i) subjects often guess in accordance with their private information at favoring and no majorities; ii) the higher their signal quality the more often *unobserved* follow their private information; iii) subjects contradict their private information more frequently with orange than with blue signals at contrary majorities of size 1, but the difference vanishes at larger contrary majorities; and iv) subjects' propensity to contradict private information increases with the size of the contrary majority.

History of public guesses		Experiment 2						Experiment 3						Experiment 4							
		Unobserved **signal quality**						***Unobserved*** **signal quality**								***Unobserved*** **signal quality**					
		Low		Medium		High		Low		Medium		High		***Observed***		Low		Medium		High	
		b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o
Favoring majority of size	$\geqslant 4$	**03%** .152 (098)	**10%** .208 (041)	**05%** .107 (096)	**00%** .162 (008)	**01%** .035 (096)	**00%** .055 (053)	**05%** .235 (646)	**02%** .186 (164)	**03%** .170 (776)	**05%** .133 (108)	**03%** .064 (812)	**02%** .048 (168)	**03%** .132 (122)	**08%** .134 (037)	**03%** .196 (478)	**06%** .152 (120)	**03%** .137 (390)	**01%** .140 (171)	**01%** .050 (447)	**02%** .067 (179)
	3	**04%** .206 (207)	**08%** .279 (099)	**05%** .147 (221)	**09%** .205 (056)	**02%** .054 (169)	**01%** .079 (113)	**07%** .235 (184)	**02%** .186 (055)	**02%** .170 (202)	**06%** .133 (033)	**03%** .064 (211)	**04%** .048 (046)	**05%** .125 (102)	**02%** .140 (045)	**05%** .180 (193)	**08%** .169 (066)	**04%** .127 (154)	**02%** .131 (087)	**02%** .045 (172)	**02%** .060 (101)
	2	**06%** .257 (220)	**09%** .340 (139)	**07%** .187 (233)	**05%** .256 (131)	**02%** .071 (178)	**02%** .103 (181)	**07%** .235 (184)	**00%** .186 (055)	**03%** .170 (202)	**03%** .133 (033)	**03%** .064 (211)	**00%** .048 (046)	**02%** .139 (171)	**04%** .156 (077)	**08%** .199 (234)	**12%** .207 (068)	**05%** .130 (216)	**04%** .165 (098)	**04%** .051 (199)	**02%** .061 (104)
	1	**04%** .315 (322)	**15%** .407 (296)	**03%** .235 (310)	**07%** .314 (367)	**03%** .093 (265)	**03%** .133 (366)	**07%** .235 (192)	**10%** .314 (142)	**05%** .170 (210)	**06%** .234 (085)	**03%** .064 (219)	**04%** .092 (094)	**03%** .187 (285)	**07%** .240 (175)	**07%** .266 (272)	**09%** .327 (164)	**04%** .192 (275)	**06%** .253 (179)	**03%** .072 (236)	**01%** .099 (165)
No majority		**08%** .380 (544)	**19%** .478 (488)	**03%** .290 (534)	**09%** .379 (538)	**03%** .120 (517)	**04%** .169 (507)	**11%** .380 (346)	**21%** .478 (318)	**06%** .290 (285)	**09%** .379 (299)	**03%** .120 (301)	**09%** .169 (275)	**05%** .292 (493)	**12%** .378 (522)	**06%** .384 (347)	**13%** .473 (375)	**02%** .296 (358)	**09%** .373 (356)	**01%** .121 (329)	**02%** .169 (377)
Contrary majority of size	1	**23%** .450 (328)	**53%** .550 (286)	**09%** .353 (361)	**21%** .449 (298)	**08%** .154 (386)	**10%** .214 (279)	**56%** .551 (154)	**67%** .647 (176)	**39%** .450 (075)	**52%** .550 (214)	**13%** .214 (082)	**28%** .289 (181)	**16%** .444 (153)	**34%** .523 (293)	**43%** .537 (134)	**51%** .613 (264)	**18%** .431 (159)	**29%** .515 (257)	**06%** .203 (141)	**07%** .265 (260)
	2	**61%** .522 (141)	**76%** .620 (204)	**32%** .421 (133)	**56%** .521 (207)	**16%** .195 (179)	**18%** .266 (182)	**93%** .711 (057)	**83%** .647 (168)	**78%** .621 (023)	**75%** .550 (206)	**50%** .353 (034)	**40%** .289 (173)	**57%** .577 (063)	**65%** .612 (174)	**66%** .676 (056)	**69%** .695 (226)	**51%** .554 (090)	**56%** .640 (208)	**26%** .290 (068)	**18%** .366 (205)
	3	**75%** .593 (101)	**83%** .685 (201)	**54%** .492 (056)	**71%** .592 (203)	**32%** .244 (095)	**30%** .326 (159)	**95%** .711 (057)	**83%** .647 (168)	**78%** .621 (023)	**79%** .550 (206)	**50%** .353 (034)	**51%** .289 (173)	**84%** .616 (037)	**85%** .633 (110)	**73%** .736 (048)	**78%** .717 (195)	**69%** .584 (083)	**68%** .637 (146)	**39%** .319 (061)	**32%** .368 (174)
	$\geqslant 4$	**74%** .678 (031)	**85%** .758 (094)	**88%** .564 (008)	**76%** .674 (080)	**42%** .318 (019)	**54%** .439 (096)	**90%** .711 (172)	**85%** .647 (602)	**82%** .621 (068)	**82%** .550 (792)	**55%** .353 (120)	**53%** .289 (660)	**96%** .635 (045)	**93%** .617 (120)	**82%** .759 (078)	**81%** .698 (458)	**70%** .575 (179)	**72%** .613 (370)	**68%** .344 (107)	**44%** .357 (451)

Note: In each cell, from top to bottom: percentage of guesses that contradict private information, $tvcPI$ (resp. $vcPI$) in Experiments 2-3 (resp. Experiment 4), and number of guesses.

Table C4: Proportions and Values of Contradicting Private Information

C.2.2 Values of Contradicting Private Information

Table C5 (resp. C6) reports the true values of contradicting private information in Experiment 2 (resp. 3) by majorities of public guesses and by roles, distinguishing between signal qualities for *unobserved*. Since all favoring majorities are grouped together, the table reports the mean of *tvcPI* in the first row.

History of public guesses		*Observed* Medium quality b	o	*Unobserved* Low quality b	o	Medium quality b	o	High quality b	o
Favoring majority		**.28** (410)	**.38** (334)	**.26** (706)	**.38** (304)	**.19** (662)	**.30** (381)	**.07** (529)	**.12** (463)
No majority		**.38** (567)	**.48** (566)	**.38** (485)	**.48** (422)	**.29** (496)	**.38** (519)	**.12** (477)	**.17** (465)
Contrary majority of size	1	**.45** (251)	**.55** (248)	**.45** (226)	**.55** (235)	**.35** (248)	**.45** (232)	**.15** (314)	**.21** (234)
	2	**.52** (061)	**.62** (106)	**.52** (063)	**.62** (172)	**.42** (082)	**.52** (147)	**.20** (117)	**.27** (132)
	3	**.59** (014)	**.68** (051)	**.59** (039)	**.68** (158)	**.49** (016)	**.59** (146)	**.24** (031)	**.33** (123)
	4	— (000)	**.74** (013)	— (000)	**.74** (051)	— (000)	**.66** (050)	**.30** (011)	**.39** (048)
	$\geqslant 5$	— (000)	— (000)	— (000)	**.79** (011)	— (000)	— (000)	— (000)	— (000)

Note: Each cell contains *tvcPI* and the number of individual observations.

Table C5: True Values of Contradicting Private Information in Experiment 2

History of public guesses		*Observed* Medium quality b	o	*Unobserved* Low quality b	o	Medium quality b	o	High quality b	o
Favoring majority		**.17** (857)	**.18** (278)	**.23** (1183)	**.23** (408)	**.17** (1382)	**.16** (251)	**.06** (1445)	**.06** (314)
No majority		**.29** (457)	**.38** (414)	**.38** (338)	**.48** (310)	**.29** (269)	**.38** (283)	**.12** (285)	**.17** (259)
Contrary majority of size	1	**.45** (133)	**.55** (215)	**.55** (146)	**.65** (165)	**.45** (067)	**.55** (206)	**.21** (074)	**.29** (173)
	2	**.62** (047)	**.55** (180)	**.71** (057)	**.65** (165)	**.62** (023)	**.55** (206)	**.35** (026)	**.29** (173)
	3	**.62** (028)	**.55** (139)	**.71** (057)	**.65** (165)	**.62** (023)	**.55** (206)	**.35** (026)	**.29** (173)
	4	**.62** (019)	**.55** (096)	**.71** (057)	**.65** (165)	**.62** (023)	**.55** (206)	**.35** (026)	**.29** (173)
	$\geqslant 5$	**.62** (011)	**.55** (091)	**.71** (115)	**.65** (437)	**.62** (045)	**.55** (586)	**.35** (078)	**.29** (487)

Note: Each cell contains *tvcPI* and the number of individual observations.

Table C6: True Values of Contradicting Private Information in Experiment 3

Table C7 reports the empirical values of contradicting private information in Experiment 4 by majorities of public guesses and by roles, distinguishing between signal qualities for *unobserved*. In each cell, the first row displays the mean of *vcPI*, the second row displays the first and ninth deciles of *vcPI*, and the third row displays the total number of individual observations included in all guessing situations.

History of public guesses		***Observed*** Medium quality		***Unobserved*** Low quality		Medium quality		High quality	
		b	*o*	*b*	*o*	*b*	*o*	*b*	*o*
Favoring majority		**.16** .12 – .18 (599)	**.19** .11 – .26 (266)	**.22** .18 – .24 (1043)	**.23** .14 – .35 (295)	**.16** .13 – .18 (1090)	**.18** .10 – .26 (395)	**.06** .05 – .07 (1200)	**.07** .04 – .12 (414)
No majority		**.29** .27 – .29 (468)	**.38** .38 – .40 (498)	**.38** .36 – .40 (309)	**.48** .46 – .50 (338)	**.29** .27 – .30 (334)	**.37** .36 – .40 (332)	**.12** .11 – .13 (306)	**.17** .16 – .18 (354)
Contrary majority of size	1	**.44** .38 – .46 (137)	**.53** .48 – .54 (282)	**.55** .51 – .56 (086)	**.62** .58 – .64 (238)	**.44** .38 – .46 (122)	**.52** .46 – .54 (205)	**.21** .17 – .22 (102)	**.27** .22 – .28 (231)
	2	**.60** .52 – .62 (052)	**.60** .58 – .61 (150)	**.68** .62 – .71 (041)	**.69** .64 – .70 (163)	**.55** .37 – .62 (069)	**.65** .61 – .61 (129)	**.29** .17 – .35 (052)	**.34** .31 – .34 (150)
	3	**.67** .67 – .67 (026)	**.64** .62 – .64 (090)	**.75** .75 – .75 (018)	**.72** .65 – .73 (134)	**.58** .35 – .67 (046)	**.64** .64 – .64 (086)	**.31** .15 – .40 (031)	**.37** .36 – .37 (120)
	4	**.66** .66 – .66 (018)	**.62** .62 – .62 (053)	**.75** .75 – .75 (018)	**.70** .65 – .71 (105)	**.57** .33 – .66 (046)	**.62** .62 – .62 (074)	**.30** .14 – .40 (030)	**.36** .36 – .36 (088)
	$\geqslant 5$	**.67** .67 – .67 (014)	**.61** .60 – .61 (053)	**.76** .75 – .77 (051)	**.69** .69 – .70 (261)	**.63** .24 – .69 (100)	**.60** .59 – .61 (216)	**.37** .10 – .43 (068)	**.34** .33 – .34 (250)

Note: In each cell, from top to bottom: mean of *vcPI*, $1^{st} - 9^{th}$ deciles of *vcPI*, and number of individual observations.

Table C7: Empirical Values of Contradicting Private Information in Experiment 4

Empirical values of contradicting private information in Experiment 4 are similar to those in Experiment 1. In particular, for each role and each signal quality, the empirically optimal guess at favoring majorities consists in following private information. The average incentives to act in accordance with private information are at least three times stronger than the average incentives to contradict private information. Further down the table incentives to contradict private information increase till the contrary majority reaches size 3, but then incentive levels hardly vary with additional contrary guesses. According to the estimated values of contradicting private information, contrary majorities of size $\geqslant 2$ aggregate about two private signals. In the last two columns of the table, the incentives to follow private information are at least 1.5 times stronger than the incentives to contradict private information. *Unobserved* should therefore always follow their private information when endowed with high quality signals. On the other hand, *unobserved* with low quality signals should herd at contrary majorities of any size and they should follow their private information otherwise (columns 4-5). Still, incentives to make the empirically optimal guess are weak in the case of no majority and an orange signal. Given the average incentive levels in columns 2-3 and 6-7 of the table, subjects with medium quality signals should follow (resp. contradict) their private information at favoring and no majorities (resp. at contrary majorities of size $\geqslant 2$). At contrary majorities of size 1 they should follow (resp. contradict) their private information with a blue signal (resp. an orange signal) though, as expected, incentives to make the empirically optimal guess are rather weak.

C.2.3 Responses to the Value of Contradicting Private Information

Here we report the regression results discussed in subsection 3.2.2 of the main text. In Experiments 2 and 3, we regress the proportion to contradict private information against a cubic polynomial in $tvcPI$ fully interacted with indicator variables for the signal quality of *unobserved*. Table C8 reports the regression results.

	Experiment 2	Experiment 3
Constant	0.525***	0.899***
	(0.097)	(0.167)
Low	0.408***	-0.154
	(0.127)	(0.216)
High	-0.474***	-0.642***
	(0.095)	(0.175)
$tvcPI$	-4.887***	-8.919***
	(0.945)	(1.824)
$(tvcPI)^2$	13.172***	25.186***
	(2.763)	(5.671)
$(tvcPI)^3$	-7.639***	-17.280***
	(2.357)	(5.287)
$Low \times tvcPI$	-2.992***	3.060
	(1.045)	(1.921)
$Low \times (tvcPI)^2$	6.395**	-11.408**
	(2.748)	(5.375)
$Low \times (tvcPI)^3$	-4.449*	10.146**
	(2.263)	(4.792)
$High \times tvcPI$	3.799***	2.931
	(0.965)	(2.317)
$High \times (tvcPI)^2$	-4.850	15.189
	(3.551)	(12.653)
$High \times (tvcPI)^3$	0.237	-42.351*
	(4.673)	(22.519)
Observations	11,520	11,520
R^2	0.326	0.491

Notes: i) Robust standard errors in parentheses, clustered at the session level.
ii) * (10%); ** (5%); and *** (1%) significance level.

Table C8: Propensity to Contradict Private Information in Experiments 2-3

In Experiment 4, we regress the proportion to contradict private information against a cubic polynomial in $vcPI$ fully interacted with indicator variables for the role and the signal quality of *unobserved*. Table C9 reports the regression results based on the IV specification for $sitcount \geqslant 10$ (minimum threshold considered in the main text) as well as for two robustness checks with $sitcount \geqslant 1$ and $sitcount \geqslant 20$. Table C10 reports the regression results based on the OLS specification for $sitcount \geqslant 10$, $sitcount \geqslant 20$, and for $sitcount \geqslant 30$.

	$sitcount \geqslant 1$	$sitcount \geqslant 10$	$sitcount \geqslant 20$
Constant	-0.004	0.007	-0.085
	(0.082)	(0.080)	(0.081)
$vcPI_1$	0.633	0.509	1.560
	(0.948)	(0.929)	(0.970)
$(vcPI_1)^2$	-3.476	-3.147	-6.644*
	(3.210)	(3.159)	(3.429)
$(vcPI_1)^3$	6.906**	6.650**	9.953***
	(2.982)	(2.942)	(3.265)
Low	0.297*	0.286	0.263
	(0.176)	(0.177)	(0.172)
$Low \times vcPI_1$	-2.698	-2.574	-2.544
	(1.650)	(1.664)	(1.684)
$Low \times (vcPI_1)^2$	7.284	6.955	7.513
	(4.862)	(4.907)	(5.115)
$Low \times (vcPI_1)^3$	-6.860*	-6.604	-7.518*
	(4.137)	(4.176)	(4.458)
$Medium$	0.236**	0.225*	0.310***
	(0.115)	(0.115)	(0.116)
$Medium \times vcPI_1$	-2.815**	-2.691**	-3.654***
	(1.251)	(1.243)	(1.296)
$Medium \times (vcPI_1)^2$	8.961**	8.632**	11.811***
	(4.141)	(4.113)	(4.438)
$Medium \times (vcPI_1)^3$	-8.467**	-8.211**	-11.210***
	(3.915)	(3.889)	(4.293)
$High$	0.059	0.048	0.124**
	(0.065)	(0.062)	(0.062)
$High \times vcPI_1$	-1.473***	-1.349***	-2.039***
	(0.473)	(0.440)	(0.484)
$High \times (vcPI_1)^2$	6.315***	5.986***	7.442***
	(1.549)	(1.510)	(1.991)
$High \times (vcPI_1)^3$	-2.037	-1.781	-1.743
	(3.771)	(3.791)	(4.854)
Observations	10,631	10,590	10,171
R^2	0.410	0.411	0.403

Notes: i) Robust standard errors in parentheses, clustered at the session level.
ii) * (10%); ** (5%); and *** (1%) significance level.

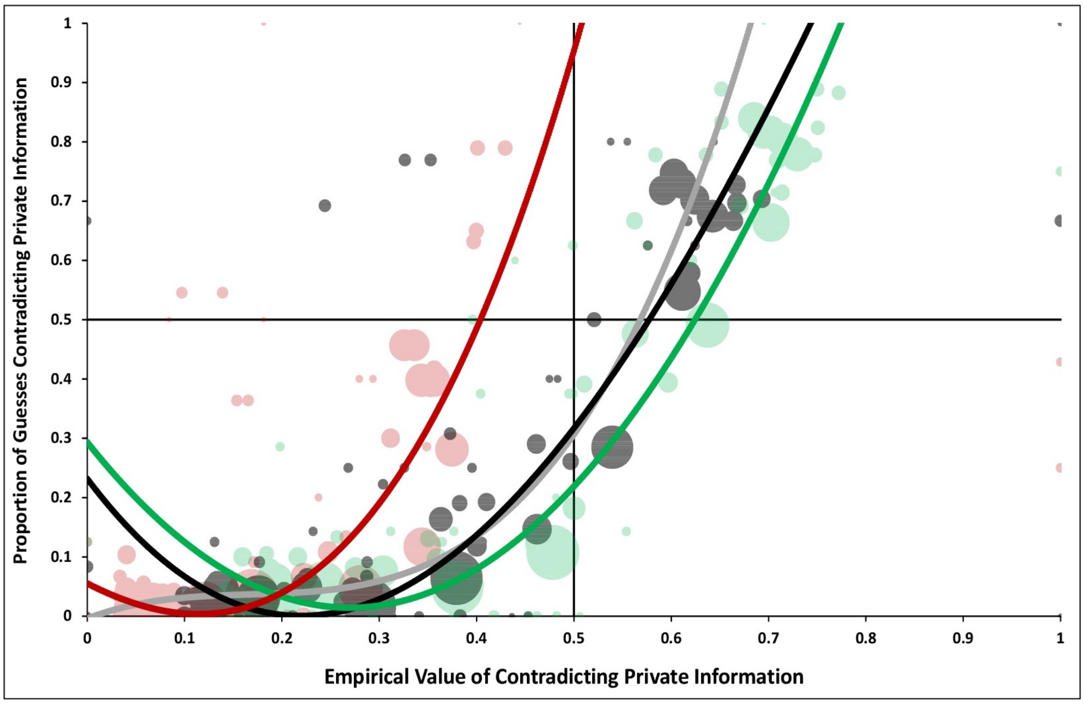
Table C9: Propensity to Contradict Private information in Experiment 4 (IV regressions)

	$sitcount \geqslant 10$	$sitcount \geqslant 20$	$sitcount \geqslant 30$
Constant	-0.079	-0.204	-0.259
	(0.095)	(0.101)	(0.133)
$vcPI$	1.514	2.925*	3.591*
	(1.059)	(1.178)	(1.540)
$(vcPI)^2$	-6.513	-11.056**	-13.344*
	(3.520)	(4.139)	(5.349)
$(vcPI)^3$	10.014**	14.086**	16.353**
	(3.322)	(4.001)	(5.220)
Low	0.506**	0.375	0.409
	(0.168)	(0.195)	(0.236)
$Low \times vcPI$	-4.868**	-3.930	-4.254
	(1.667)	(2.003)	(2.290)
$Low \times (vcPI)^2$	13.934**	12.210	13.203
	(5.061)	(6.286)	(7.017)
$Low \times (vcPI)^3$	-12.832**	-12.002*	-13.052*
	(4.406)	(5.586)	(6.304)
$Medium$	0.380***	0.413**	0.407*
	(0.080)	(0.148)	(0.186)
$Medium \times vcPI$	-4.723***	-4.880**	-4.716*
	(0.820)	(1.698)	(2.085)
$Medium \times (vcPI)^2$	16.181***	15.850**	15.140*
	(2.671)	(5.832)	(7.041)
$Medium \times (vcPI)^3$	-15.960***	-15.033**	-14.298*
	(2.651)	(5.682)	(6.841)
$High$	0.057	0.243**	0.322**
	(0.079)	(0.086)	(0.127)
$High \times vcPI$	-0.273	-3.407**	-4.819**
	(0.794)	(0.894)	(1.586)
$High \times (vcPI)^2$	-3.347	11.592**	19.265*
	(3.703)	(4.017)	(7.828)
$High \times (vcPI)^3$	16.424**	-4.955	-17.411
	(6.272)	(7.503)	(13.262)
Observations	11,650	10,595	9,775
R^2	0.411	0.407	0.409

Notes: i) Robust standard errors in parentheses, clustered at the session level.
ii) * (10%); ** (5%); and *** (1%) significance level.

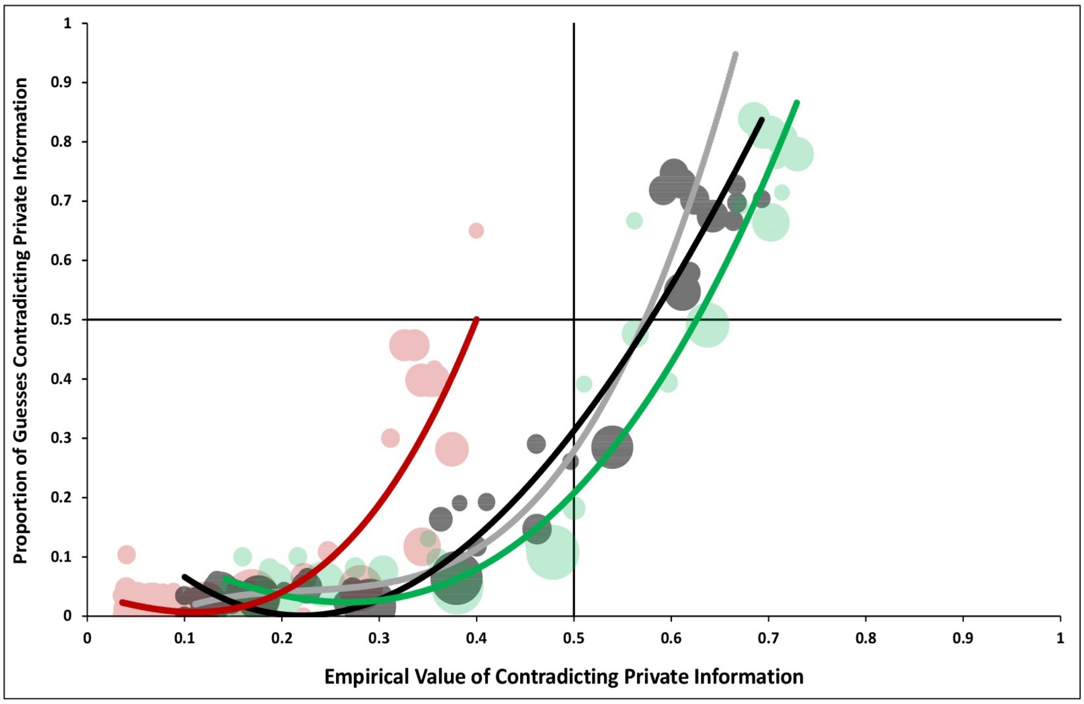
Table C10: Propensity to Contradict Private information in Experiment 4 (OLS regressions)

Figure C1 (resp. C2) plots $vcPI$ against the proportion of contradictions collected in identical guessing situations with $sitcount \geqslant 1$ (resp. $sitcount \geqslant 20$), and it superimposes fitted curves from the IV regressions. Clearly, the average responses to $vcPI$ predicted by IV regressions are basically identical for samples of guessing situations with $sitcount \geqslant 1$, 10 or 20.



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

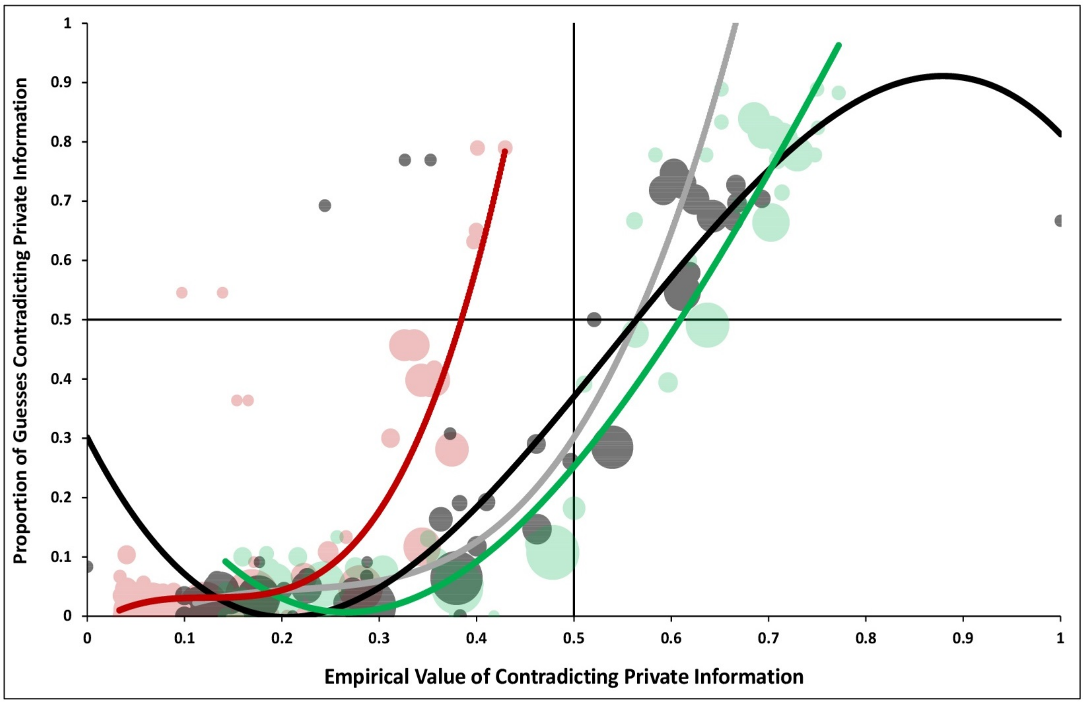
Figure C1: Responses to $vcPI$ in Experiment 4 ($sitcount \geqslant 1$)



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

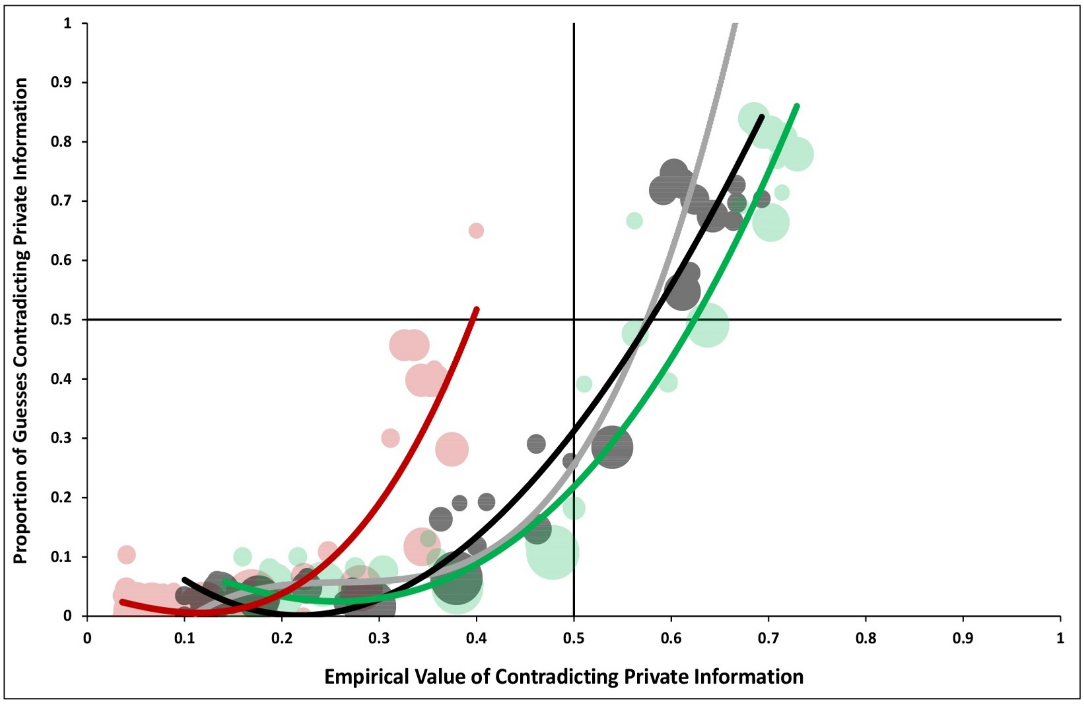
Figure C2: Responses to $vcPI$ in Experiment 4 ($sitcount \geqslant 20$)

Figure C3 (resp. C4 and C5) plots $vcPI$ against the proportion of contradictions collected in identical guessing situations with $sitcount \geqslant 10$ (resp. $sitcount \geqslant 20$ and $sitcount \geqslant 30$), and it superimposes fitted curves from the OLS regressions.



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

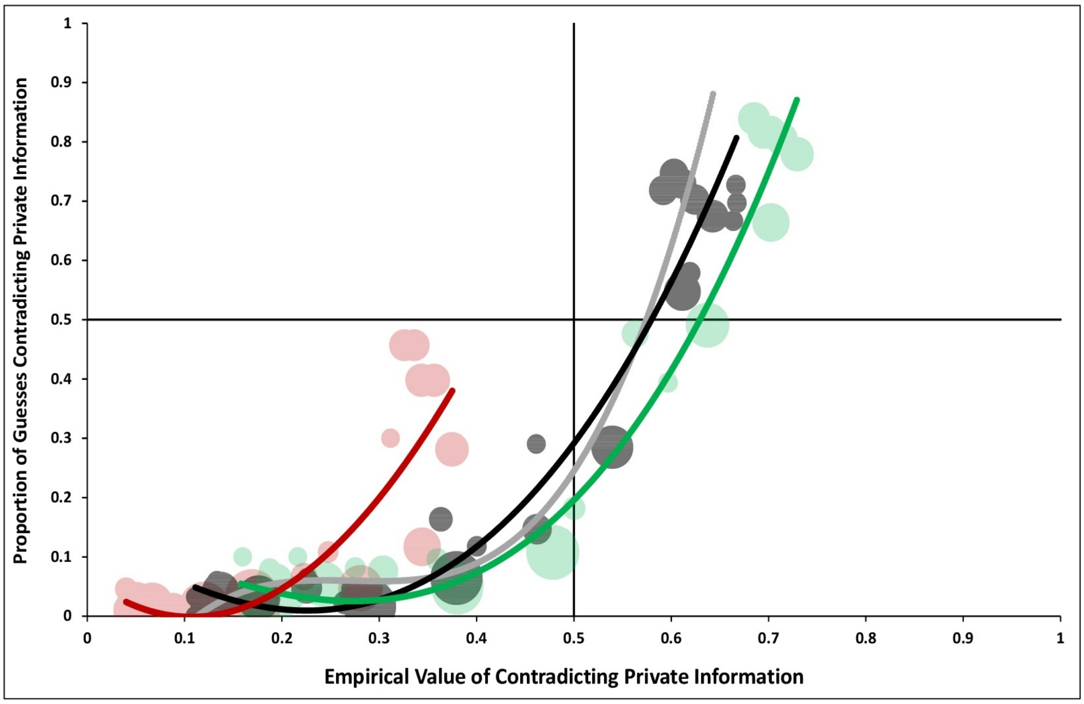
Figure C3: Responses to $vcPI$ in Experiment 4 (OLS regressions, $sitcount \geqslant 10$)



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

Figure C4: Responses to $vcPI$ in Experiment 4 (OLS regressions, $sitcount \geqslant 20$)

The average responses to $vcPI$ predicted by OLS regressions are similar to those predicted by IV regressions, with the minor exception that *observed* are predicted to be slightly more reluctant to contradict their private information when incentives to follow others are weak.



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively; ●: *Observed* guesses.

Figure C5: Responses to $vcPI$ in Experiment 4 (OLS regressions, $sitcount \geqslant 30$)

C.3. Dynamics of *Unobserved* Guesses

Table C11 (resp. C12 and C13) reports the percentage of *unobserved* guesses that contradict private information in the first (resp. second and third) part of each experiment by the signal and for different majorities of public guesses.

The comparison of the percentages of contradictions in the three parts of Experiments 2-4 reveals the absence of strong dynamics in the herding behavior of *unobserved*. Indeed, for a given signal quality, we never observe a similar variation of the percentage of contradictions as the session progresses in all three experiments. And, in a given experiment, we never observe a similar variation of the percentage of contradictions as the session progresses for all three qualities. In Experiments 3 and 4, however, *unobserved* with high quality signals who face contrary majorities tend to contradict their private information less as the session progresses. There is no such tendency in Experiment 2. Also, we note that the regularities described in subsection C.2.1 are (almost) systematically observed in each of the three parts: i) subjects often guess in accordance with their private information at favoring and no majorities; ii) the higher their signal quality the more often *unobserved* follow their private information; iii) subjects contradict their private information more frequently with orange than with blue signals at contrary majorities of size 1, but the difference vanishes at larger contrary majorities; and iv) subjects' propensity to contradict private information increases with the size of the contrary majority.

History of public guesses		Experiment 2						Experiment 3						Experiment 4					
		Unobserved **signal quality**						***Unobserved*** **signal quality**						***Unobserved*** **signal quality**					
		Low		Medium		High		Low		Medium		High		Low		Medium		High	
		b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o
Favoring majority of size	$\geqslant 4$	**00%** (010)	**09%** (033)	**13%** (024)	— (000)	— (000)	**00%** (045)	**06%** (118)	**00%** (052)	**03%** (136)	**08%** (060)	**01%** (236)	**05%** (056)	**00%** (070)	**08%** (049)	**00%** (042)	**00%** (068)	**00%** (062)	**01%** (092)
	3	**00%** (015)	**04%** (027)	**10%** (029)	— (000)	**00%** (017)	**02%** (041)	**15%** (040)	**00%** (015)	**00%** (034)	**12%** (017)	**03%** (059)	**14%** (014)	**07%** (043)	**14%** (021)	**00%** (028)	**00%** (034)	**00%** (023)	**02%** (048)
	2	**11%** (044)	**07%** (027)	**18%** (049)	**04%** (027)	**00%** (026)	**04%** (045)	**13%** (040)	**00%** (015)	**09%** (034)	**06%** (017)	**02%** (059)	**00%** (014)	**06%** (052)	**14%** (014)	**04%** (028)	**00%** (044)	**00%** (025)	**04%** (050)
	1	**03%** (098)	**10%** (040)	**05%** (094)	**08%** (079)	**06%** (065)	**04%** (054)	**15%** (040)	**03%** (030)	**06%** (034)	**14%** (021)	**03%** (059)	**00%** (014)	**07%** (068)	**04%** (047)	**00%** (033)	**03%** (070)	**00%** (032)	**04%** (048)
No majority		**07%** (136)	**13%** (080)	**06%** (118)	**12%** (122)	**06%** (109)	**05%** (099)	**13%** (082)	**19%** (054)	**02%** (045)	**12%** (059)	**02%** (061)	**06%** (035)	**05%** (074)	**11%** (102)	**01%** (077)	**09%** (075)	**00%** (048)	**01%** (096)
Contrary majority of size	1	**31%** (072)	**53%** (062)	**22%** (073)	**23%** (082)	**15%** (074)	**15%** (079)	**48%** (042)	**54%** (024)	**18%** (011)	**34%** (038)	**00%** (002)	**29%** (021)	**24%** (017)	**53%** (060)	**14%** (050)	**40%** (015)	**08%** (024)	**07%** (056)
	2	**66%** (029)	**68%** (028)	**45%** (029)	**57%** (023)	**28%** (043)	**23%** (030)	**94%** (017)	**75%** (024)	**71%** (007)	**74%** (038)	**00%** (002)	**52%** (021)	**100%** (002)	**80%** (044)	**56%** (036)	**90%** (020)	**43%** (014)	**23%** (031)
	3	**72%** (029)	**78%** (009)	— (000)	**91%** (011)	**30%** (023)	**14%** (007)	**94%** (017)	**79%** (024)	**57%** (007)	**79%** (038)	**50%** (002)	**71%** (021)	**100%** (003)	**87%** (045)	**80%** (030)	**90%** (020)	**88%** (008)	**36%** (025)
	$\geqslant 4$	**74%** (023)	**83%** (006)	— (000)	**100%** (008)	**36%** (011)	— (000)	**87%** (060)	**84%** (074)	**65%** (020)	**82%** (152)	**50%** (008)	**71%** (084)	**71%** (007)	**96%** (050)	**72%** (076)	**82%** (022)	**100%** (020)	**48%** (066)

Note: Each cell contains the percentage of guesses that contradict private information and the number of guesses.

Table C11: Percentage of *Unobserved* Guesses that Contradict Private Information in Part 1

History of public guesses		Experiment 2						Experiment 3						Experiment 4					
		Unobserved signal quality						*Unobserved* signal quality						*Unobserved* signal quality					
		Low		Medium		High		Low		Medium		High		Low		Medium		High	
		b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o	b	o
Favoring majority of size	$\geqslant 4$	**04%** (024)	— (000)	**25%** (008)	— (000)	**00%** (032)	— (000)	**03%** (064)	**02%** (064)	**01%** (176)	— (000)	**02%** (112)	**00%** (064)	**03%** (124)	— (000)	**06%** (064)	**03%** (032)	**02%** (101)	**00%** (016)
	3	**06%** (048)	**17%** (024)	**15%** (048)	**13%** (008)	**00%** (008)	**00%** (024)	**04%** (024)	**00%** (024)	**02%** (048)	— (000)	**00%** (032)	**00%** (016)	**04%** (047)	— (000)	**04%** (023)	**13%** (008)	**00%** (046)	**00%** (008)
	2	**04%** (048)	**03%** (032)	**07%** (056)	**08%** (024)	**00%** (024)	**00%** (056)	**04%** (024)	**00%** (024)	**00%** (048)	— (000)	**03%** (032)	**00%** (016)	**04%** (047)	**00%** (008)	**04%** (053)	**13%** (008)	**05%** (039)	**00%** (008)
	1	**03%** (064)	**10%** (048)	**05%** (056)	**08%** (080)	**00%** (040)	**03%** (104)	**08%** (024)	**13%** (056)	**04%** (048)	**00%** (008)	**00%** (032)	**08%** (024)	**04%** (055)	**04%** (024)	**03%** (093)	**13%** (016)	**04%** (055)	**00%** (024)
No majority		**06%** (096)	**15%** (096)	**04%** (104)	**07%** (104)	**04%** (096)	**05%** (096)	**06%** (080)	**26%** (080)	**04%** (056)	**09%** (056)	**04%** (056)	**13%** (056)	**08%** (071)	**08%** (071)	**03%** (079)	**09%** (079)	**00%** (079)	**05%** (079)
Contrary majority of size	1	**21%** (048)	**48%** (064)	**05%** (080)	**25%** (056)	**08%** (104)	**03%** (040)	**55%** (056)	**79%** (024)	**50%** (008)	**56%** (048)	**08%** (024)	**38%** (032)	**54%** (024)	**44%** (055)	**25%** (016)	**26%** (093)	**04%** (024)	**07%** (055)
	2	**72%** (032)	**71%** (048)	**21%** (024)	**66%** (056)	**14%** (056)	**13%** (024)	**92%** (024)	**83%** (024)	— (000)	**79%** (048)	**56%** (016)	**53%** (032)	**75%** (008)	**57%** (047)	**50%** (008)	**38%** (053)	**25%** (008)	**23%** (039)
	3	**88%** (024)	**85%** (048)	**38%** (008)	**79%** (048)	**38%** (024)	**38%** (008)	**96%** (024)	**79%** (024)	— (000)	**81%** (048)	**50%** (016)	**66%** (032)	— (000)	**74%** (047)	**63%** (008)	**61%** (023)	**25%** (008)	**37%** (046)
	$\geqslant 4$	— (000)	**88%** (024)	— (000)	**75%** (008)	— (000)	**63%** (032)	**92%** (064)	**83%** (064)	— (000)	**85%** (176)	**50%** (064)	**65%** (112)	— (000)	**73%** (124)	**50%** (032)	**72%** (064)	**38%** (016)	**47%** (101)

Note: Each cell contains the percentage of guesses that contradict private information and the number of guesses.

Table C12: Percentage of *Unobserved* Guesses that Contradict Private Information in Part 2

History of public guesses		Experiment 2						Experiment 3						Experiment 4					
		Unobserved signal quality						*Unobserved* signal quality						*Unobserved* signal quality					
		Low		Medium		High		Low		Medium		High		Low		Medium		High	
		b	*o*	*b*	*o*	*b*	*o*	*b*	*o*	*b*	*o*	*b*	*o*	*b*	*o*	*b*	*o*	*b*	*o*
Favoring majority of size	$\geqslant 4$	**03%** (064)	**13%** (008)	**00%** (064)	**00%** (008)	**02%** (064)	**00%** (008)	**05%** (464)	**04%** (048)	**05%** (464)	**00%** (048)	**04%** (464)	**00%** (048)	**04%** (284)	**04%** (071)	**02%** (284)	**01%** (071)	**01%** (284)	**04%** (071)
	3	**03%** (144)	**06%** (048)	**01%** (144)	**08%** (048)	**03%** (144)	**00%** (048)	**05%** (120)	**06%** (016)	**03%** (120)	**00%** (016)	**04%** (120)	**00%** (016)	**05%** (103)	**04%** (045)	**05%** (103)	**02%** (045)	**03%** (103)	**02%** (045)
	2	**05%** (128)	**13%** (080)	**02%** (128)	**05%** (080)	**02%** (128)	**03%** (080)	**05%** (120)	**00%** (016)	**03%** (120)	**00%** (016)	**04%** (120)	**00%** (016)	**10%** (135)	**13%** (046)	**06%** (135)	**07%** (046)	**04%** (135)	**00%** (046)
	1	**05%** (160)	**17%** (208)	**01%** (160)	**07%** (208)	**02%** (160)	**03%** (208)	**04%** (128)	**11%** (056)	**05%** (128)	**04%** (056)	**03%** (128)	**04%** (056)	**07%** (149)	**13%** (093)	**05%** (149)	**08%** (093)	**03%** (149)	**00%** (093)
No majority		**09%** (312)	**22%** (312)	**01%** (312)	**09%** (312)	**01%** (312)	**03%** (312)	**11%** (184)	**20%** (184)	**07%** (184)	**08%** (184)	**04%** (184)	**09%** (184)	**06%** (202)	**15%** (202)	**02%** (202)	**09%** (202)	**01%** (202)	**02%** (202)
Contrary majority of size	1	**22%** (208)	**54%** (160)	**06%** (208)	**18%** (160)	**05%** (208)	**09%** (160)	**64%** (056)	**67%** (128)	**41%** (056)	**55%** (128)	**16%** (056)	**25%** (128)	**44%** (093)	**53%** (149)	**18%** (093)	**30%** (149)	**05%** (093)	**06%** (149)
	2	**55%** (080)	**80%** (128)	**31%** (080)	**51%** (128)	**10%** (080)	**17%** (128)	**94%** (016)	**85%** (120)	**81%** (016)	**74%** (120)	**50%** (016)	**34%** (120)	**63%** (046)	**69%** (135)	**48%** (046)	**59%** (135)	**22%** (046)	**16%** (135)
	3	**71%** (048)	**83%** (144)	**56%** (048)	**67%** (144)	**29%** (048)	**31%** (144)	**94%** (016)	**84%** (120)	**88%** (016)	**78%** (120)	**50%** (016)	**43%** (120)	**71%** (045)	**76%** (103)	**62%** (045)	**66%** (103)	**33%** (045)	**28%** (103)
	$\geqslant 4$	**75%** (008)	**84%** (064)	**88%** (008)	**73%** (064)	**50%** (008)	**50%** (064)	**92%** (048)	**86%** (464)	**90%** (048)	**80%** (464)	**63%** (048)	**47%** (464)	**83%** (071)	**82%** (284)	**77%** (071)	**71%** (284)	**66%** (071)	**42%** (284)

Note: Each cell contains the percentage of guesses that contradict private information and the number of guesses.

Table C13: Percentage of *Unobserved* Guesses that Contradict Private Information in Part 3

Appendix D. Classification of *Unobserved* to Decision Rules

In this appendix, we first describe the groups of guessing situations used to classify *unobserved* in our four experiments. Second, we detail the guesses made by the decision rules in the different groups of guessing situations. Finally, we report the results of our classification.

D.1. Guessing Situations

Table D1 describes the groups of guessing situations used for the classification in Experiments 1, 3 and 4. Since *unobserved* learn from public signals of low quality, a different set of groups of guessing situations is used in Experiment 2, which is described in Table D2. The tables also report for each group of guessing situations the (empirically) optimal guess and the average incentives to guess optimally. Note that we rely on the true incentives to make the optimal guess in Experiments 2 and 3. On the other hand, incentives to herd in Experiments 1 and 4 are given by the empirical value of contradicting private information, $vcPI$, measured at the precision level of *sitcount* $\geqslant 10$, and incentives to follow private information are given by $1 - vcPI$ (in a given group, the empirical value of contradicting private information is the average weighted by the frequency of observations in the different guessing situations).

	Group of guessing situations									
	1	2	3	4	5	6	7	8	9	10
Quality of signal(s)	Low	Medium	High	Low	Low	Medium	Medium	Medium	High	High
Signal(s)	{b, o}	{b, o}	{b, o}	{b, o}	{b, o}	b	o ¦ {b, o}	{b, o}	{b, o}	{b, o}
Size of the contrary majority		$\{-7, -6, \ldots, -1, 0\}$		$\{1, 2\}$	$\geqslant 3$	1	1 ¦ 2	$\geqslant 3$	$\{1, 2\}$	$\geqslant 3$
Optimal guess		FPI		HERD	HERD	FPI	HERD	HERD	FPI	FPI
Incentives in Exp. 1	0.717	0.766	0.904	0.649	0.750	0.567	0.552	0.578	0.729	0.663
Incentives in Exp. 3	0.711	0.789	0.918	0.628	0.662	0.550	0.554	0.556	0.720	0.702
Incentives in Exp. 4	0.711	0.778	0.909	0.634	0.710	0.563	0.567	0.611	0.722	0.653

Table D1: The 10 Groups of Guessing Situations in Experiments 1, 3 and 4

There are groups of guessing situations where *unobserved* should follow their private information, referred to as FPI-groups, and groups of guessing situations where they should follow the majority and contradict their private information, referred to as HERD-groups. In the guessing situations that compose the first three groups *unobserved* face either a favoring majority or no majority meaning that the string of guesses they observe don't conflict with their private information and guessing optimally by following the latter is arguably straightforward. On the other hand, in the guessing situations of the remaining groups *unobserved* face contrary majorities and the conflict between their private information and the string of guesses they observe implies that guessing optimally is rather challenging. Among the HERD-groups, situations where *unobserved* face short contrary majorities with medium quality signals are characterized by low incentives to follow the majority (e.g. group 7 in Experiment 1 and group 8 in Experiment 2) whereas situations where *unobserved* face large contrary majorities with low quality signals are characterized by strong incentives to follow the majority (e.g. group 5 in Experiment 1 and group 6 in Experiment 2). Among the FPI-groups in

	Group of guessing situations												
	1	2	3	4	5		6	7		8	9	10	11
Quality of signal(s)	Low	Medium	High	Low	Low		Low	Medium		Medium	Medium	High	High
Signal(s)	{b, o}	{b, o}	{b, o}	b	o	{b, o}	{b, o}	{b, o}	b	o	{b, o}	{b, o}	{b, o}
Size of the contrary majority	$\{-7, -6, \ldots, -1, 0\}$			1	1	2	$\geqslant 3$	1	$\{2, 3\}$	$\{2, 3\}$	$\geqslant 4$	$\{1, 2\}$	$\geqslant 3$
Optimal guess	FPI			FPI	HERD		HERD	FPI		HERD	HERD	FPI	FPI
Incentives	0.644	0.718	0.881	0.550	0.572		0.687	0.595		0.556	0.659	0.804	0.672

Table D2: The 11 Groups of Guessing Situations in Experiment 2

Experiment 1 (3 or 4), the situation where *unobserved* are endowed with a blue medium quality signal and face a contrary majority of size 1 entails low incentives to follow private information whereas the situations where *unobserved* are endowed with high quality signals are characterized by strong incentives to follow private information whether the contrary majority is short or long.

D.2. Non-noisy Decision Rules

We consider five non-noisy decision rules which prescribe to follow private information when facing either a favoring majority or no majority (groups 1, 2 and 3), but guess differently in at least some of the situations where facing a contrary majority.

The first decision rule is the successful observational learning rule (henceforth SOL) which guesses optimally in every group of guessing situations.

The next two rules herd excessively compared to SOL. In Experiment 1 (3 or 4), the weak conformism rule (henceforth WC) guesses like SOL except that it herds in groups 6 and 10 of guessing situations. In Experiment 2, WC guesses like SOL except that it herds in groups 4, 7 and 11 of guessing situations. In Experiment 1 (3 or 4), the strong conformism rule (henceforth SC) guesses like WC except that it also herds in group 9 of guessing situations. In Experiment 2, SC guesses like WC except that it also herds in group 10 of guessing situations.

The last two decision rules follow private information excessively compared to SOL. In Experiment 1 (3 or 4), the weak following-private-information rule (henceforth WFPI) guesses like SOL except that it follows private information in groups 4 and 7 of guessing situations. In Experiment 2, WFPI guesses like SOL except that it follows private information in groups 5 and 8 of guessing situations. In Experiment 1 (3 or 4), the strong following-private-information rule (henceforth SFPI) guesses like WFPI except that it also follows private information in groups 5 and 8 of guessing situations. In Experiment 2, SFPI guesses like WFPI except that it also follows private information in groups 6 and 9 of guessing situations.

Tables D3 and D4 report the guesses made by the non-noisy decision rules in Experiment 1 (3 or 4) and 2 respectively.

D.3. Classification Procedure and Results

Our classification proceeds in two steps. First, for each *unobserved*, we compute the proportion of her optimal guesses averaged across the first three groups of situations and the proportion of her optimal guesses averaged

	Group of guessing situations									
	1	2	3	4	5	6	7	8	9	10
SOL	FPI	FPI	FPI	HERD	HERD	FPI	HERD	HERD	FPI	FPI
WC	FPI	FPI	FPI	HERD	HERD	HERD	HERD	HERD	FPI	HERD
SC	FPI	FPI	FPI	HERD	HERD	HERD	HERD	HERD	HERD	HERD
WFPI	FPI	FPI	FPI	FPI	HERD	FPI	FPI	HERD	FPI	FPI
SFPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI

Table D3: Guesses Made by the Five Non-noisy Decision Rules in Experiments 1, 3 and 4

	Group of guessing situations										
	1	2	3	4	5	6	7	8	9	10	11
SOL	FPI	FPI	FPI	FPI	HERD	HERD	FPI	HERD	HERD	FPI	FPI
WC	FPI	FPI	FPI	HERD	HERD	HERD	HERD	HERD	HERD	FPI	HERD
SC	FPI	FPI	FPI	HERD	HERD	HERD	HERD	HERD	HERD	HERD	HERD
WFPI	FPI	FPI	FPI	FPI	FPI	HERD	FPI	FPI	HERD	FPI	FPI
SFPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI	FPI

Table D4: Guesses Made by the Five Non-noisy Decision Rules in Experiment 2

across the last seven (resp. eight) groups of situations. If both fractions are less than or equal to 50%, the *unobserved* is classified as noisy. Second, for each non-noisy *unobserved*, we compute 5 scores where each score reflects the adequacy between her guesses and the guesses made by one of the 5 non-noisy decision rules. Concretely, if in a given situation her guess matches the guess of the decision rule then the score increases by one unit, otherwise the score remains unchanged. The *unobserved* is said to be of the decision rule that has the highest score.

For each *unobserved* we use all the guesses she submitted except those of the practice part. In Experiment 4, *unobserved* with ID 4109 only submitted 48 guesses and is therefore excluded from the classification.

Ties

In Experiment 1, two subjects achieve the same score with rules SOL and WFPI. Subject 1514 is classified as WFPI whereas subject 1910 is classified as SOL. In Experiment 2, subject 2412 achieves the same score with rules SOL and WFPI, and subject 2611 achieves the same score with rules WFPI and SFPI. Subject 2412 is classified as SOL and subject 2611 is classified as WFPI. In Experiment 3, subject 3408 achieves the same score with rules WC and SC and is classified as WC. In Experiment 4, there is no tie.

Results

Table D5 shows for each experiment the percentage of *unobserved* assigned to each of the five non-noisy decision rules as well as the percentage of *unobserved* classified as noisy.

	SOL	WC	SC	WFPI	SFPI	Noisy
Experiment 1	32%	33%	03%	25%	06%	01%
Experiment 2	54%	10%	04%	19%	13%	00%
Experiment 3	38%	19%	25%	06%	10%	02%
Experiment 4	15%	32%	02%	21%	28%	02%

Table D5: Classification of *Unobserved* in the Four Experiments

Appendix E. Complements on Intuitive Observational Learning

In this appendix, we characterize the representativeness of Bayes-rational guesses for signals and we prove that it satisfies four relevant properties. We also prove that the representativeness of quantal-response equilibrium guesses for signals satisfies the two main properties and we offer numerical results suggesting that the other two properties are satisfied too. Finally, we discuss some graphical illustrations of our model of intuitive observational learning.

E.1. Characterization and Properties of Representativeness

Bayes-rational Guesses

When public guesses are Bayes-rational, the representativeness of guess $g_t \in \{B, O\}$ for signal $s_t \in \{b, o\}$, $t \in \{1, \ldots, T\}$, is given by

$$\begin{aligned} R(s_t, g_t) = Pr(g_t \mid s_t) &= \sum_{h_t \in H_t} \sigma^*(g_t \mid s_t, h_t) \Pr(h_t \mid s_t) \\ &= \sum_{h_t \in H_t} \sigma^*(g_t \mid s_t, h_t) \sum_{\mathbf{s}_{t-1}} \prod_{\tau=1}^{t-1} \sigma^*(g_\tau \mid s_\tau, h_\tau) \Pr(\mathbf{s}_{t-1} \mid s_t) \end{aligned}$$

where $\mathbf{s}_{t-1} = (s_1, s_2, \ldots, s_{t-1}) \in \{b, o\}^{t-1}$, and $R(s_1, g_1) = \sigma^*(g_1 \mid s_1)$. Note that $R(s_t, g_t) = 1 - R(s_t, \bar{g}_t)$ where $\bar{g}_t \in \{B, O\} \backslash \{g_t\}$. Furthermore, $R(s_1 = b, g_1 = B) = R(s_1 = o, g_1 = O) = 1$, and, $\forall\, t \geqslant 2$,

$$\begin{aligned} R(s_t = b, g_t = B) &= 1 - \frac{\Pr(\mathcal{B})(1 - q_{\text{PUB}}) + \Pr(\mathcal{O})\, q_{\text{PUB}}}{\Pr(\mathcal{B})\, q_{\text{PUB}} + \Pr(\mathcal{O})(1 - q_{\text{PUB}})} \sum_{i=1}^{\lfloor (t-1)/2 \rfloor} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^i \\ \text{and} \qquad R(s_t = o, g_t = O) &= 1 - \frac{1}{\Pr(\mathcal{B})(1 - q_{\text{PUB}}) + \Pr(\mathcal{O})\, q_{\text{PUB}}} \sum_{i=1}^{\lfloor t/2 \rfloor} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^i. \end{aligned}$$

where $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leqslant x\}$ and $\sum_{i=1}^{0} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^i = 0$.

Proof. First, $\sigma^*(B \mid s_t = b, h_t) = 1$ at all histories for which no O-cascade has started in period t, and $\sigma^*(B \mid s_t = b, h_t) = 0$ otherwise. Hence,

$$\begin{aligned} R(s_t = b, g_t = B) &= 1 - \Pr(O - \text{cascade in period } t \mid s_t = b) \\ &= 1 - \frac{\sum_{\theta \in \{\mathcal{B}, \mathcal{O}\}} \Pr(\theta)\ \Pr(s_t = b \mid \theta)\ \Pr(O - \text{cascade in period } t \mid \theta)}{\Pr(\mathcal{B})\, q_{\text{PUB}} + \Pr(\mathcal{O})(1 - q_{\text{PUB}})}. \end{aligned}$$

Second, an O-cascade requires two consecutive Os not canceled out by previous guesses and may start in any odd period. Thus,

$$\begin{aligned} \Pr(O - \text{cascade in period } t \mid \theta) &= \sum_{\tau \leqslant t; \tau \text{ odd}} \Pr(O - \text{cascade starts in period } \tau \mid \theta) \\ &= \sum_{\tau \leqslant t; \tau \text{ odd}} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^{(\tau-3)/2} \Pr(s_{\tau-2} = o, s_{\tau-1} = o \mid \theta). \end{aligned}$$

Similarly, $\sigma^*(O \mid s_t = o, h_t) = 1$ at all histories for which no B-cascade has started in period t, and $\sigma^*(O \mid s_t = o, h_t) = 0$ otherwise. Hence, $R(s_t = o, g_t = O) = 1 - \Pr(B - \text{cascade in period } t \mid s_t = o)$. A

B-cascade may start in any even period after one B not canceled out by previous guesses. Thus,

$$R(s_t = o, g_t = O) = 1 - \frac{\sum_\theta \Pr(\theta) \sum_{\tau \leqslant t, \tau \text{ even}} \Pr(s_t = o, s_{\tau-1} = b \mid \theta) \, [q_{\text{PUB}}(1-q_{\text{PUB}})]^{(\tau-2)/2}}{\Pr(s_t = o)}.$$

□

We now present a series of relevant properties that are satisfied by the representativeness of guesses for signals.

Property 1. For each $t \geqslant 1$, $R(s_t = b, g_t = B) > R(s_t = o, g_t = B)$.

Proof. First, we have that $R(s_1 = b, g_1 = B) = 1 > R(s_1 = o, g_1 = B) = 1 - R(s_1 = o, g_1 = O) = 0$. Second, we have that $R(s_2 = b, g_2 = B) = 1 > R(s_2 = o, g_2 = B) = \Pr(s_1 = b \mid s_2 = o)$. Third, we prove by induction that the property is satisfied for the case $t \geqslant 3$. Let us assume that $R(s_{t-1} = b, g_{t-1} = B) > R(s_{t-1} = o, g_{t-1} = B)$. We have that

$$\begin{aligned}
& R(s_{t+1} = b, g_{t+1} = B) - R(s_{t+1} = o, g_{t+1} = B) \\
&= R(s_{t+1} = b, g_{t+1} = B) - [1 - R(s_{t+1} = o, g_{t+1} = B)] \\
&= 1 - \frac{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}}{\Pr(\mathcal{B}) q_{\text{PUB}} + \Pr(\mathcal{O})(1-q_{\text{PUB}})} \sum_{i=1}^{\lfloor (t-1)/2 \rfloor} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i - \frac{\sum_{i=1}^{\lfloor t/2 \rfloor} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i}{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}} \\
&= 1 - [q_{\text{PUB}}(1-q_{\text{PUB}})] \left\{ \frac{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}}{\Pr(\mathcal{B}) q_{\text{PUB}} + \Pr(\mathcal{O})(1-q_{\text{PUB}})} \sum_{i=0}^{\lfloor (t-3)/2 \rfloor} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i \right. \\
&\qquad \left. + \frac{1}{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}} \sum_{i=0}^{\lfloor (t-2)/2 \rfloor} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i \right\} \\
&= 1 - [q_{\text{PUB}}(1-q_{\text{PUB}})] \left[1 + \frac{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}}{\Pr(\mathcal{B}) q_{\text{PUB}} + \Pr(\mathcal{O})(1-q_{\text{PUB}})} + \frac{1}{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}} \right] \\
&\quad + [q_{\text{PUB}}(1-q_{\text{PUB}})] \left[R(s_{t-1} = b, g_{t-1} = B) - R(s_{t-1} = o, g_{t-1} = B) \right]
\end{aligned}$$

where the first line is positive since $q_{\text{PUB}} > \Pr(\mathcal{B}) > 1/2$ and the second line is positive by induction. □

Property 2. For each $t \geqslant 2$,

$$R(s_t = b, g_t = B) - R(s_t = o, g_t = B) < R(s_{t-1} = b, g_{t-1} = B) - R(s_{t-1} = o, g_{t-1} = B).$$

Proof. For $t \geqslant 2$ even, $R(s_t = b, g_t = B) = R(s_{t-1} = b, g_{t-1} = B)$, and

$$R(s_t = o, g_t = B) = \frac{\sum_{i=1}^{t/2} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i}{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}} > \frac{\sum_{i=1}^{(t-2)/2} [q_{\text{PUB}}(1-q_{\text{PUB}})]^i}{\Pr(\mathcal{B})(1-q_{\text{PUB}}) + \Pr(\mathcal{O}) q_{\text{PUB}}}.$$

Similarly, for $t \geqslant 3$ odd, $R(s_t = o, g_t = B) = R(s_{t-1} = o, g_{t-1} = B)$, and

$$R(s_t = b, g_t = B) = 1 - \frac{\Pr(\mathcal{B})(1 - q_{\text{PUB}}) + \Pr(\mathcal{O})\, q_{\text{PUB}}}{\Pr(\mathcal{B})\, q_{\text{PUB}} + \Pr(\mathcal{O})(1 - q_{\text{PUB}})} \sum_{i=1}^{(t-1)/2} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^i$$
$$< 1 - \frac{\Pr(\mathcal{B})(1 - q_{\text{PUB}}) + \Pr(\mathcal{O})\, q_{\text{PUB}}}{\Pr(\mathcal{B})\, q_{\text{PUB}} + \Pr(\mathcal{O})(1 - q_{\text{PUB}})} \sum_{i=1}^{(t-3)/2} \left[q_{\text{PUB}}(1 - q_{\text{PUB}})\right]^i.$$

□

Property 3. For each $t \geqslant 1$, $R(s_t = o, g_t = O) > R(s_t = b, g_t = O)$.

Proof. For each $t \geqslant 1$, we have that $R(s_t = b, g_t = B) > R(s_t = o, g_t = B)$ (property 1). This is equivalent to $1 - R(s_t = b, g_t = B) < 1 - R(s_t = o, g_t = B)$ or $R(s_t = b, g_t = O) < R(s_t = o, g_t = O)$ for each $t \geqslant 1$. □

Property 4. For each $t \geqslant 2$, $R(s_{t-1} = o, g_{t-1} = O) - R(s_{t-1} = b, g_{t-1} = O) > R(s_t = o, g_t = O) - R(s_t = b, g_t = O)$.

Proof. For each $t \geqslant 2$, we have that $R(s_{t-1} = b, g_{t-1} = B) - R(s_{t-1} = o, g_{t-1} = B) > R(s_t = b, g_t = B) - R(s_t = o, g_t = B)$ (property 2). This is equivalent to $1 - R(s_{t-1} = b, g_{t-1} = O) - (1 - R(s_{t-1} = o, g_{t-1} = O)) > 1 - R(s_t = b, g_t = O) - (1 - R(s_t = o, g_t = O))$ which implies that for each $t \geqslant 2$

$$R(s_{t-1} = o, g_{t-1} = O) - R(s_{t-1} = b, g_{t-1} = O) > R(s_t = o, g_t = O) - R(s_t = b, g_t = O).$$

□

Quantal-Response Equilibrium Guesses

When public guesses derive from logit quantal-response equilibrium strategies, the representativeness of guess $g_t \in \{B, O\}$ for signal $s_t \in \{b, o\}$, $t \in \{1, \ldots, T\}$, is given by

$$R(s_t, g_t; \lambda_{\text{PUB}}^{E_i}) = Pr(g_t \mid s_t) = \sum_{h_t \in H_t} \sigma_{\text{PUB}}^{E_i}(B \mid s_\tau, h_\tau; \lambda_{\text{PUB}}^{E_i}) \Pr(h_t \mid s_t)$$

where $\lambda_{\text{PUB}}^{E_i}$ denotes the payoff-responsiveness and

$$\sigma_{\text{PUB}}^{E_i}(B \mid s_t, h_t; \lambda_{\text{PUB}}^{E_i}) = 1 - \sigma_{\text{PUB}}^{E_i}(O \mid s_t, h_t; \lambda_{\text{PUB}}^{E_i}) = \frac{1}{1 + \exp\left(\lambda_{\text{PUB}}^{E_i}\left(1 - 2\,\mu_{\text{PUB}}^{E_i}(\mathbf{p}, s_t, h_t; \lambda_{\text{PUB}}^{E_i})\right)\right)}$$

with $\mu_{\text{PUB}}^{E_i}(\mathbf{p}, s_t, h_t; \lambda_{\text{PUB}}^{E_i}) = \left[1 + \frac{\Pr(\mathcal{O})}{\Pr(\mathcal{B})} \frac{\Pr(s_t|\mathcal{O})}{\Pr(s_t|\mathcal{B})} \frac{\Pr(h_t|\mathcal{O})}{\Pr(h_t|\mathcal{B})}\right]^{-1}$ and $\Pr(h_t \mid s_t) = \sum_{\theta \in \{\mathcal{B}, \mathcal{O}\}} \Pr(h_t \mid \theta) \Pr(\theta \mid s_t)$ and where $\Pr(h_t \mid \theta)$ for each $t \in \{1, \ldots, T\}$ and each $\theta \in \{\mathcal{B}, \mathcal{O}\}$ is given by

$$\Pr(h_t \mid \theta) = \prod_{\tau < t} \sum_{s_\tau \in \{b, o\}} \Pr(s_\tau \mid \theta)\, \sigma_{\text{PUB}}^{E_i}(g_\tau \mid s_\tau, h_\tau; \lambda_{\text{PUB}}^{E_i}).$$

Below we prove two relevant properties that are satisfied by the representativeness of quantal-response equilibrium guesses for signals. Note that $R(s_t, g_t; \lambda_{\text{PUB}}^{E_i}) = 1 - R(s_t, \bar{g}_t; \lambda_{\text{PUB}}^{E_i})$ where $\bar{g}_t \in \{B, O\} \setminus \{g_t\}$.

Property 1. For each $t \geqslant 1$ and $\lambda_{\text{PUB}}^{E_i} > 0$ sufficiently large, $R(s_t = b, g_t = B; \lambda_{\text{PUB}}^{E_i}) > R(s_t = o, g_t = B; \lambda_{\text{PUB}}^{E_i})$.

Proof. To simplify notation, let $\lambda = \lambda^{E_i}_{\text{PUB}}$, $\sigma(g_t \mid s_t, h_t; \lambda) = \sigma^{E_i}_{\text{PUB}}(g_t \mid s_t, h_t; \lambda^{E_i}_{\text{PUB}})$, and $\mu(\mathbf{p}, s_t, h_t; \lambda) = \mu^{E_i}_{\text{PUB}}(\mathbf{p}, s_t, h_t; \lambda^{E_i}_{\text{PUB}})$. We have that $R(s_t = b, g_t = B; \lambda) > R(s_t = o, g_t = O; \lambda)$ if for each $h_t \in H_t$

$$\frac{\sigma(B \mid s_t = b, h_t; \lambda)}{\sigma(B \mid s_t = o, h_t; \lambda)} \geqslant \frac{\Pr(h_t \mid s_t = o)}{\Pr(h_t \mid s_t = b)} \tag{1}$$

with strict inequality for at least one history h_t. Note that the LHS is strictly larger than one in quantal-response equilibrium for each $t \in \{1, \ldots, T\}$ and each $h_t \in H_t$. Hence, it suffices to focus on histories such that $\Pr(h_t \mid s_t = o) \geqslant \Pr(h_t \mid s_t = b)$ or equivalently $\Pr(h_t \mid \mathcal{O}) \geqslant \Pr(h_t \mid \mathcal{B})$. The latter implies that

$$\begin{aligned} \mu(\mathbf{p}, s_t = b, h_t; \lambda) &\leqslant \frac{\Pr(\mathcal{B})\ q_{\text{PUB}}}{\Pr(\mathcal{B})\ q_{\text{PUB}} + \Pr(\mathcal{O})\ (1 - q_{\text{PUB}})} \\ \text{and} \quad \mu(\mathbf{p}, s_t = o, h_t; \lambda) &\leqslant \frac{\Pr(\mathcal{B})\ (1 - q_{\text{PUB}})}{\Pr(\mathcal{B})\ (1 - q_{\text{PUB}}) + \Pr(\mathcal{O})\ q_{\text{PUB}}} < \frac{1}{2}. \end{aligned}$$

We prove below that for these beliefs and λ sufficiently large $\sigma(B \mid s_t = b, h_t; \lambda) / \sigma(B \mid s_t = o, h_t; \lambda) \geqslant \mu(\mathbf{p}, s_t = b, h_t; \lambda) / \mu(\mathbf{p}, s_t = b, h_t; \lambda)$. Since

$$\frac{\mu(\mathbf{p}, s_t = b, h_t; \lambda)}{\mu(\mathbf{p}, s_t = b, h_t; \lambda)} = \frac{q_{\text{PUB}}}{1 - q_{\text{PUB}}} \frac{\Pr(s_t = o)}{\Pr(s_t = b)} \frac{\Pr(h_t \mid s_t = o)}{\Pr(h_t \mid s_t = b)}$$

and $q_{\text{PUB}} / (1 - q_{\text{PUB}}) > \Pr(s_t = b) / \Pr(s_t = o)$ this suffices to prove (1).

Let $G(\mu; \lambda) = f(\mu)/\mu$ where $f(\mu) = [1 + \exp(\lambda(1 - 2\mu))]^{-1}$. We have that G is increasing whenever

$$(2\,\lambda\,\mu - 1)\ e^{\lambda} - e^{2\,\lambda\,\mu} > 0.$$

First, this holds for any $\mu = 1/\lambda$ and $\mu = 1/2$ if $\lambda > 2$. Second, for sufficiently large λ, it also holds for $\mu = z/\lambda$ where $0.5 < z < 1$. Third, the LHS is increasing (decreasing) if $\mu < (>)\ 1/2$. We therefore conclude that G is increasing for every $\mu \in [\underline{\mu}, \overline{\mu}]$ where $\overline{\mu} > 1/2$ and $\underline{\mu}$ can be arbitrarily small for sufficiently large λ. Hence, $G(\mu(\mathbf{p}, s_t = b, h_t; \lambda)) > G(\mu(\mathbf{p}, s_t = o, h_t; \lambda))$ provided $\mu(\mathbf{p}, s_t = b, h_t; \lambda) < \overline{\mu}$. Indeed, if $1/2 < \mu(\mathbf{p}, s_t = b, h_t; \lambda) < q_{\text{PUB}}$ and $\mu(\mathbf{p}, s_t = b, h_t; \lambda)$ is on the decreasing part of G then $G(\mu(\mathbf{p}, s_t = b, h_t; \lambda)) > G(q_{\text{PUB}}) > G(1 - q_{\text{PUB}}) > G(\mu(\mathbf{p}, s_t = o, h_t; \lambda))$ where the second inequality holds if $1 - q_{\text{PUB}} > f(1 - q_{\text{PUB}})$ and the third inequality holds since $\mu(\mathbf{p}, s_t = b, h_t; \lambda) < q_{\text{PUB}}$ implies $\mu(\mathbf{p}, s_t = o, h_t; \lambda) < 1 - q_{\text{PUB}}$. □

Property 2. For each $t \geqslant 1$ and $\lambda^{E_i}_{\text{PUB}} > 0$ sufficiently large, $R(s_t = o, g_t = O; \lambda^{E_i}_{\text{PUB}}) > R(s_t = b, g_t = O; \lambda^{E_i}_{\text{PUB}})$.

Proof. This directly follows from Property 1 and $R(s_t, g_t; \lambda^{E_i}_{\text{PUB}}) = R(s_t, \bar{g}_t; \lambda^{E_i}_{\text{PUB}})$ where $\bar{g}_t \in \{\mathcal{B}, \mathcal{O}\} \setminus \{g_t\}$. □

Table E1 provides the representativeness of quantal-response equilibrium guesses for signals for different values of $\lambda^{E_i}_{\text{PUB}}$ and for the parametrization of the observational learning game employed in Experiments 1 and 4 where $\Pr(\mathcal{B}) = 1 - \Pr(\mathcal{O}) = 0.55$, $q_{\text{PUB}} = 14/21$, and $T = 7$ (recall that *observed* only act in periods 1 to 7). The table shows that properties 1 and 2 hold even for small values of $\lambda^{E_i}_{\text{PUB}}$. Moreover, the results suggest that $R(s_t = b, g_t = B; \lambda^{E_i}_{\text{PUB}}) - R(s_t = o, g_t = B; \lambda^{E_i}_{\text{PUB}})$ and $R(s_t = o, g_t = O; \lambda^{E_i}_{\text{PUB}}) - R(s_t = b, g_t = O; \lambda^{E_i}_{\text{PUB}})$ decrease in t, as in the case of representativeness of Bayes-rational guesses for signals.

E.2. Illustrative Predictions of Intuitive Observational Learning

To illustrate our model of intuitive observational learning, we predicted the responses to $vcPI$ for 18 different behavioral types across the guessing situations of Experiment 4 with *sitcount* $\geqslant 10$. Figure E1 shows the associated plots of the empirical value of contradicting private information against the predicted probability

Period	$\lambda^{E_i}_{\text{PUB}} = 1$		$\lambda^{E_i}_{\text{PUB}} = 2.5$		$\lambda^{E_i}_{\text{PUB}} = 5$		$\lambda^{E_i}_{\text{PUB}} = 10$	
	$R(b, B; \lambda^{E_i}_{\text{PUB}})$	$R(o, B; \lambda^{E_i}_{\text{PUB}})$	$R(b, B; \lambda^{E_i}_{\text{PUB}})$	$R(o, B; \lambda^{E_i}_{\text{PUB}})$	$R(b, B; \lambda^{E_i}_{\text{PUB}})$	$R(o, B; \lambda^{E_i}_{\text{PUB}})$	$R(b, B; \lambda^{E_i}_{\text{PUB}})$	$R(o, B; \lambda^{E_i}_{\text{PUB}})$
1	0.6033	0.4399	0.7405	0.3536	0.8906	0.2303	0.9851	0.0821
2	0.6033	0.4400	0.7372	0.3567	0.8565	0.2773	0.8995	0.3047
3	0.6032	0.4400	0.7343	0.3593	0.8323	0.3004	0.8205	0.3302
4	0.6032	0.4401	0.7317	0.3616	0.8151	0.3180	0.7966	0.3786
5	0.6031	0.4401	0.7293	0.3637	0.8015	0.3309	0.7713	0.3920
6	0.6031	0.4401	0.7272	0.3656	0.7906	0.3411	0.7584	0.4088
7	0.6030	0.4402	0.7252	0.3673	0.7816	0.3494	0.7467	0.4172

Table E1: Representativeness of QRE Guesses for Signals in the Laboratory Observational Learning Game

of contradicting. As before black (green and red) bubbles denote guessing situations with a medium (low and high) signal quality. For readability, we also plot curves from a logistic regression. All illustrations assume that $\lambda_i = 10$.

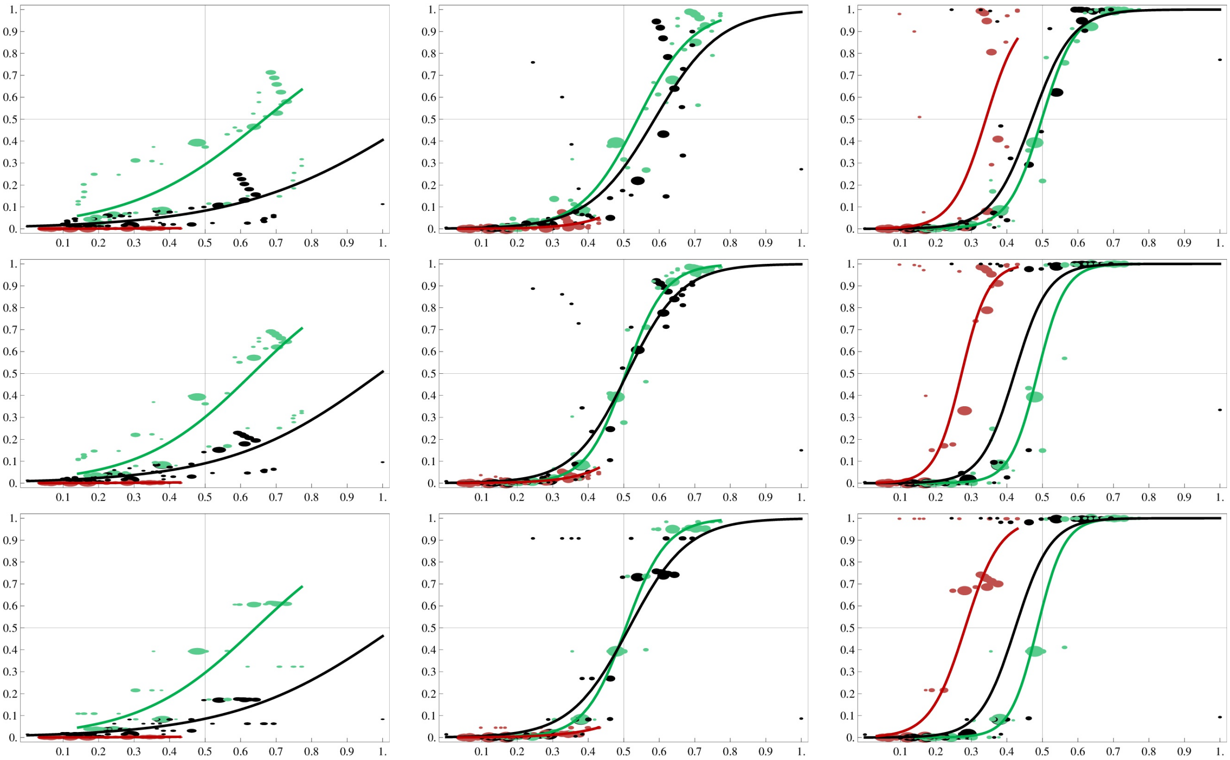
In a given panel, the public information weight $w_i \in \{0.25, 1, 2.5\}$ varies across columns increasing from left to right, and the payoff-responsiveness attributed to others $\lambda^{E_i}_{\text{PUB}} \in \{2.5, 10, 40\}$ varies across rows increasing from top to bottom. There is a null degree of local thinking ($\ell_i = 0$) in the upper panel, and an extreme degree of local thinking ($\ell_i \to \infty$) in the lower panel.

First, we find that behavioral types with $w_i = 0.25$ are always reluctant to contradict their low and medium quality signals (left column). These behavioral types contradict their low and medium quality signals more often when they are extreme local thinkers (lower panel) than in the absence of local thinking (upper panel). On the other hand, predicted responses to $vcPI$ in the left column hardly vary across the rows of a given panel. This indicates that the level of noise attributed to public guesses hardly affects the behavior of types with $w_i = 0.25$.

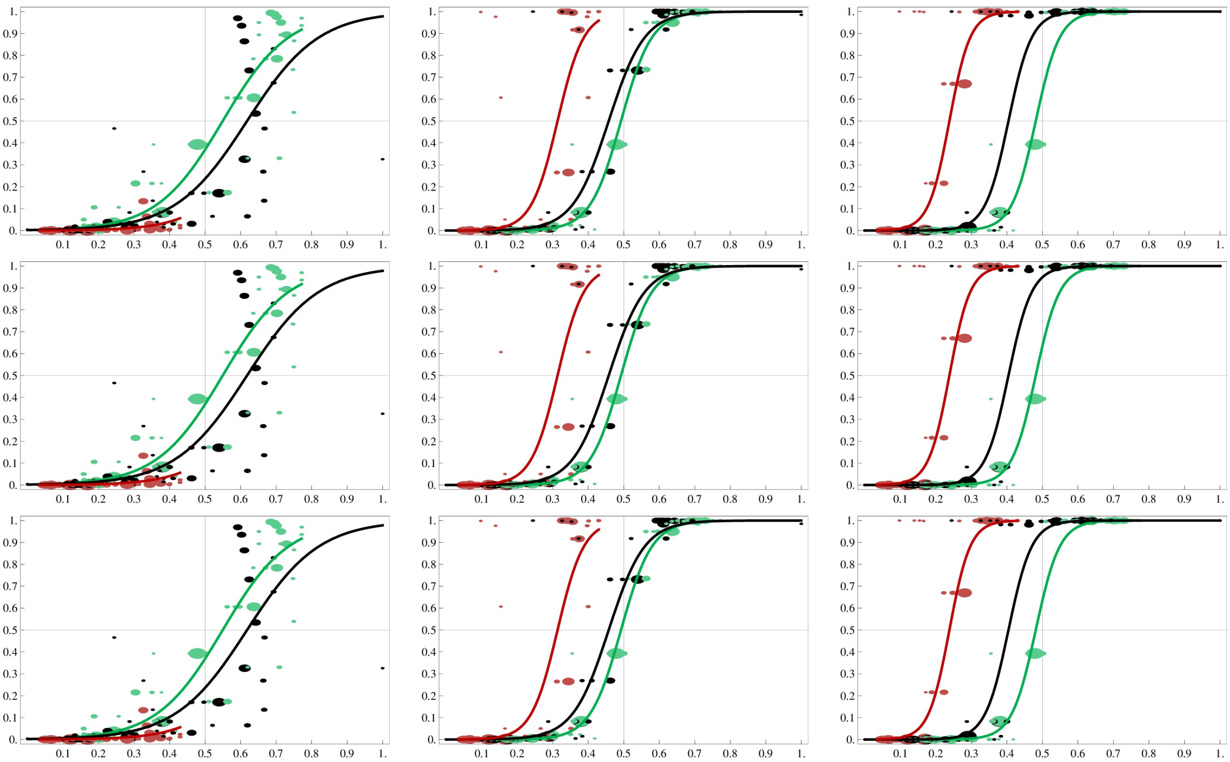
Second, reluctance to contradict low and medium quality signals is never predicted when $w_i > 0.25$, except for the behavioral type ($w_i = 1, \ell_i = 0, \lambda^{E_i}_{\text{PUB}} = 2.5$) which mildly does so. The first two observations suggest that a small public information weight best captures the reluctance to contradict low and medium quality signals.

Third, for any $\lambda^{E_i}_{\text{PUB}} \in \{2.5, 10, 40\}$, excessive herding with high quality signals results either from $w_i \geqslant 1$ and extreme local thinking or from $w_i = 2.5$ and the absence of local thinking. The difference is that behavioral types with $w_i = 2.5$ and $\ell_i = 0$ exhibit a stronger tendency to herd excessively with low and medium quality signals than extreme local thinkers who properly weight public information. Moreover, large public information weights lead to excessive herding with high quality signals even at short contrary majorities, a prediction that does not hold for large degrees of local thinking. Thus, both pronounced local thinking and large public information weights predict excessive herding with high quality signals.

Fourth, the level of noise assigned to public guesses has little impact on the predictions in the absence of local thinking, and basically no impact in case of extreme local thinking.



(a) Absence of Local Thinking ($\ell_i = 0$)



(b) Extreme Local Thinking ($\ell_i \to \infty$)

Figure E1: Illustrative Predictions of Intuitive Observational Learning

Appendix F. Estimation and Prediction Procedures

In this appendix, we first detail the estimation procedure for the four models of intuitive observational learning introduced in subsection 4.3 of the main text. Second, we report the estimates for models 1λ-QRE, 2λs-QRE, and 3λs-QR (IOL's estimates are in the main text). Third, we detail our prediction framework and we report prediction results complementary to those in the main text.

F.1. Estimation Procedure

Index subjects by $i = 1, \ldots, I$ and subject i's guessing contexts by $c = 1, \ldots, C_i$ where C_i is the number of guesses submitted by subject i in the non-practice rounds. Each guessing context c is characterized by the tuple (s_{ic}, q_{ic}, h_{ic}) where $s_{ic} \in \{b, o\}$ is subject i's private signal of quality $q_{ic} \in \{12/21, 14/21, 18/21\}$ and h_{ic} is the history of public guesses she observes, and $g_{ic} \in \{B, O\}$ denotes subject i's guess in context c. Let $\boldsymbol{Y}_i = (s_{ic}, q_{ic}, h_{ic})_{c=1}^{C_i}$ denote the collection of subject i's guessing contexts and $\boldsymbol{g}_i = (g_{ic})_{c=1}^{C_i}$ the vector of her guesses. Given her behavioral type $(\Psi_i, \lambda_i) = \left(w_i, \ell_i, \lambda_{\text{PUB}}^{E_i}, \lambda_{\text{PUB}}^{E_i^2}, \lambda_i\right)$, subject i's likelihood function is given by

$$\mathcal{L}_i\left(\boldsymbol{g}_i \mid \boldsymbol{Y}_i; \Psi_i, \lambda_i\right) = \prod_{c=1}^{C_i} \sigma_i\left(g_{ic} \mid \mu_i\left(\boldsymbol{p}, s_{ic}, q_{ic}, h_{ic}; \Psi_i\right); \lambda_i\right) \tag{F1}$$

where $\sigma_i\left(g_{ic} \mid \mu_i\left(\boldsymbol{p}, s_{ic}, q_{ic}, h_{ic}; \Psi_i\right); \lambda_i\right)$ is her quantal response with $\mu_i\left(\boldsymbol{p}, s_{ic}, q_{ic}, h_{ic}; \Psi_i\right)$ her belief.

We estimate the model parameters at the individual level to avoid making restrictive assumptions about the joint distribution of these parameters, i.e., we maximize (F1) for each subject $i = 1, \ldots, I$. And to mitigate empirical identification problems, we employ a step-wise estimation procedure for models IOL, 2λs-QRE, and 3λs-QR.

In the case of models IOL and 2λs-QRE, we repeatedly estimate the parameters (w_i, ℓ_i, λ_i), for each subject i, while holding the ratio $\lambda_{\text{PUB}}^{E_i}/\lambda_i$ fixed at each value in the grid $\Xi = \{0.1, 0.2, \ldots, 0.8, 0.9, 1\} \cup \{1/0.9, 1/0.8, \ldots, 5, 10\}$. Concretely, for each value in Ξ we maximize (F1) with respect to (w_i, ℓ_i, λ_i) using the Newton-Raphson-algorithm and we repeat the procedure at least 10 times with random starting values to rule out local maxima.[2] We then select the ratio that maximizes the log-likelihood across all 19 estimation runs. For IOL, we only keep parameter estimates with $\hat{\lambda}_{\text{PUB}}^{E_i} \neq \hat{\lambda}_i$ if they significantly improve the fit of the model over the parameter estimates with $\hat{\lambda}_{\text{PUB}}^{E_i} = \hat{\lambda}_i$.

In the case of model 3λs-QR, we adapt the estimation procedure. For each subject i, we repeatedly estimate (w_i, ℓ_i, λ_i) while holding the pair of ratios $\lambda_{\text{PUB}}^{E_i}/\lambda_i$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i}$ fixed at each point in the grid Ξ^2. Hence, we perform (at least) 361 estimation runs for each subject i. We then select the parameter estimates from the estimation run which achieves the highest log-likelihood ratio across these 361 runs. Furthermore, we perform additional estimation runs if the log-likelihood ratio is maximized on the boundaries of the grid where either (i) $\lambda_{\text{PUB}}^{E_i}/\lambda_i > 1$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i} = 0.1$, or (ii) $\lambda_{\text{PUB}}^{E_i}/\lambda_i < 1$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i} = 10$. In the first (second) case we additionally estimate (w_i, ℓ_i, λ_i) for ratios $\lambda_{\text{PUB}}^{E_i}/\lambda_i > 1$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i} < 0.1$ ($\lambda_{\text{PUB}}^{E_i}/\lambda_i < 1$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i} > 10$) with $\lambda_{\text{PUB}}^{E_i}/\lambda_i \in \Xi$ and $\left[\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i}\right] / \left[\lambda_{\text{PUB}}^{E_i}/\lambda_i\right] \in \Xi$.

Standard Errors

We use the bootstrap method of Efron and Tibshirani (1993) to obtain standard errors for our parameter estimates except $\lambda_{\text{PUB}}^{E_i}$ and $\lambda_{\text{PUB}}^{E_i^2}$. We bootstrap parameter by parameter, fixing the other parameters at their estimated values. In each run of the bootstrap for a given subject, we draw 240 times with replacement from

[2] We use the Broyden-Fletcher-Goldfarb-Shannon algorithm in case of convergence problems.

the set of decisions and we attempt 500 replications for each subject and each parameter (resp. 1,000 for IOL). Due to convergence issues not all replications may be used. Each standard error is based on at least 200 replications (resp. 400 for IOL). Note also that, contrary to the procedure for the estimates, we did not repeat the estimation in each bootstrap run which implies that the estimation procedure may have converged to local maxima in some of the runs. Thus, bootstrapped standard errors are likely to overestimate the true standard deviation of the parameter estimates.

F.2. Complementary Estimation Results

Tables F1, F2, and F3 report the estimation results for models 1λ-QRE, 2λs-QRE, and 3λs-QR respectively. For each *unobserved* in Experiment 4 (except subject 4109), each table contains (i) the estimates and bootstrapped standard errors of w_i, ℓ_i, and λ_i, (ii) the grid values of $\lambda_{\text{PUB}}^{E_i}/\lambda_i$ and $\lambda_{\text{PUB}}^{E_i^2}/\lambda_{\text{PUB}}^{E_i}$ that maximize the log-likelihood ratio across all estimation runs, and (iii) the maximum log-likelihood ratio.

We make two observations. First, as the model of expectations about others' strategy becomes richer, the distribution of estimated public information weights shifts towards higher values—the first (second and third) quartile of this distribution is 0.20 (0.58 and 1.00) for 1λ-QRE, 0.24 (0.78 and 1.40) for 2λs-QRE, and 0.34 (0.74 and 1.38) for 3λs-QR. Thus, a richer model of expectations leads to fewer estimated weights of low value. This first observation confirms that the reluctance to contradict private information can be captured either through attributing to others a smaller payoff-responsiveness than one's own or through the underweighting of the signals inferred from public guesses.

Second, lower estimated degrees of local thinking are associated with richer models of expectations. While 40 subjects satisfy $\hat{\ell}_i/\left(1+\hat{\ell}_i\right) > 0.1$ and 11 subjects are almost full local thinkers ($\hat{\ell}_i/\left(1+\hat{\ell}_i\right) > 0.9$) for 1$\lambda$-QRE, these numbers decrease to 37 and 6 for 2λs-QRE and to 19 and 2 for 3λs-QR. This second observation confirms that informational overinferences from public guesses can be captured either through the belief that others systematically make informative guesses or through local thinking.

In sum, allowing for rich expectations about others' strategy can partly substitute for non-Bayesian updating and local thinking in capturing the behavior of *unobserved.*

F.3. Prediction Framework

To measure the predictive power of an estimated model of intuitive observational learning, we evaluate the accuracy of its predictions relative to the guesses made by *unobserved* in Experiment 4.

Measuring Predictive Power

The aggregate behavior of *unobserved* in Experiment 4 can be summarized with the help of matrix $X = (t_r, h_r, s_r, q_r, value\_contra\_PI_r, prop\_contra_r, sitcount_r)_{r=1}^{R}$ where each row r represents a guessing situation with period $t_r \in \{1, 2, \ldots, 8\}$, history $h_r \in H_{t_r}$, private signal $s_r \in \{b, o\}$ of quality $q_r \in \{12/21, 14/21, 18/21\}$, empirical value of contradicting private information $value\_contra\_PI_r$, fraction of guesses that contradict private information $prop\_contra_r$, and the number of occurrences of the guessing situation $sitcount_r$. Given an estimated model of intuitive observational learning M and the matrix X, our prediction exercise proceeds as follows.

First, we compute for each guessing situation r the predicted probability to contradict private information for each of the K behavioral types that comprise model M. Denote by $pred\_prop\_contra_r^k$ the predicted probability that type $k = 1, \ldots, K$ contradicts its private signal in guessing situation r. Second, we take the average across these K probabilities, $pred\_prop\_contra_r = \sum_{k=1}^{K} pred\_prop\_contra_r^k/K$. Third, we calculate a weighted sum of squared differences (SSD) between $pred\_prop\_contra_r$ and $prop\_contra_r$ where the sum

	$\hat{w}$		$\hat{\ell}/\left(1+\hat{\ell}\right)$		$\hat{\lambda}$		Expectations		
Unobserved	Est.	SE	Est.	SE	Est.	SE	$\hat{\lambda}^{E}_{PUB}/\hat{\lambda}$	$\hat{\lambda}^{E^2}_{PUB}/\hat{\lambda}^{E}_{PUB}$	*LL*
4108	0.729	(0.095)	0.588	(0.089)	5.694	(0.592)	1	1	-61.8
4110	1.046	(0.103)	0.761	(0.115)	6.681	(0.988)	1	1	-44.4
4111	0.578	(0.071)	0.521	(0.090)	13.625	(4.609)	1	1	-29.6
4112	0.446	(0.067)	0.699	(0.143)	7.662	(1.457)	1	1	-55.4
4113	0.177	(0.089)	0.000	(0.175)	5.777	(1.024)	1	1	-88.1
4114	0.321	(0.064)	0.787	(0.165)	7.289	(1.156)	1	1	-59.6
4115	0.016	(0.028)	0.606	(0.267)	7.382	(0.855)	1	1	-72.0
4208	0.372	(0.109)	0.000	(0.095)	5.732	(0.751)	1	1	-82.3
4209	2.020	(0.282)	0.317	(0.082)	4.308	(0.414)	1	1	-66.3
4210	0.870	(0.053)	0.733	(0.041)	29.655	(10.149)	1	1	-14.2
4211	0.089	(0.032)	0.584	(0.215)	9.531	(2.289)	1	1	-53.7
4212	0.052	(0.036)	0.953	(0.460)	4.427	(0.456)	1	1	-101.2
4213	2.646	(0.238)	0.122	(0.070)	6.538	(0.973)	1	1	-37.2
4214	0.041	(0.008)	0.766	(0.107)	71.592	(49.977)	1	1	-15.3
4215	0.767	(0.057)	0.539	(0.053)	9.774	(1.745)	1	1	-37.8
4308	0.346	(0.027)	0.987	(0.103)	14.810	(5.457)	1	1	-27.0
4309	1.349	(0.191)	0.433	(0.107)	7.514	(1.689)	1	1	-28.0
4310	0.067	(0.053)	0.275	(0.210)	6.102	(3.178)	1	1	-81.0
4311	0.045	(0.016)	0.993	(0.280)	14.527	(8.866)	1	1	-38.7
4312	0.242	(0.022)	0.967	(0.062)	21.097	(6.104)	1	1	-23.3
4313	1.104	(0.079)	0.990	(0.103)	46.877	(1.2E+02)	1	1	-5.6
4314	0.544	(0.073)	0.899	(0.121)	4.861	(0.565)	1	1	-67.6
4315	1.000	(0.077)	0.303	(0.048)	11.717	(11.510)	1	1	-24.3
4408	0.196	(0.068)	0.691	(0.208)	8.438	(1.544)	1	1	-54.3
4409	0.612	(0.024)	0.661	(0.018)	24.061	(5.032)	1	1	-19.2
4410	0.001	(7.303)	0.047	(0.482)	0.170	(0.212)	1	1	-166.2
4411	0.562	(0.200)	0.000	(0.118)	4.811	(0.543)	1	1	-91.8
4412	0.999	(0.115)	0.519	(0.079)	7.682	(1.353)	1	1	-50.6
4413	0.667	(0.052)	0.968	(0.080)	7.173	(0.914)	1	1	-51.2
4414	0.842	(0.092)	0.614	(0.079)	4.920	(0.583)	1	1	-70.3
4415	0.653	(0.088)	0.719	(0.092)	6.116	(0.797)	1	1	-62.6
4508	1.019	(0.062)	0.447	(0.043)	14.903	(2.210)	1	1	-23.1
4509	0.368	(0.041)	0.943	(0.097)	8.355	(1.311)	1	1	-49.1
4510	0.000	(0.035)	0.032	(0.270)	6.833	(0.941)	1	1	-73.0
4511	0.838	(0.073)	0.307	(0.059)	8.004	(1.058)	1	1	-51.7
4512	0.307	(0.034)	0.790	(0.047)	11.838	(1.460)	1	1	-38.8
4513	0.000	(0.000)	0.413	(0.000)	2.524	(0.342)	1	1	-134.9
4514	0.136	(0.024)	0.700	(0.085)	17.074	(3.232)	1	1	-31.1
4515	1.078	(0.101)	0.501	(0.087)	5.827	(0.670)	1	1	-44.7
4608	1.857	(0.162)	0.068	(0.062)	6.114	(0.633)	1	1	-43.8
4609	1.394	(0.169)	0.970	(0.122)	9.957	(26.639)	1	1	-20.7
4610	0.675	(0.034)	0.842	(0.057)	13.719	(3.844)	1	1	-27.0
4611	2.181	(3.897)	0.987	(0.053)	8.203	(5.033)	1	1	-20.4
4612	0.651	(0.254)	0.000	(0.148)	2.280	(0.336)	1	1	-137.3
4613	0.286	(0.051)	0.963	(0.196)	3.896	(0.505)	1	1	-100.7
4614	0.437	(0.108)	0.332	(0.143)	4.951	(0.508)	1	1	-84.9
4615	1.014	(0.080)	0.903	(0.079)	10.969	(3.349)	1	1	-23.8

Note: Bootstrapped standard errors in parentheses.

Table F1: Parameter Estimates for 1λ-QRE

	$\hat{w}$		$\hat{\ell}/\left(1+\hat{\ell}\right)$		$\hat{\lambda}$		Expectations		
Unobserved	Est.	SE	Est.	SE	Est.	SE	$\hat{\lambda}^{E}_{PUB}/\hat{\lambda}$	$\hat{\lambda}^{E^2}_{PUB}/\hat{\lambda}^{E}_{PUB}$	*LL*
4108	1.279	(0.159)	0.465	(0.070)	6.163	(0.552)	0.3	1	-60.1
4110	2.278	(0.204)	0.500	(0.048)	7.584	(0.577)	0.2	1	-41.2
4111	1.089	(0.104)	0.044	(0.049)	15.796	(1.942)	0.2	1	-23.6
4112	1.228	(0.173)	0.323	(0.099)	8.009	(0.856)	0.2	1	-53.7
4113	0.216	(0.084)	0.000	(0.218)	5.816	(0.911)	10	1	-87.1
4114	0.861	(0.160)	0.418	(0.117)	7.399	(0.753)	0.2	1	-59.2
4115	0.031	(0.040)	0.666	(0.309)	7.380	(0.855)	10	1	-71.9
4208	0.382	(0.116)	0.000	(0.138)	5.705	(0.732)	3.33	1	-81.9
4209	2.478	(0.354)	0.330	(0.087)	4.563	(0.437)	0.6	1	-65.2
4210	1.426	(0.098)	0.289	(0.053)	28.911	(8.863)	0.1	1	-10.7
4211	0.094	(0.035)	0.740	(0.207)	9.530	(2.281)	10	1	-53.7
4212	0.052	(0.036)	0.991	(0.375)	4.427	(0.466)	0.1	1	-101.2
4213	3.053	(0.285)	0.094	(0.112)	6.772	(0.887)	0.7	1	-36.8
4214	0.041	(0.008)	0.767	(0.117)	71.389	(30.884)	2.5	1	-15.3
4215	0.744	(0.049)	0.710	(0.045)	9.686	(2.213)	2.5	1	-37.6
4308	0.784	(0.069)	0.560	(0.047)	14.726	(1.247)	0.1	1	-26.2
4309	1.397	(22.117)	0.666	(0.084)	6.912	(2.207)	10	1	-27.5
4310	0.193	(0.118)	0.000	(0.180)	6.188	(1.413)	10	1	-80.1
4311	0.059	(0.021)	0.797	(0.179)	14.481	(5.390)	0.1	1	-38.7
4312	0.242	(0.018)	0.969	(0.032)	21.096	(5.687)	5	1	-23.3
4313	1.104	(0.077)	0.968	(0.151)	46.877	(3.5E+02)	0.1	1	-5.6
4314	0.555	(0.070)	0.915	(0.050)	4.876	(0.563)	10	1	-67.5
4315	0.977	(0.071)	0.538	(0.062)	11.337	(4.037)	10	1	-23.9
4408	0.201	(0.069)	0.717	(0.203)	8.430	(1.684)	1.43	1	-54.3
4409	0.621	(0.031)	0.706	(0.017)	22.172	(7.017)	10	1	-17.8
4410	0.305	(0.740)	0.256	(0.343)	0.185	(0.191)	10	1	-166.2
4411	0.562	(0.196)	0.000	(0.119)	4.811	(0.543)	1	1	-91.8
4412	1.115	(0.101)	0.687	(0.048)	7.596	(1.204)	10	1	-45.5
4413	1.413	(0.107)	0.592	(0.034)	7.435	(0.522)	0.2	1	-49.8
4414	0.853	(0.079)	0.799	(0.037)	5.089	(0.611)	10	1	-68.8
4415	0.742	(0.077)	0.815	(0.037)	6.310	(0.718)	10	1	-59.7
4508	0.933	(0.058)	0.641	(0.052)	15.567	(4.279)	10	1	-22.6
4509	3.001	(0.283)	0.000	(0.051)	8.394	(0.579)	0.1	1	-48.7
4510	0.000	(0.029)	0.032	(0.246)	6.833	(0.920)	1	1	-73.0
4511	0.841	(0.072)	0.335	(0.060)	7.779	(1.002)	1.25	1	-51.5
4512	0.686	(0.072)	0.266	(0.057)	13.287	(1.304)	0.2	1	-34.4
4513	0.000	(0.000)	0.413	(0.000)	2.524	(0.333)	1	1	-134.9
4514	0.145	(0.026)	0.762	(0.069)	16.996	(3.033)	10	1	-30.6
4515	1.807	(0.166)	0.365	(0.056)	7.231	(0.593)	0.3	1	-40.3
4608	2.130	(0.199)	0.000	(0.055)	6.553	(0.667)	0.7	1	-43.4
4609	2.221	(0.311)	0.775	(0.035)	10.443	(88.665)	0.1	1	-20.4
4610	1.270	(0.065)	0.338	(0.035)	14.682	(1.172)	0.2	1	-22.8
4611	2.181	(0.320)	0.992	(0.046)	8.203	(5.086)	0.1	1	-20.4
4612	0.614	(0.237)	0.000	(0.145)	2.279	(0.339)	1.11	1	-137.3
4613	0.286	(0.050)	0.994	(0.027)	3.896	(0.503)	0.1	1	-100.7
4614	0.376	(0.091)	0.681	(0.121)	5.184	(0.587)	10	1	-82.8
4615	3.146	(0.257)	0.407	(0.053)	12.769	(0.933)	0.1	1	-21.7

Note: Bootstrapped standard errors in parentheses.

Table F2: Parameter Estimates for 2λs-QRE

	$\hat{w}$		$\hat{\ell}/\left(1+\hat{\ell}\right)$		$\hat{\lambda}$		Expectations		
Unobserved	Est.	SE	Est.	SE	Est.	SE	$\hat{\lambda}^{E}_{PUB}/\hat{\lambda}$	$\hat{\lambda}^{E^2}_{PUB}/\hat{\lambda}^{E}_{PUB}$	*LL*
4108	1.160	(0.154)	0.000	(0.078)	6.346	(0.651)	0.6	0.1	-58.5
4110	1.722	(0.150)	0.240	(0.081)	7.830	(0.634)	0.4	0.1	-39.6
4111	0.671	(0.058)	0.000	(0.035)	17.862	(12.099)	0.3	0.2	-21.0
4112	0.630	(0.079)	0.106	(0.132)	8.069	(1.420)	0.6	0.1	-52.8
4113	0.180	(0.070)	0.000	(0.098)	5.871	(1.090)	5	0.2	-86.9
4114	0.358	(0.064)	0.728	(0.139)	7.295	(1.119)	10	0.09	-59.0
4115	0.063	(0.049)	0.000	(0.115)	7.412	(0.801)	10	0.06	-71.4
4208	0.383	(0.115)	0.000	(0.142)	5.700	(0.744)	3.33	1.25	-81.8
4209	2.196	(0.301)	0.000	(0.065)	4.767	(0.452)	1	0.1	-61.2
4210	0.775	(0.024)	0.000	(0.042)	64.257	(13.480)	0.3	0.1	-9.1
4211	0.113	(0.036)	0.000	(0.240)	9.662	(2.162)	10	0.04	-53.0
4212	0.052	(0.036)	0.985	(0.407)	4.427	(0.449)	1	1	-101.2
4213	4.022	(0.398)	0.057	(0.062)	6.741	(0.679)	0.5	1.67	-36.6
4214	0.101	(0.019)	0.000	(0.051)	68.131	(43.148)	0.1	5	-14.4
4215	0.766	(0.060)	0.524	(0.072)	9.533	(1.056)	10	0.07	-36.6
4308	0.920	(0.078)	0.281	(0.048)	15.533	(2.004)	0.2	50	-24.7
4309	1.284	(20.134)	0.364	(0.105)	7.940	(0.697)	10	0.08	-26.7
4310	0.194	(0.119)	0.000	(0.189)	6.185	(1.184)	10	0.4	-80.1
4311	0.094	(0.032)	0.271	(0.195)	14.457	(91.039)	0.2	0.1	-38.6
4312	0.335	(0.018)	0.498	(0.039)	21.110	(1.167)	10	0.03	-21.0
4313	1.155	(0.058)	0.000	(0.436)	50.687	(9.827)	0.5	0.1	-5.4
4314	0.579	(0.063)	0.000	(0.304)	5.000	(0.497)	10	0.02	-66.5
4315	0.977	(0.072)	0.538	(0.044)	11.337	(5.098)	10	0.3	-23.9
4408	0.270	(0.059)	0.082	(0.295)	8.672	(1.472)	10	0.05	-53.1
4409	0.664	(0.039)	0.137	(0.034)	27.040	(6.639)	1	0.1	-13.9
4410	0.486	(0.770)	0.000	(0.291)	0.207	(0.199)	10	0.1	-166.1
4411	0.501	(0.180)	0.000	(0.118)	4.799	(0.521)	1.11	0.6	-91.8
4412	0.935	(0.062)	0.000	(0.075)	8.365	(0.930)	10	0.03	-40.5
4413	1.471	(0.107)	0.188	(0.064)	7.593	(0.502)	0.4	0.1	-48.6
4414	1.546	(0.193)	0.081	(0.092)	5.063	(0.423)	0.7	0.1	-68.5
4415	0.743	(0.078)	0.815	(0.038)	6.294	(0.771)	10	0.3	-59.7
4508	0.822	(0.022)	0.095	(0.028)	26.191	(9.384)	1.67	0.06	-20.6
4509	3.025	(0.284)	0.000	(0.051)	8.383	(0.579)	0.1	1.11	-48.7
4510	0.000	(0.034)	0.032	(0.285)	6.833	(0.920)	1	1	-73.0
4511	0.784	(0.063)	0.185	(0.084)	8.378	(0.639)	10	0.06	-47.6
4512	0.498	(0.046)	0.000	(0.037)	16.008	(1.537)	0.3	0.1	-30.4
4513	0.000	(0.000)	0.000	(0.000)	2.524	(0.333)	1	1	-134.9
4514	0.145	(0.026)	0.762	(0.069)	16.996	(3.033)	10	1	-30.6
4515	1.381	(0.118)	0.000	(0.046)	7.860	(0.633)	0.6	0.1	-36.2
4608	1.618	(0.176)	0.000	(0.061)	6.647	(0.724)	0.9	0.4	-42.7
4609	2.731	(0.367)	0.712	(0.039)	10.447	(1.926)	0.1	5	-20.3
4610	1.384	(0.092)	0.057	(0.058)	16.214	(1.360)	0.2	0.1	-21.8
4611	2.181	(58.796)	0.964	(0.052)	8.203	(5.086)	1	1	-20.4
4612	0.694	(0.244)	0.000	(0.111)	2.292	(0.327)	1.67	10	-136.8
4613	5.729	(0.990)	0.000	(0.076)	3.895	(0.367)	0.1	100	-100.2
4614	0.376	(0.090)	0.683	(0.114)	5.173	(0.583)	10	0.4	-82.8
4615	1.906	(0.158)	0.346	(0.069)	12.571	(0.931)	0.2	0.1	-21.0

Note: Bootstrapped standard errors in parentheses.

Table F3: Parameter Estimates for 3λs-QR is taken across guessing situations and each guessing situation is weighted by $sitcount_r$. The SSD of model M is thus given by $SSD_M = \sum_{r=1}^{R} sitcount_r \, (prop\_contra_r - pred\_prop\_contra_r)^2$. Finally, we compute the predictive power of model M as

$$PP_M = 1 - SSD_M / SSD_B,$$

where SSD_B denotes the weighted sum of squared differences of our theoretical benchmark. The predictive power of model M is thus the reduction in the SSD compared to the theoretical benchmark.

Theoretical Benchmark

Our theoretical benchmark assumes that players know the informational value of public guesses, whether these guesses are compatible with Bayesian rationality or not, that they form beliefs using Bayes' rule, and that they make probabilistic money-maximizing guesses conditional on their beliefs. To compute SSD_B, we assume more specifically that benchmark guesses are logit quantal-responses to an extended version of $value\_contra\_PI$. Thus, we assume that benchmark player k contradicts her private information with probability

$$\sigma_k^B (B \mid s_k = o, h_t) = \sigma_k^B (O \mid s_k = b, h_t) = \left\{1 + \exp\left(\lambda_k^B \left[1 - 2\, value\_contra\_PI'\right]\right)\right\}^{-1}$$

where λ_k^B is benchmark player k's payoff-responsiveness, and $value\_contra\_PI'$ is the extended version of $value\_contra\_PI$. We have that $value\_contra\_PI'$ equals $value\_contra\_PI$ whenever the latter exists. Otherwise, $value\_contra\_PI'$ equals

$$\begin{cases} \left[1 + \frac{11\,q}{9\,(1-q)} \frac{\widehat{\Pr}(h_t|\mathcal{B})}{\widehat{\Pr}(h_t|\mathcal{O})}\right]^{-1} & \text{if } s_k = b \\ \left[1 + \frac{9\,q}{11\,(1-q)} \frac{\widehat{\Pr}(h_t|\mathcal{O})}{\widehat{\Pr}(h_t|\mathcal{B})}\right]^{-1} & \text{if } s_k = o \end{cases}$$

where q is the quality of benchmark player k's private signal, and $\widehat{\Pr}\,(h_t \mid \theta)$ is the fraction of rounds in which history h_t occurred among all rounds in which the state is $\theta \in \{\mathcal{B}, \mathcal{O}\}$. Extending the empirical value of contradicting private information allows us to use the same set of guesses to estimate a model and to measure its predictive power.

Simulated Confidence Intervals

To check whether differences in predictive power are statistically significant we perform a simulation exercise. For each model and each guessing situation $r = 1, \ldots, R$ we randomly draw a new set of guesses according to a binomial distribution where the probability to draw a single guess contradicting the private signal is given by the probability to contradict private information predicted by the model. We then calculate predictive power from the relative frequencies of new guesses contradicting the private signal in each guessing situation. Formally, the relative frequency of new guesses contradicting the private signal in guessing situation r is a random variable $F\tilde{C}PI_r / sitcount_r$ where $F\tilde{C}PI_r \sim B\,(sitcount_r, pred\_prop\_contra_r)$. We repeat the simulation process 1,000 times to construct the 90%-confidence intervals of the model's predictive power.

We note that the simulation-average predictive power (average predictive power across the 1,000 simulation runs) differs from the predictive power calculated directly from the average predicted probabilities to contradict private information. This follows from

$$\begin{aligned}
SSD'_M &= \sum_{r=1}^{R} (pc_r - \tilde{pc}_r)^2 \cdot sc_r \\
&= \sum_{r=1}^{R} \Big[(pc_r - ppc_r)^2 + 2\,(\tilde{pc}_r - ppc_r)\,(ppc_r - pc_r) + (\tilde{pc}_r - ppc_r)^2 \Big] \cdot sc_r \\
&= SSD_M + \sum_{r=1}^{R} 2\,(\tilde{pc}_r - ppc_r)\,(ppc_r - pc_r)\,sc_r + \sum_{r=1}^{R} ppc_r\,[1 - ppc_r]
\end{aligned}$$

where $pc_r = prop\_contra_r$, $ppc_r = pred\_prop\_contra_r$, $\tilde{pc}_r = \tilde{FCPI}_r / sitcount_r$, and $sc_r = sitcount_r$. While the second term on the RHS is likely to be small, the third term is positive. Accordingly, the average SSD of model M across simulation runs is likely to be larger than the SSD calculated directly from the predicted probabilities to contradict private information, and the difference depends on the predictions and thus on the model M. We therefore systematically state the simulation-average predictive power alongside the confidence intervals.

F.4. Complementary Prediction Results

We here report prediction results for Experiment 4 that are discussed in subsection 4.3 of the main text. Table F4 reports the simulated confidence-intervals and the simulation-average predictive power of our 4 models of intuitive observational learning. Table F5 shows the predictive power of IOL by signal quality, size of the contrary majority, and $value\_contra\_PI$. To ease interpretation, the table also contains the number of guesses for each subset of guessing situations and the weighted sum of squared deviations of the benchmark model (SSD_B). Note that for a given set of guessing situations X and subsets $\{X_j\}_{j=1}^{J}$ such that $X = \bigcup_{j=1}^{J} X_j$ the predictive power of IOL on set X is a weighted sum of the predictive power on the subsets X_j where the weight of each subset is given by the corresponding SSD_B. Finally, Table F6 contains the empirical, IOL predicted, and benchmark proportions of contradictions in the different guessing situations distinguished in Table F5.

Signal Quality	*IOL*	1λ-QRE	2λs-QRE	3λs-QR
Low	53.4%	52.8%	53.6%	54.3%
	(44.6%,61.4%)	(44.1%,61.0%)	(45.6%,61.6%)	(45.6%,62.5%)
Medium	76.0%	75.3%	76.3%	76.5%
	(70.5%,81.3%)	(69.6%,80.7%)	(70.7%,81.3%)	(71.1%,81.4%)
High	78.2%	78.1%	78.8%	79.4%
	(74.0%,82.3%)	(73.6%,82.3%)	(74.2%,82.9%)	(75.1%,83.4%)
All	69.4%	68.9%	69.7%	70.2%
	(65.7%,72.7%)	(65.2%,72.3%)	(66.1%,73.1%)	(66.9%,73.3%)

Table F4: Simulated Confidence Intervals of Models of Intuitive Observational Learning

Signal Quality	Majority		Range of $value\_contra\_PI$ $[0, 0.4)$	$[0.4, 0.5)$	$[0.5, 0.6)$	$[0.6, 1]$
Low	All	Pred. Power	82.4%	38.4%	61.8%	54.8%
		SSD_B	[54.1]	[42.6]	[15.8]	[38.1]
		Nb. of Guesses	(1,932)	(386)	(236)	(1,206)
	Favoring	Pred. Power	74.8%	86.1%	99.3%	94.8%
		SSD_B	[19.4]	[4.2]	[4.4]	[7.9]
		Nb. of Guesses	(1,526)	(32)	(12)	(18)
	No	Pred. Power	82.3%	35.9%	76.0%	97.1%
		SSD_B	[20.7]	[37.3]	[5.8]	[5.4]
		Nb. of Guesses	(347)	(312)	(49)	(10)
	Small Contrary	Pred. Power	73.3%	-101.2%	17.3%	27.6%
		SSD_B	[2.8]	[0.8]	[5.6]	[17.0]
		Nb. of Guesses	(34)	(34)	(175)	(435)
	Large Contrary	Pred. Power	97.8%	81.6%		44.1%
		SSD_B	[11.2]	[0.3]	–	[7.8]
		Nb. of Guesses	(25)	(8)		(743)
Medium	All	Pred. Power	89.8%	88.6%	68.4%	86.1%
		SSD_B	[60.6]	[14.6]	[13.8]	[56.9]
		Nb. of Guesses	(2,348)	(228)	(278)	(906)
	Favoring	Pred. Power	90.3%	99.3%		98.2%
		SSD_B	[13.8]	[2.5]	–	[17.5]
		Nb. of Guesses	(1,522)	(14)		(28)
	No	Pred. Power	91.2%	98.6%		100.0%
		SSD_B	[28.1]	[2.0]	–	[5.2]
		Nb. of Guesses	(661)	(42)		(8)
	Small Contrary	Pred. Power	24.6%	83.9%	68.8%	83.2%
		SSD_B	[1.8]	[10.0]	[12.3]	[21.9]
		Nb. of Guesses	(100)	(172)	(181)	(257)
	Large Contrary	Pred. Power	94.0%		65.6%	68.5%
		SSD_B	[16.9]	–	[1.4]	[12.3]
		Nb. of Guesses	(65)		(97)	(613)
High	All	Pred. Power	66.5%	36.4%		96.3%
		SSD_B	[52.7]	[10.6]	–	[103.4]
		Nb. of Guesses	(3,477)	(58)		(225)
	Favoring	Pred. Power	83.0%			99.9%
		SSD_B	[8.6]	–	–	[13.8]
		Nb. of Guesses	(1,575)			(16)
	No	Pred. Power	98.0%			99.9%
		SSD_B	[6.8]	–	–	[8.6]
		Nb. of Guesses	(693)			(10)
	Small Contrary	Pred. Power	45.8%	-136.6%		97.1%
		SSD_B	[12.1]	[0.4]	–	[38.7]
		Nb. of Guesses	(599)	(1)		(73)
	Large Contrary	Pred. Power	62.3%	42.4%		93.5%
		SSD_B	[25.1]	[10.2]	–	[42.3]
		Nb. of Guesses	(610)	(57)		(126)

Note: A small (large) contrary majority is of size 1 or 2 (3 or more).

Table F5: Predictive Power of IOL in Diverse Guessing Situations

Signal Quality	Majority	Model	Range of $value\_contra\_PI$ $[0, 0.4)$	$[0.4, 0.5)$	$[0.5, 0.6)$	$[0.6, 1]$
Low	All	Empirical	0.077	0.155	0.403	0.716
		IOL	0.110	0.387	0.491	0.788
		Benchmark	0.176	0.452	0.586	0.800
	Favoring	Empirical	0.058	0.062	0.000	0.278
		IOL	0.078	0.103	0.051	0.155
		Benchmark	0.147	0.412	0.607	0.930
	No	Empirical	0.063	0.118	0.163	0.200
		IOL	0.158	0.377	0.331	0.325
		Benchmark	0.297	0.455	0.501	0.930
	Small Contrary	Empirical	0.471	0.471	0.498	0.607
		IOL	0.552	0.666	0.565	0.739
		Benchmark	0.208	0.464	0.608	0.774
	Large Contrary	Empirical	0.880	0.625		0.797
		IOL	0.832	0.707	–	0.838
		Benchmark	0.226	0.435		0.811
Medium	All	Empirical	0.068	0.175	0.477	0.634
		IOL	0.089	0.228	0.561	0.627
		Benchmark	0.163	0.414	0.592	0.773
	Favoring	Empirical	0.035	0.000		0.143
		IOL	0.045	0.035	–	0.064
		Benchmark	0.120	0.426		0.930
	No	Empirical	0.051	0.119		0.125
		IOL	0.099	0.116	–	0.134
		Benchmark	0.250	0.338		0.930
	Small Contrary	Empirical	0.270	0.203	0.335	0.545
		IOL	0.339	0.271	0.454	0.531
		Benchmark	0.222	0.431	0.567	0.778
	Large Contrary	Empirical	0.708		0.742	0.701
		IOL	0.632	–	0.762	0.699
		Benchmark	0.202		0.639	0.763
High	All	Empirical	0.105	0.793		0.271
		IOL	0.095	0.464	–	0.196
		Benchmark	0.084	0.367	–	0.930
	Favoring	Empirical	0.019			0.000
		IOL	0.021	–	–	0.024
		Benchmark	0.084			0.930
	No	Empirical	0.017			0.000
		IOL	0.026	–	–	0.028
		Benchmark	0.116		–	0.930
	Small Contrary	Empirical	0.107	1.000		0.205
		IOL	0.082	0.085	–	0.102
		Benchmark	0.195	0.405		0.930
	Large Contrary	Empirical	0.423	0.789		0.365
		IOL	0.379	0.470	–	0.286
		Benchmark	0.256	0.367		0.930

Table F6: Empirical, IOL Predicted, and Benchmark Proportions of Contradictions

Appendix G. The Predictive Value of Heterogeneity in Belief Distortions

Our estimation results for IOL show that there is a rich diversity in the weighting of public information and in the degree of local thinking whereas most *unobserved* believe that others have the same payoff-responsiveness as them. This appendix first evaluates the predictive value of heterogeneous, rather than homogeneous, non-Bayesian updating or local thinking. Second, we measure the loss in the predictive power of IOL when all belief distortions are homogeneous.

G.1. Heterogeneous versus Homogeneous Non-Bayesian Updating or Local Thinking

To assess the predictive value of heterogeneity in non-Bayesian updating, we compare the predictive power of heterogeneous and homogeneous non-Bayesian updating in Experiment 2. In heterogeneous non-Bayesian updating, the two parameters w and λ are individual-specific, and, in homogeneous non-Bayesian updating, the public information weight is common across the 48 *unobserved* while the payoff-responsiveness remains individual-specific. Similarly, to assess the predictive value of heterogeneity in local thinking, we compare the predictive power of heterogeneous and homogeneous local thinking, combined with heterogeneous non-Bayesian updating, in Experiment 3. In heterogeneous local thinking, the three parameters w, ℓ, and λ are individual-specific, and, in homogeneous local thinking, the degree of local thinking is common across the 48 *unobserved* while the public information weight and the payoff-responsiveness are both individual-specific. Thus, in the homogeneous cases, we estimate some of the parameters jointly across *unobserved* by maximizing the joint likelihood function with respect to 49 variables in Experiment 2 (1 w and 48 λs) and with respect to 97 variables in Experiment 3 (1 ℓ, 48 ws, and 48 λs).[3]

We derive the predictive power of heterogeneous and homogeneous non-Bayesian updating (resp. local thinking combined with heterogeneous non-Bayesian updating) from the average predicted probabilities to contradict private information in the guessing situations of Experiment 2 (resp. Experiment 3) where the average is taken across behavioral types. Table G1 reports for the four cases the predictive power by the quality of private signals and averaged across qualities.[4]

In heterogeneous non-Bayesian updating, we find that the median estimate of w is close to 1 (it equals 1.13), and that the first and third quartile is 0.51 and 1.73 respectively. This rich diversity in public information weights is comparable to the one observed in Experiment 4 (the first, second and third quartile of the distribution of estimated weights in Experiment 4 is 0.23, 0.71 and 1.30 respectively.) Moreover, assuming that subjects share a common public information weight decreases the predictive power of non-Bayesian updating by approximately 20% across signal qualities and by at least 14% for each signal quality (the loss in predictive power is larger the higher the signal quality). In fact, the normative model predicts on average as well as non-Bayesian updating with a common weight and the latter also fails to capture excessive herding with high quality signals (the common public information weight is 1.091).

Figure G1 shows the responses to the true value of contradicting private information ($tvcPI$) predicted by heterogeneous and homogeneous non-Bayesian updating. The upper (lower) subfigure plots $tvcPI$ against the probability to contradict private information predicted by homogeneous (heterogeneous) non-Bayesian updating. Each subfigure relies on the entire sample of guessing situations and it superimposes fitted lines from a weighted linear regression that includes a cubic polynomial in $tvcPI$ fully interacted with indicator variables for the signal quality. The figure clearly shows that non-Bayesian updating with a common weight

[3]Comparing the predictive power of IOL with heterogeneous and homogeneous non-Bayesian updating in Experiment 4 biases the predictive value of heterogeneity in non-Bayesian updating since the estimated parameters for the two other belief distortions differ in the two versions of IOL. Similarly, a biased predictive value of heterogeneity in local thinking is obtained when comparing the predictive power of IOL with heterogeneous and homogeneous local thinking.

[4]Estimation results are available from the authors upon request.

	Non-Bayesian Updating (Exp. 2)		Local Thinking (Exp. 3)	
	Homogeneous	Heterogeneous	Homogeneous	Heterogeneous
Low Signal Quality	-6.0% [-3.7%] (-24.0%,15.0%)	8.4% [3.9%] (-15.2%,20.1%)	23.6% [21.7%] (-3.5%,42.1%)	26.9% [24.5%] (-1.9%,44.2%)
Medium Signal Quality	-8.2% [-5.5%] (-28.5%,13.8%)	11.6% [7.0%] (-14.1%,26.7%)	75.0% [67.1%] (55.5%,77.6%)	74.7% [66.8%] (54.9%,76.5%)
High Signal Quality	12.1% [7.7%] (-9.3%,23.5%)	41.6% [27.9%] (11.3%,41.5%)	92.0% [88.2%] (84.2%,91.6%)	93.3% [89.4%] (85.6%,92.6%)
All Signal Qualities	-2.0% [-1.2%] (-13.3%,9.9%)	17.9% [10.7%] (-0.8%,20.5%)	77.9% [72.2%] (67.6%,76.6%)	79.1% [73.3%] (68.9%,77.6%)

Notes: Unbracketed numbers are predictive powers based on mean predicted probabilities to contradict private information; numbers in square brackets are predictive powers based on simulated contradictions averaged across 1,000 runs; and 90%-confidence intervals exclude the 50 runs with the lowest and the 50 runs with the highest predictive power.

Table G1: Predictive Power of Non-Bayesian Updating and Local Thinking With & W/O Heterogeneity

is neither able to predict excessive herding with high quality signals nor a reluctance to contradict low or medium quality signals at short contrary majorities. In contrast, excessive herding with high quality signals is predicted by heterogeneous non-Bayesian updating, though the effect is small.

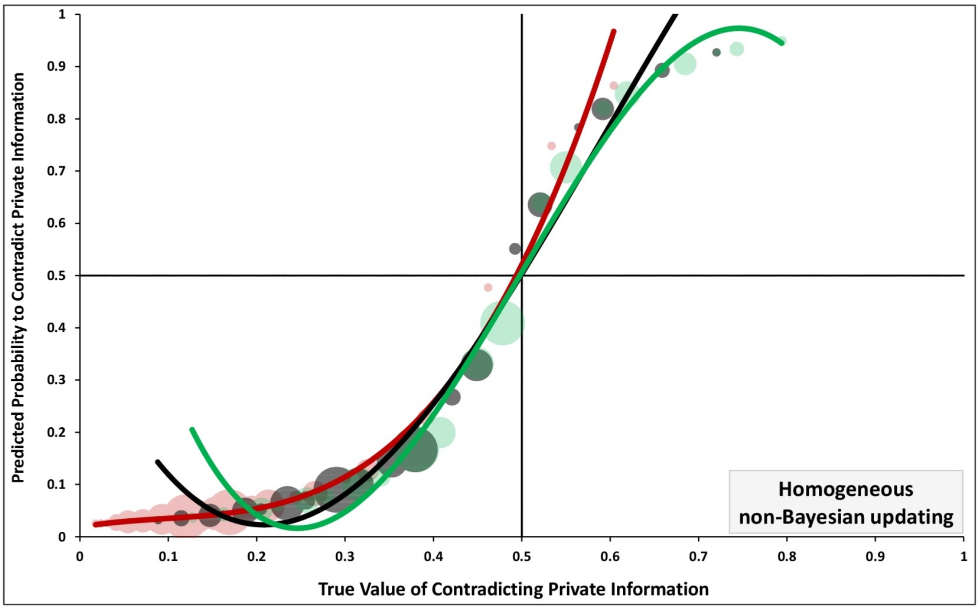
When allowing for heterogeneous local thinking, we find considerable diversity in the subjects' ability to learn successfully from Bayes-rational guesses. While 8 subjects make proper informational inferences ($\hat{\ell}/\left(1+\hat{\ell}\right) < 0.1$), 14 subjects are extreme local thinkers ($\hat{\ell}/\left(1+\hat{\ell}\right) > 0.9$), and 24 subjects are partial local thinkers.[5] Despite this rich diversity, which is comparable to the one observed in Experiment 4, we observe that homogeneous local thinking predicts as well as heterogeneous local thinking (the common degree of local thinking equals 0.572). Thus, there is a strong heterogeneity in the degree of local thinking, but this heterogeneity has little predictive value.

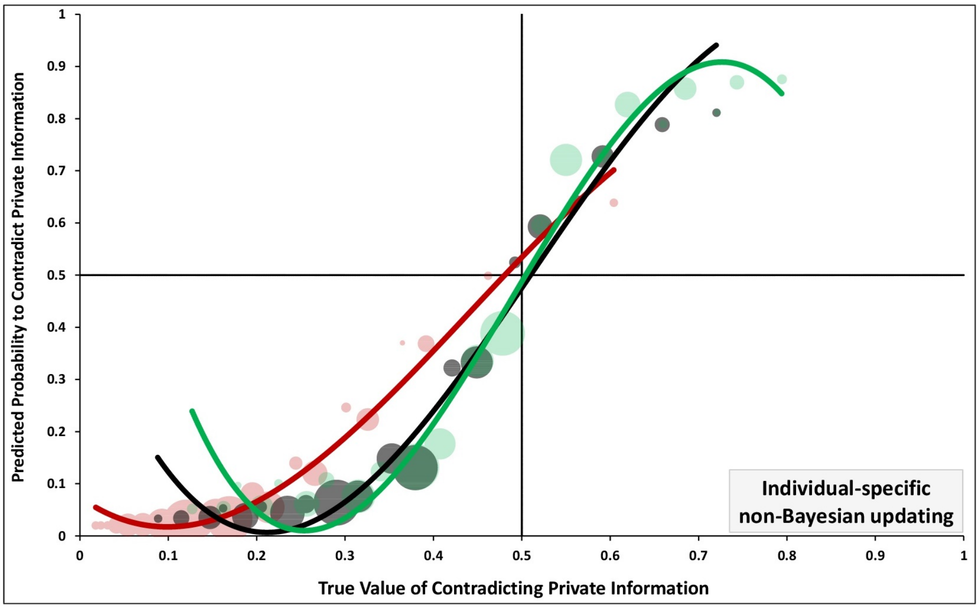
Figure G2 shows the responses to the true value of contradicting private information ($tvcPI$) predicted by heterogeneous and homogeneous local thinking. The upper (lower) subfigure plots $tvcPI$ against the probability to contradict private information predicted by homogeneous (heterogeneous) local thinking. Each subfigure relies on the entire sample of guessing situations and it superimposes fitted lines from a weighted linear regression that includes a cubic polynomial in $tvcPI$ fully interacted with indicator variables for the signal quality. The two subfigures are hardly distinguishable but for a slightly larger degree of excessive herding with a common degree of local thinking especially at very long contrary majorities of size 6 or 7. In these guessing situations the predicted probability to contradict private information with a common degree of local thinking is slightly too large.

G.2. Heterogeneous versus Homogeneous Intuitive Observational Learning

Here, we compare the predictive power of IOL in Experiment 4 to the predictive power of a variant with a common public information weight, a common degree of local thinking, and a common ratio $\lambda^{E}_{\text{PUB}}/\lambda$. We find that IOL with a single belief distortion type achieves a predictive power of 68.8%, considerably less

[5] We fail to estimate the degree of local thinking for two subjects. One subject always guesses in line with private information while the other subject is a noisy observational learner.



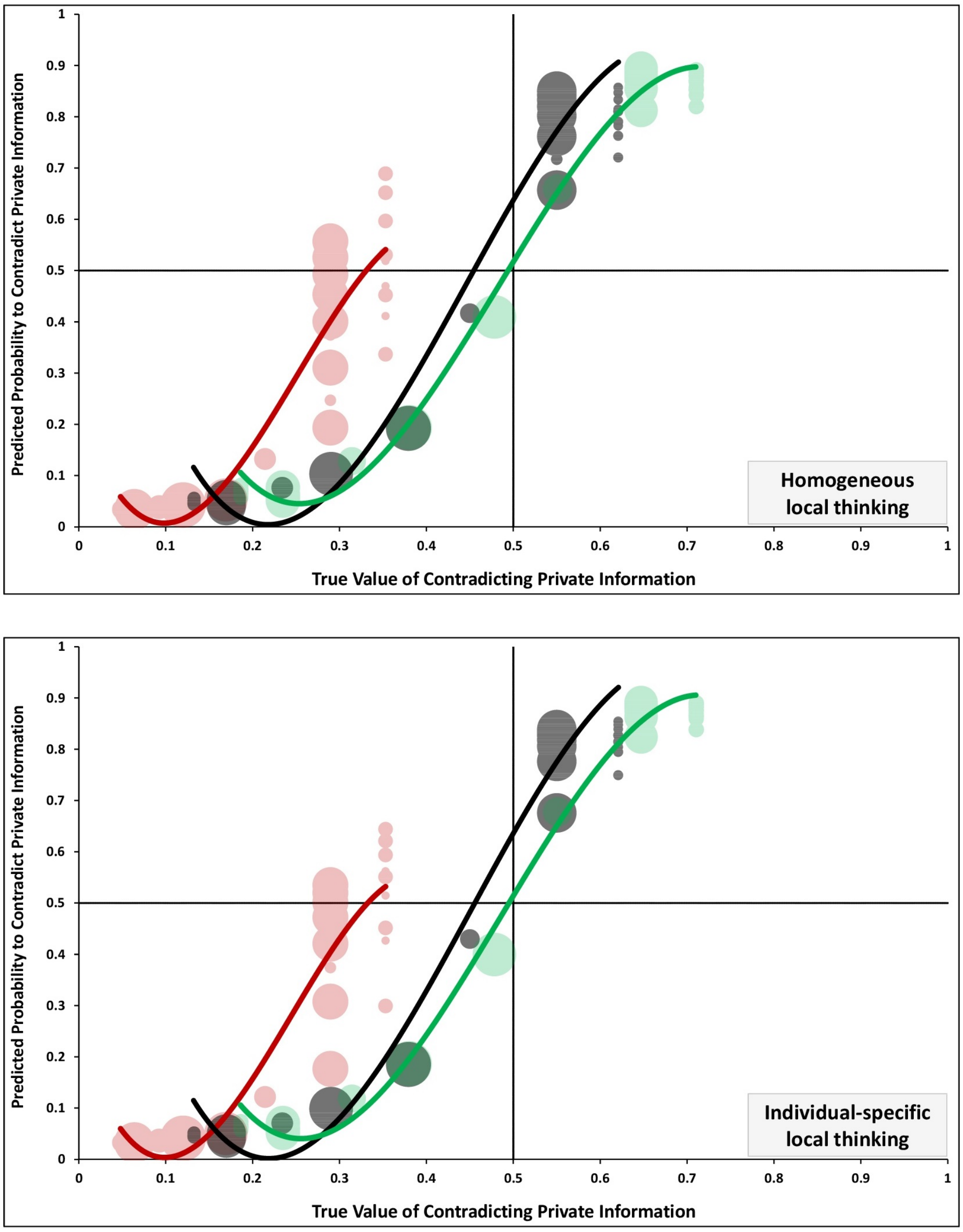


Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively.

Figure G1: Predicted Responses to the True Value of Contradicting Private Information in Experiment 2

than heterogeneous IOL (76.8%).[6] This difference is marginally significant, as the 90%-confidence interval for homogeneous IOL is $[57.9\%, 66.2\%]$ compared to $[65.7\%, 72.7\%]$ for heterogeneous IOL.

[6]The common parameter estimates are given by $\hat{w} = 5.34$, $\hat{\ell}/\left(1+\hat{\ell}\right) = 0$, and $\hat{\lambda}^E_{PUB}/\hat{\lambda} = 0.08$. The log-likelihood of the homogeneous model equals -3,143.2.



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively.

Figure G2: Predicted Responses to the True Value of Contradicting Private Information in Experiment 3

Appendix H. The Predictive Power of Restricted Versions of Intuitive Observational Learning

In this appendix, we first measure the loss in predictive power that results from excluding either non-Bayesian updating or local thinking from IOL. This allows us to investigate whether two belief distortions are sufficient to capture well the behavior of *unobserved* in Experiment 4. Then, we measure the loss in predictive power that results from excluding either non-Bayesian updating or local thinking from 1λ-QRE. Our discussion of the prediction results in subsection 4.3 of the main text has made clear that 1λ-QRE predicts (almost) as well as IOL for each signal quality. By comparing the loss in predictive power for IOL to the loss in predictive power for 1λ-QRE we assess the role played by mildly flexible expectations about others' strategy in compensating for the missing component.

We compare the predictive power of IOL (referred to as IOL-*Full* in this appendix) to the predictive power of two restricted versions of IOL, one where belief updating is Bayesian (referred to as IOL-*Bayesian*) and one where local thinking is absent (referred to as IOL-*Sophisticated*). For all three models, we first estimate the parameters for each *unobserved* in Experiment 4 (except subject 4109) and then we compare the predictive power of the three models based on the predicted probabilities to contradict private information in each guessing situation. Table H1 reports for each of the three models the predictive power by the quality of private signals and averaged across signal qualities.[7]

Signal	IOL-			1λ-QRE-		
Quality	*Full*	*Bayesian*	*Sophisticated*	*Full*	*Bayesian*	*Sophisticated*
	60.8%	59.7%	60.1%	60.2%	39.2%	47.4%
Low	[53.4%]	[52.2%]	[52.6%]	[52.8%]	[35.2%]	[40.7%]
	(44.6%,61.4%)	(43.7%,59.8%)	(43.5%,61.0%)	(44.1%,61.0%)	(24.3%,45.0%)	(30.4%,50.3%)
	86.3%	86.2%	85.4%	85.6%	61.5%	69.9%
Medium	[76.0%]	[75.7%]	[75.2%]	[75.3%]	[54.4%]	[60.6%]
	(70.5%,81.3%)	(70.0%,80.7%)	(69.4%,80.8%)	(69.6%,80.7%)	(45.9%,62.0%)	(52.1%,68.2%)
	83.1%	82.1%	83.3%	82.9%	84.5%	80.5%
High	[78.2%]	[77.3%]	[78.4%]	[78.1%]	[78.3%]	[75.3%]
	(74.0%,82.3%)	(73.0%,81.4%)	(73.9%,82.7%)	(73.6%,82.3%)	(74.0%,82.1%)	(70.2%,79.9%)
	76.8%	76.1%	76.4%	76.4%	62.6%	66.4%
All	[69.4%]	[68.6%]	[68.9%]	[68.9%]	[56.5%]	[59.2%]
	(65.7%,72.7%)	(65.1%,71.9%)	(65.3%,72.4%)	(65.2%,72.3%)	(52.1%,61.0%)	(54.8%,63.5%)

Notes: Unbracketed numbers are predictive powers based on mean predicted probabilities to contradict private information; numbers in square brackets are predictive powers based on simulated contradictions averaged across 1,000 runs; and 90%-confidence intervals exclude the 50 runs with the lowest and the 50 runs with the highest predictive power.

Table H1: Predictive Power of Restricted Versions of IOL and 1λ-QRE in Experiment 4

The results in columns 2 to 4 of Table H1 show that the restricted versions of IOL predict as well as the full model. This holds across signal qualities and regardless of the size of the contrary majority. Accordingly, herding behavior is well captured by a model of QRE expectations, where the noise level attributed to others is possibly different than one's own, combined with either non-Bayesian updating or local thinking.

As for IOL, we compare the predictive power of 1λ-QRE (referred to as 1λ-QRE-*Full* in this appendix) to the predictive power of two restricted versions of 1λ-QRE, one where belief updating is Bayesian (referred to as 1λ-QRE-*Bayesian*) and one where local thinking is absent (referred to as 1λ-QRE-*Sophisticated*). Predictive powers are reported in the last three columns of Table H1 by the quality of private signals and averaged across signal qualities. Contrary to IOL, the predictive power markedly decreases when 1λ-

[7] Estimation results are available upon request.

QRE is combined only with either non-Bayesian updating or local thinking. For example, 1λ-QRE with Bayesian updating predicts rather poorly when the signal quality is low or medium. Thus, a model of QRE expectations, where the noise level attributed to others is identical to one's own, combined with either non-Bayesian updating or local thinking fails to adequately capture herding behavior.

In sum, we find that, as alluded to in Appendix F, sufficiently rich expectations about others' strategy can substitute for non-Bayesian updating or local thinking in describing herding behavior. Still, the fact that expectation models of how others learn from public guesses are flexible enough to be descriptively accurate does not entail that they pinpoint the main principles behind herding behavior. Actually, the evidence gathered in Experiments 2-3 strongly suggests that non-Bayesian updating and local thinking are more relevant principles governing herding behavior.

Appendix I. Intuitive Observational Learning with Efficiency Concerns

In former appendices we measured the within-sample predictive power of IOL. Here, we measure the out-of-sample predictive power of (extensions of) IOL by calibrating the model from *unobserved* guesses in Experiment 4 and predicting *observed* guesses in Experiment 4 as well as *observed* and *unobserved* guesses in Experiment 1. Concretely, we rely on the model's parameters estimated for each *unobserved* in Experiment 4 (except subject 4109) to predict the mean probabilities to contradict private information in the guessing situations of *observed* in Experiment 4 as well as in the guessing situations of both *observed* and *unobserved* in Experiment 1.

All results are collected in Table I1. The table reports for each of the two experiments the predictive power of three variants of IOL by role, quality of private signals, and majority types. We discuss those results step-by-step below.

		Experiment 4			Experiment 1		
	Signal Quality	All Situations	Favoring Majority	Contrary Majority	All Situations	Favoring Majority	Contrary Majority
				IOL: *Observed*			
	Medium	64.7%	86.5%	49.4%	64.6%	89.5%	36.9%
(I)		[56.5%]	[80.8%]	[38.1%]	[58.8%]	[85.4%]	[29.8%]
		(46.9%,64.8%)	(74.5%,86.3%)	(21.0%,53.2%)	(51.1%,65.9%)	(81.1%,89.5%)	(13.1%,44.8%)
				IOL: *Unobserved*			
	Low	60.8%	66.7%	47.2%	52.6%	62.7%	24.7%
		[53.4%]	[62.3%]	[34.0%]	[45.4%]	[58.7%]	[13.0%]
		(44.6%,61.4%)	(52.7%,70.7%)	(16.9%,50.6%)	(36.0%,54.6%)	(49.0%,67.4%)	(-10.4%,33.1%)
	Medium	86.3%	93.9%	79.3%	56.1%	90.6%	27.7%
		[76.0%]	[89.0%]	[64.2%]	[48.5%]	[84.7%]	[18.2%]
(II)		(70.5%,81.3%)	(84.9%,92.0%)	(53.5%,73.9%)	(38.3%,57.2%)	(80.1%,88.8%)	(-2.8%,35.0%)
	High	83.1%	95.7%	79.3%	72.8%	93.8%	64.1%
		[78.2%]	[92.4%]	[73.6%]	[65.6%]	[89.0%]	[54.7%]
		(74.0%,82.3%)	(89.6%,94.6%)	(67.9%,79.0%)	(56.9%,73.0%)	(83.9%,92.7%)	(41.2%,66.0%)
	All	76.8%	80.8%	73.5%	58.1%	74.5%	38.7%
		[69.4%]	[76.7%]	[63.0%]	[51.0%]	[70.6%]	[28.7%]
		(65.7%,72.7%)	(72.0%,81.0%)	(57.6%,67.9%)	(45.7%,56.1%)	(65.2%,75.6%)	(17.9%,38.4%)
				IOL with Efficiency Concerns (IOL-EC): *Observed*			
	Medium	69.8%	86.2%	58.3%	76.5%	88.0%	63.8%
(III)		[60.8%]	[80.4%]	[46.0%]	[69.7%]	[84.0%]	[54.0%]
		(51.7%,68.7%)	(73.6%,85.8%)	(31.3%,59.1%)	(63.3%,75.3%)	(79.3%,88.2%)	(41.1%,64.9%)
				IOL with Expected Efficiency Concerns (IOL-ExpEC): *Unobserved*			
	Low	60.8%	66.7%	47.2%	53.5%	63.6%	25.4%
		[53.4%]	[62.3%]	[34.0%]	[46.3%]	[59.6%]	[14.0%]
		(44.6%,61.4%)	(52.7%,70.7%)	(16.9%,50.6%)	(36.6%,55.1%)	(49.7%,68.4%)	(-8.1%,34.0%)
	Medium	86.3%	93.9%	79.3%	58.4%	90.7%	31.8%
		[76.0%]	[89.0%]	[64.2%]	[50.5%]	[84.9%]	[21.7%]
(IV)		(70.5%,81.3%)	(84.9%,92.0%)	(53.5%,73.9%)	(40.6%,59.0%)	(80.1%,88.8%)	(2.8%,38.3%)
	High	83.1%	95.7%	79.3%	77.5%	93.8%	70.6%
		[78.2%]	[92.4%]	[73.6%]	[68.8%]	[88.9%]	[59.3%]
		(74.0%,82.3%)	(89.6%,94.6%)	(67.9%,79.0%)	(60.1%,76.4%)	(84.4%,92.7%)	(45.5%,70.2%)
	All	76.8%	80.8%	73.5%	60.3%	75.0%	42.7%
		[69.4%]	[76.7%]	[63.0%]	[52.8%]	[71.1%]	[31.9%]
		(65.7%,72.7%)	(72.0%,81.0%)	(57.6%,67.9%)	(46.9%,58.1%)	(65.3%,76.1%)	(20.9%,41.3%)

Notes: Unbracketed numbers are predictive powers based on mean predicted probabilities to contradict private information; numbers in square brackets are predictive powers based on simulated contradictions averaged across 1,000 runs; and 90%-confidence intervals exclude the 50 runs with the lowest and the 50 runs with the highest predictive power.

Table I1: Predictive Power of IOL and its Extensions in Experiments 1 and 4

I.1. Predicting *Observed* Guesses in Experiment 4

First, we discuss the results for *observed* in Experiment 4, which are reported in panel (I). Since the parameters are estimated from *unobserved* guesses across three signal qualities, this exercise is clearly more demanding for the model than predicting *unobserved* guesses in Experiment 4. Unsurprisingly, we find that the predictive power of IOL is lower for *observed* than for *unobserved* guesses in Experiment 4 and that the difference is statistically significant. This indicates that fitting the model across all three signal qualities limits the extent to which we can predict guessing behavior for any single quality. The predictive power for *observed* guesses is also significantly lower than the predictive power for *unobserved* guesses with medium quality signals. This suggests that the behavior of the two roles, though similar on average, is driven in part by different forces. Note also that the difference in predictive power is largest at contrary majorities.

Though the out-of-sample predictive power (about 65%) is lower than the within-sample predictive power (about 77%), IOL's out-of-sample predictions are considerably and significantly more accurate than the out-of-sample predictions of our theoretical benchmark. We now investigate the robustness of this finding by turning to IOL's predictive power in Experiment 1 whose procedures slightly differ from those of Experiment 4.

I.2. Predicting Guesses in Experiment 1

The last three columns of panel (II) report IOL's predictive powers for *unobserved* in Experiment 1. For comparison, columns 2-4 of the panel report IOL's predictive powers for *unobserved* in Experiment 4.

Remember that *unobserved* are less inclined to follow others when they should in Experiment 4 than in Experiment 1. Lower predictive powers are therefore expected in Experiment 1 than in Experiment 4 as the calibration is done in the latter experiment. Indeed, we find that IOL's predictive powers for *unobserved* are considerably smaller in Experiment 1 than in Experiment 4. This holds especially for the medium signal quality and at contrary majorities. In those situations IOL mostly underpredicts the propensity to contradict private information.[8]

Moreover, contrary to the results for Experiment 4, IOL's predictive power in Experiment 1 is higher for *observed* than for *unobserved* endowed with medium quality signals. This corroborates our finding that *observed* in Experiment 1 rely more strongly on their private information than *unobserved* in Experiment 1 since *unobserved* in Experiment 4 also show a higher propensity to follow private information than *unobserved* in Experiment 1. As argued before, this suggests that *observed* take into account the informational benefits of their guesses for their successors. Below we evaluate the increase in predictive power when considering an extension of IOL that takes into account these efficiency concerns.

I.3. Accounting for Efficiency Concerns

To capture the behavioral differences between *observed* and *unobserved*, we propose an extension of IOL that incorporates efficiency concerns. Though we do not adopt the equilibrium approach introduced in March and Ziegelmeyer (2016), our overly simple extension captures the main regularities they identified. First, altruistic players act more informatively than selfish players when the monetary incentives to follow others are sufficiently weak and they have sufficiently many successors who can benefit from the revelation of their private information. Second, future informational benefits of actions do not alter behavior if the monetary incentives to follow others are strong or if players act late in the sequence. Obviously, these two regularities cannot be captured when calibrating the model from *unobserved* guesses in Experiment 4. Moreover, belief

[8] Concretely, this holds for 74% (82%) of the guessing situations and 71% (80%) of the guesses in which *unobserved* endowed with low (medium) quality signals face a contrary majority.

distortion parameters cannot be estimated in an unbiased manner from guesses which take into account the future informational benefits of guesses since belief distortions are likely to capture in part this influence.

Therefore, we consider an extension of IOL, referred to as IOL-EC, which comprises K *Intuitive* where each *Intuitive* $k = 1, \ldots, K$ is characterized by four parameters $(w_k, \ell_k, \lambda^{E}_{\text{PUB}}, \lambda_k)$ and the vector $\boldsymbol{\alpha}_k = \big(\alpha_k^2, \alpha_k^3, \ldots, \alpha_k^T\big)$. The parameter α_k^t determines how the public information weight is distorted in period $t \geqslant 2$ due to efficiency concerns. Concretely, in period $t \geqslant 2$ *Intuitive* k assigns weight $w_k \cdot \alpha_k^t$ to the public likelihood ratio. Accordingly, the weight assigned to the public likelihood ratio by *Intuitive* k has two components: (i) a cognitive component w_k which derives from the inability to combine multiple signals in a Bayesian way, and (ii) a foresight component α_k^t which derives from *Intuitive* k taking into account future informational benefits of her guesses. To properly account for the impact of efficiency concerns we assume that $0 < \alpha_k^2 \leqslant \alpha_k^3 \leqslant \ldots \leqslant \alpha_k^T \leqslant 1$.

We merely aim at finding out whether taking into account the future informational benefits of guesses improves IOL's predictive power for *observed* in Experiments 1 and 4. We therefore make the following simplifying assumptions. First, we assume that $\boldsymbol{\alpha}_k$ is common across *Intuitive*, i.e., $\boldsymbol{\alpha}_k \equiv \boldsymbol{\alpha} = \big(\alpha^2, \ldots, \alpha^7\big)$. Second, we assume that $\alpha^t = 1$ for $t \geqslant 5$. Third, we do not estimate the parameters of IOL-EC. Instead, we rely on the set of parameter estimates for *unobserved* in Experiment 4 and we perform a coarse grid search for the vectors $\big(\alpha^2, \alpha^3, \alpha^4\big) \in \{0, 0.1, 0.2, \ldots, 1\}^3$ with $\alpha^2 \leqslant \alpha^3 \leqslant \alpha^4$ that achieve the highest predictive power for *observed* in Experiments 1 and 4 respectively.

Panel (III) reports the predictive power of IOL-EC for *observed* guesses in each experiment. The results show that allowing *observed* to take into account the future informational benefits of guesses improves considerably the predictive power of IOL. In Experiment 1, the grid search delivers the largest predictive power for $\alpha^2 = 0.2$ and $\alpha^3 = \alpha^4 = 1$, and IOL-EC's predictive power is almost 12% larger than IOL's predictive power. We also find an improvement in Experiment 4, though the grid search suggests a smaller influence of efficiency concerns ($\alpha^2 = 0.6$ and $\alpha^3 = \alpha^4 = 1$) and the difference in predictive power between IOL-EC and IOL is lower (only 5%). The results confirm that *observed* guessing behavior is partially driven by altruism and that the impact is stronger in Experiment 1 than in Experiment 4. This raises the question whether *unobserved* expect *observed* to act altruistically. We address this question next.

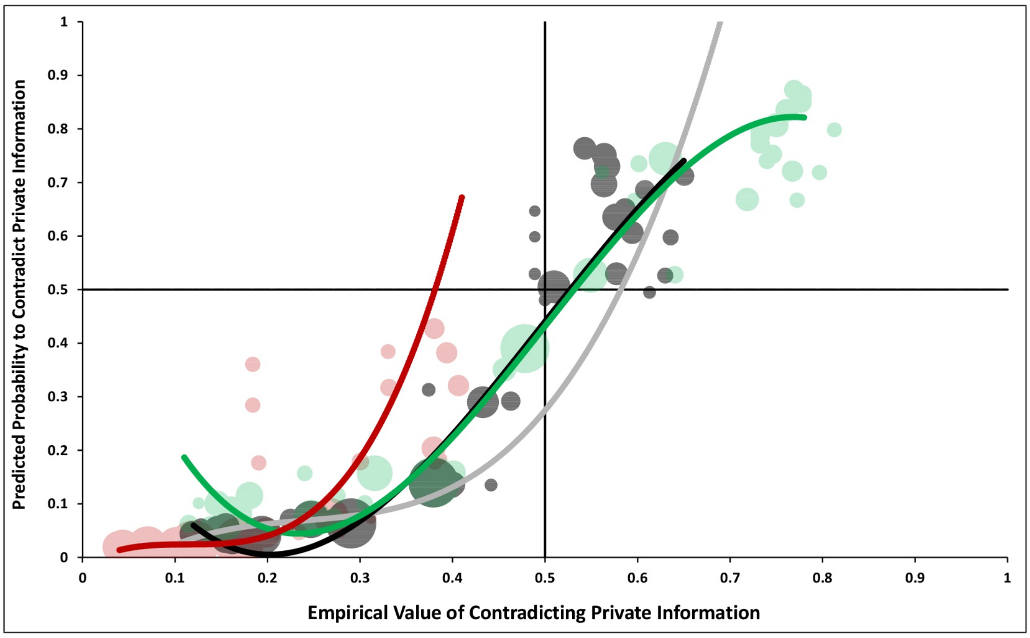
I.4. Are Efficiency Concerns Expected?

We consider a second extension of IOL, referred to as IOL-ExpEC, which comprises K *Intuitive* where each *Intuitive* $k = 1, \ldots, K$ is characterized by the four parameters $(w_k, \ell_k, \lambda^{E}_{\text{PUB}}, \lambda_k)$ and the vector $\boldsymbol{\gamma}_k = \big(\gamma_k^1, \gamma_k^2, \ldots, \gamma_k^T\big)$. The parameter γ_k^t determines how the payoff-responsiveness assigned to *observed* in period t is distorted due to the expectation that *observed* act altruistically. More precisely, *Intuitive* k expects *observed* to play a quantal-response equilibrium where the (commonly known) payoff-responsiveness in period t is given by $\lambda^{E}_{\text{PUB}} \cdot \gamma_k^t$. We assume that $\gamma_k^1 \geqslant \gamma_k^2 \geqslant \ldots \geqslant \gamma_k^T \geqslant 1$, in line with the fact that altruistic observational learners make guesses that are more informative in early than in late periods. Additionally, we assume that (i) $\boldsymbol{\gamma}_k$ is common across *Intuitive*, i.e., $\boldsymbol{\gamma}_k \equiv \boldsymbol{\gamma} = \big(\gamma^1, \gamma^2, \ldots, \gamma^8\big)$, (ii) $\gamma^t = 1$ for $t \geqslant 5$, and (iii) $\gamma^3 = \gamma^4$. Furthermore, we rely on the set of parameter estimates for *unobserved* in Experiment 4 and we perform a coarse grid search for the vectors $\big(\gamma^1, \gamma^2, \gamma^3, \gamma^4\big) \in \{1, 1.5, 2, \ldots, 5\}^4$ with $\gamma^1 \geqslant \gamma^2 \geqslant \gamma^3 = \gamma^4$ that induce the highest predictive power for *unobserved* in Experiments 1 and 4 respectively.

Panel (IV) reports the predictive power of IOL-ExpEC for *unobserved* in each experiment. We find that IOL-ExpEC achieves a slightly larger predictive power than IOL for *unobserved* in Experiment 1. The largest predictive power is obtained for $\gamma^t = 2$ for $t = 1, \ldots, 4$ and the increase in IOL's predictive power is about 2%. The difference is larger at contrary than at favoring majorities and largest for high quality signals. Indeed, across guessing situations in which *unobserved* are endowed with high quality signals the (out-of-sample) predictive power of IOL-ExpEC for *unobserved* in Experiment 1 is not significantly smaller than the (within-sample) predictive power of either IOL or IOL-ExpEC for *unobserved* in Experiment 4. We therefore conclude that *unobserved* in Experiment 1 have a partial understanding that *observed* are influenced by the future informational benefits of guesses.

In contrast, whatever the parameter constellation, IOL-ExpEC's predictive power is never larger than IOL's predictive power for *unobserved* in Experiment 4. Hence, *unobserved* in Experiment 4 fail to expect that *observed* act altruistically when learning from others. This finding may explain why *unobserved* are less prone to follow others in Experiment 4 than in Experiment 1.

Figure I1 shows the responses to *vcPI* predicted by IOL-ExpEC in Experiment 1. The figure considers only guessing situations with *sitcount* $\geqslant 10$ and superimposes fitted lines from a weighted IV regression that includes a cubic polynomial in the value of contradicting private information fully interacted with indicator variables for the signal quality and the role. The figure shows that the predicted guesses for *unobserved* with high quality signals closely match the empirical guesses. On the other hand, IOL-ExpEC fails to predict the pronounced reluctance to contradict low quality signals. Also, IOL-ExpEC predicts probabilities to contradict medium quality signals at *vcPI* $\in]0.5, 0.6]$ that are much lower than the empirical ones. Thus, procedural differences between Experiments 1 and 4 cannot be fully accounted for by extending intuitive observational learning.



Note: ●, ●, ●: *Unobserved* guesses made with high, medium and low quality signals respectively.
●: *Observed* guesses.

Figure I1: IOL-ExpEC Predicted Responses to *vcPI* in Experiment 1

References

CAMERON, A. C., AND D. L. MILLER (2010): "Robust Inference with Clustered Data," in *Handbook of Empirical Economics and Finance*, ed. by A. Ullah, and D. E. Giles, pp. 1–28. Chapman & Hall/CRC.

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics*. New York: Cambridge University Press.

EFRON, B., AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap*. New York: Chapman & Hall.

MARCH, C., AND A. ZIEGELMEYER (2016): "Altruistic Observational Learning," CESifo Working Paper Series 5792, CESifo Group Munich.

WEIZSÄCKER, G. (2010): "Do We Follow Others when We Should? A Simple Test of Rational Expectations," *American Economic Review*, 100, 2340–60.

ZIEGELMEYER, A., C. MARCH, AND S. KRUEGEL (2013): ""Do We Follow Others when We Should? A Simple Test of Rational Expectations": Comment," *American Economic Review*, 103, 2633–42.