

Why Are Rational Expectations Violated in Social Interactions?

Mohsen Foroughifar*

June 12, 2021

Abstract

Individuals often interact with each other through observation — they observe the choices of other people who possess private information. In empirical models of social interactions, it is typically assumed that decision makers are Bayesian rational, therefore their beliefs are identified by the empirical distribution of other agents' decisions. In this study, I assess the validity of the rational expectations assumption in a social interaction experiment. I use a simple and transparent experimental setting to show that decision makers often fail to exhibit rational expectations in social interactions and this behavior is independent of commonly documented errors in statistical reasoning: subjects exhibit a higher level of irrationality in the presence than in the absence of social interaction, even when they receive informationally equivalent signals across the two conditions. A series of treatments aimed at identifying mechanisms suggests that a decision maker is often uncertain about the behavior of other people. So, she might have difficulty in inferring the information contained in others' choices. A general decision-making process is then non-parametrically estimated and sheds light on the identification of three types of error in subjects' behavior.

Keywords— Rational expectation, Social interaction, Decision-making error

JEL Codes: C91, C92, D01, D81, D83

*Rotman School of Management, University of Toronto, ON, Canada, mohsen.foroughifar@rotman.utoronto.ca. This research has been approved by the University of Toronto Research Ethics Board under Human Participant Ethics Protocol #36847. Financial Support from the BEAR center (Behavioural Economics in Action at Rotman) is gratefully acknowledged. I am grateful to Ryan Webb for supervising this project. I would like to thank Yoram Halevy, Tanjim Hossain, David Soberman, Sandro Ambuehl, Guillaume Fréchette, Ben Greiner, Deborah Small, and participants of the Marketing Brown Bag at Rotman, BEAR Summer Research Retreat, and Marketing Science Conference for helpful feedback. Johannes Hoelzemann and Pavandeep Singh kindly helped with the logistics of the experiment.

1 Introduction

Understanding the interactions of economic agents who know that other agents possess private information is a central concern in the economics of information ([Manski, 2000](#)). Observation of chosen actions by other agents may reveal private information: product choices may reveal consumers' private knowledge about product quality; acceptance of job offers may reveal workers' skills ([Akerlof, 1970](#)); in the rookie market for professors, later schools may give job interviews and offers based on the known interest of earlier schools ([Bikhchandani et al., 1992](#)); when the referee of a paper finds out that a paper is rejected in another journal, his/her opinion may be tilted toward rejection. There are also numerous examples in the marketing literature which include interactions between consumers ([Godes et al., 2005](#)) or between firms ([Draganska et al., 2008](#); [Hartmann et al., 2008](#)).

In theory, a decision maker's action may impact the actions of other decision makers through constraints, preferences, and expectations ([Schelling, 1971](#); [Manski, 2000](#)). Econometric analysis of choice data often reduces the empirical inference to revelation of preferences by assuming that individuals have objectively correct expectations, namely rational expectations ([Manski, 2004](#)). In applications of games of incomplete information using field data, this assumption identifies players' beliefs from the actual probability distribution of the other players' actions ([Draganska et al., 2008](#); [Aguirregabiria & Xie, 2021](#)). However, decision makers often do not directly observe the expectations of other decision makers, so they may not have perfect forecast about other decision makers' behavior ([Manski, 2004](#); [Orhun, 2012](#)).

There are many lab and field studies that show individuals may not adequately use the information contained in other people's actions ([Eyster, 2019](#)). For example, [Stone and Zafar \(2014\)](#) use data from Associated Press (AP) Top 25 college American football poll and present evidence that sports journalists who rank American football teams for the AP top 25, under-value subjective ranking of others. So, identifying the sources of this phenomenon is important for our understanding of social interactions and for counterfactual predictions.

Studies that impose the assumption of rational expectations often do not explain why decision makers may have such optimal forecasts. Conditions that may or may not lead to the failure of this

assumption has been the focus of a limited literature in economics (Cyert & DeGroot, 1974; Smith, 1991; Kalai & Lehrer, 1993). Yet there is no consensus on why individuals may fail to exhibit rational expectations in social interactions. This is specially important in that it can lead to a socially inefficient outcome in situations where information is transmitted by observation (Eyster, 2019).

In this paper, I employ a set of relatively simple and transparent laboratory experiments to uncover why the rational expectations assumption might be violated in social interactions (SI). In the experiment, subjects need to guess about an *ex ante* unknown state of the world and are paid for accuracy. The state is binary and its possible realizations are represented by two boxes that contain ten balls of black and white color. The combination of black and white balls can be different across the two boxes and the content of each box is known to subjects. The true state is randomly realized in the beginning of the experiment, i.e., one box is selected by flipping a fair coin. Subjects do not observe the true state, but they receive a signal about it. They then guess which box is the true state. The key manipulation of the experimental design is the source of the signals: across conditions, subjects receive informationally-equivalent signals that vary only in whether or not the signal arises from a social interaction. In the *Individual* condition (control), a subject observes a ball randomly drawn from the true state. In the *Social* condition (treatment), the subject does not directly observe a ball, but she observes the choice of another participant, called *neighbor* in the experiment, who has observed a ball randomly drawn from the true state. Subjects know the precise signal-generating process and that all participants are incentivized to make a correct choice. So, the provided signal is informationally equivalent across the two conditions: under rational expectations, the choices of the subjects should be identical in the two conditions.

In order to identify individual errors that are associated with the failure of rational expectations, I use a within-subject design. I compare subjects' choices across the individual and the social conditions and present clean evidence that, despite extensive instructions, subjects exhibit on average a higher level of irrationality in the social condition (in the presence of SI) than in the individual condition (in the absence of SI) when they receive informationally equivalent signals across the two conditions. That is, they neglect the provided information relatively more in the social condition

than in the individual condition.¹

This finding extends the results of a long literature in psychology and economics, that examines why individuals may deviate from the neoclassical theory of probabilistic decision-making in isolated environments (Phillips & Edwards, 1966; Kahneman & Tversky, 1972; Ambuehl & Li, 2018), to a social interaction setting. The within subject design plays a critical role here, because it allows the identification of the errors that are associated with SI (i.e. the rational expectations assumption) while controlling for any other error that is independent of SI, e.g., errors in statistical reasoning (Benjamin, 2019).² In my setting, SI has a particularly simple form so that there is little concern regarding the complexity of the decision problem, which has previously been shown to affect individuals’ mistakes (Charness & Levin, 2009; Enke & Zimmermann, 2019).

A plausible explanation for the unexpected tendency to neglect information in the SI is the subject’s inability in inferring the relationship between what other participants choose and what they know. That is, subjects might not be able to predict how their neighbors make decision based on their private information. This uncertainty may lead subjects to mistrust their neighbors’ choices or assign a higher error rate to the neighbors’ choices than what neighbors actually exhibit (Huck & Weizsäcker, 2002; Weizsäcker, 2003; Kübler & Weizsäcker, 2004). If this is the reason behind the extra neglect in the SI, then providing additional information about neighbors’ behavior may help subjects to better extract the information contained in their neighbors’ choices.

To test the role of uncertainty about neighbor’s behavior in the subject’s irrationality, I develop three sets of treatment variations. In each treatment, I exogenously manipulate subjects’ knowledge about their neighbors by providing additional information about the neighbors. First, I present to the subjects the demographic information of their neighbors. Specifically, subjects in this treatment observe their neighbors’ age, gender, years of education, and whether the neighbors have taken Probability/Statistics courses. If demographics provide information about the behavior

¹Throughout, I use “irrationality” and “neglect” interchangeably.

²A notable difference between my work and the belief updating literature is that I examine individuals’ sub-optimal “choice” under uncertainty, while this literature examines how individuals form their posterior “belief”. The papers in this literature either do not collect data on actual choices or ignore the fact that having a Bayesian (non-Bayesian) belief is not necessarily equivalent to making a correct (incorrect) choice. As a result, the notion of “bias” in the belief updating literature is completely different than the notion of “error” in my study. A *biased belief* is defined as a belief that is not Bayesian. However, an *error in choice* is defined as an action that fails to optimize the individual’s payoff based on the available information. It can be shown that neither does a biased belief necessarily lead to an error in the choice, nor is an error in the choice necessarily a result of a bias in the belief.

of neighbors, one expects to see a lower level of neglect in the SI for this treatment than the base experiment. The results indicate that providing demographic information about the neighbor has a small and insignificant effect on the subjects’ irrationality in the SI.³

Second, I design a treatment in which the neighbor is replaced by a “computer bot”, whose behavior is clearly described to the subjects. The idea here is to create a social environment where there is less uncertainty in the neighbor’s behavior than the baseline experiment. In this treatment, subjects are told that their neighbor is a computer bot that chooses the box with more black balls when it observes a black ball, and chooses a box with more white balls when it observes a white ball. The results of this low uncertainty treatment show that the neglect in the social condition significantly drops compared to the baseline experiment. This finding highlights that the uncertainty of neighbor’s behavior can play an important role in individual decision-making when social interaction happens through observation. In fact, the violation of rational expectations might be a result of the ambiguity in other decision makers’ behavior.

Third, I devise a treatment where the subject observes both her neighbor’s choice and the ball that her neighbor has seen. If the additional neglect in the SI was largely about the uncertainty in the neighbor’s behavior, then the observed difference between the irrationality across the social and the individual condition should disappear in this treatment. The results support this prediction: when subjects are provided with both their neighbors’ choices and the signals behind those choices, there is no statistically significant difference between the level of irrationality (neglect) across the social and the individual condition. This suggests that the failure of rational expectations in the social condition is mainly driven by the ambiguity of other people’s behavior, i.e., subjects behave as if they lack knowledge about how others make choice based on their private information. As a consequence, even though subjects know that others’ choices are based on useful information, they may choose to neglect them.

Researchers are increasingly interested in the mechanisms behind reduced-form errors in decision-making due to the view that this may help develop new behavioral models that can explain real world behavior (Enke, 2020). In the final part of the paper, I develop and estimate a non-parametric

³I also examine the role of subjects’ observable characteristics in driving the unexpected irrationality observed in the SI. I find that typical demographics such as: age, gender, education, and taking Probability/Statistics courses cannot explain the additional neglect in the social condition.

model of decision-making to explain the sources of individual error in social interactions. This model is based on the theoretical work of Walliser (1989) and uncovers the mechanisms that may lead a subject to make an irrational choice in the context of my experiment, and identifies the channel associated with the violation of the rational expectations.

In the current study, the individual’s decision-making process can be modelled as a two stage procedure: upon obtaining a signal, a subject needs to 1) update her belief, and 2) make a binary choice based on her belief. Here, an irrational choice can be a result of one of the following two errors. First, it might be the case that the subject updates her belief in a wrong direction, i.e., the updated belief contradicts the observed signal. In the context of my experiment, this happens when the signal supports a specific state, but the subject mistakenly believes that the other state is more likely, and thus she makes a choice that, from the experimenter’s point of view, is irrational. I call this mistake a *posterior error*.⁴

Second, the subject may update her belief in a correct direction, but she may fail to pick the right choice based on her belief.⁵ In my experiment, this happens when both the signal and the subject’s posterior belief favor the same state, but the subject makes a choice that contradicts her belief. I call this mistake a *reasoning error*.

I first show that posterior errors and reasoning errors are not separately identified from choice data and one needs additional sources of variation to be able to distinguish between them. I then add a survey question to each choice that subjects make during the experiment. The purpose of this survey question is to measure the “relative” direction of the subject’s posterior belief.⁶ Specifically, it asks about the probability that the subject believes her choice is correct. This survey question along with the subject’s actual choice allow me to identify both the posterior error and the reasoning error in subjects’ behavior. I find that both of these errors contribute to subjects’ mistakes in the experiment, with posterior error being the dominant one.⁷ I then compare the magnitude of these errors across the individual condition and the social condition. The result shows that the additional

⁴The posterior error is a choice that is consistent with the subject’s belief, but from the experimenter’s point of view is irrational because it fails to optimize the subject’s payoff.

⁵The idea that subjects may choose all option with positive probability has been documented in prior literature (McKelvey & Palfrey, 1995).

⁶See Manski (2004) for a detailed discussion of how survey data on probabilistic expectations can enable experimental economists to overcome identification problems.

⁷The joint identification of reasoning error and posterior error has not been explored in the prior literature.

neglect in the social condition is mainly driven by posterior error. That is, while the magnitude of the reasoning error remains unchanged across the individual and the social conditions, the posterior errors are significantly higher in the social condition. This evidence is consistent with the earlier finding that there is uncertainty in other people’s behavior and provides an explanation for why an individual may fail to exhibit rational expectations in SI.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the experimental design. Section 4 presents the main results of the paper. In Section 5, I develop a model to explain the sources of irrational choices in individual behavior and identify the channel that is influenced by SI. In section 6, I non-parametrically estimate a two-step decision making process and highlight what variation in the data identifies which error in decision-making. Finally, Section 7 concludes the paper.

2 Related Literature

Empirical models of choice data often assume decision makers are Bayesian (Erdem & Keane, 1996). In social environments and games of incomplete information, the assumption of rational expectations is further imposed for identification of beliefs (Draganska et al., 2008; Aguirregabiria & Xie, 2021). However, there are a lot of studies in the literature that document the opposite: that individuals do not always behave as a Bayesian rational agent would do.

Plausible explanations of the violation of Bayesian rational behavior include three categories of errors: **A) *belief updating error***: this error has been widely documented in isolated environments and is not necessarily related to the social environment. It can arise due to several reasons such as *conservatism* (Edwards, 1968; Huck & Weizsäcker, 2002), *overconfidence* (Nöth & Weber, 2003), *base rate fallacy* (Goeree et al., 2007), or more broadly any form of non-Bayesian updating (Grether, 1980; Jindal & Aribarg, 2021; Ching et al., 2021). **B) *reasoning error***: this type of error often happens randomly when a subject makes decision based on her updated belief, and again is not specific to the social environment. The well-known logistic choice function is a special case of this where errors are due to random shocks that follow an extreme-value type I distribution. **C) *violation of rational expectations***: this error is specifically associated with the social

environment and can be due to wrong beliefs about others’ strategies (Kübler & Weizsäcker, 2004; Weizsäcker, 2010).⁸

Despite numerous studies in the literature, it remains unclear whether and how one can non-parametrically identify these three types of errors using empirical data. Many studies either ignore some of these errors or put additional assumptions to overcome the identification challenges. For example, the need to separate the inability to best respond to one’s beliefs (type B error) from potentially wrong beliefs about others (type C error) has been recognized in prior literature (Stahl & Wilson, 1995; Camerer et al., 2004). However, most of these papers approach the identification problem by imposing parametric functional forms on beliefs and/or choice probabilities (Kübler & Weizsäcker, 2004; Goeree et al., 2007; March & Ziegelmeyer, 2018). These approaches might suffer from misspecification and generate misleading results (Dominitz & Hung, 2009). Some papers such as Weizsäcker (2010) utilize a reduced-form approach and do not impose parametric assumptions, but these studies are not able to precisely identify the contribution of three types of errors mentioned earlier to the subject’s behavior. Hence, they provide limited explanation about the nature of the suboptimality.

The current paper contributes to the literature by developing a social interaction experiment which non-parametrically identifies these errors in subject’s behavior. The framework in my study is built on the theoretical work of Walliser (1989) and sheds light on what variation in the data identifies which error type. The idea is that type A and type C errors often appear when a subject forms expectation about the environment (stage 1 in decision-making), but type B error emerges when these expectations are used to find out a selected action (stage 2 in decision-making). To distinguish between these two stages, I collect both choice data and belief data so that I am able to separate type B errors from the sum of type A and type B errors. This highlights the fact that one cannot separately identify errors associated with the first stage of decision-making process from those of the second stage using only the choice data or only the belief data.

To further disentangle between type A and type B errors, I separately ask subjects to process a private signal and the choice of a neighbor across two conditions with a within-subject design. The

⁸Some papers suggest a combination of these errors (March & Ziegelmeyer, 2018; Angrisani et al., 2020; Filippis et al., 2021).

separation and within-subject design are important because without them, one may not be able to study the “absolute” effect of each source of information. In fact, as [Eyster \(2019\)](#) points out, many studies in the social learning literature are only able to study the “relative” importance of private versus social information.⁹

While the above features of the experiment help to separate different types of error, they are not sufficient for the non-parametric identification of beliefs and choices. One needs enough variation in the information structures (e.g. precision of signals) so that beliefs cover the whole range of probabilities in $[0, 1]$. The wide variety of information structures in my experiment further facilitates the non-parametric estimation and provides insights about what parametric assumptions might or might not be appropriate for modeling decisions in social interactions. This variation in the data is closely related to the literature on non-parametric identification of non-equilibrium beliefs in games of incomplete information ([Aguirregabiria & Magesan, 2020](#); [Aguirregabiria & Xie, 2021](#)). The novelty is in the source of variation which has not been sufficiently explored in prior literature. In addition, I show that there is substantial heterogeneity in decision-making across subjects. So, one needs to naturally account for individual-specific heterogeneity when modeling decision-making in social interactions. The heterogeneity in beliefs has been documented in other contexts, see for instance [Orhun \(2012\)](#), [Ching et al. \(2021\)](#), [Jindal and Aribarg \(2021\)](#), and [Benetton and Compiani \(2021\)](#). My results are consistent with these findings and extends them to a more general framework in which there is no strategic incentive among subjects.¹⁰

3 Experimental Design

Subjects are randomly assigned to one of four treatments (Figure 1). The experiment in each treatment consists of two consecutive parts and the order of these two parts is randomized. In one

⁹My study has two other advantages compared to social learning experiments ([Anderson & Holt, 1997](#); [Kübler & Weizsäcker, 2004](#); [Goeree et al., 2007](#)). First, the decision task in my experiment is very simple so that there is no concern about the complexity of the decision in the social condition. The complexity of the decision problem has been shown to be an important factor in driving individual errors ([Charness & Levin, 2009](#); [Weizsäcker, 2010](#); [Enke & Zimmermann, 2019](#)). Second, by separating the people whose choices are used in the social condition from those of the main experiment, my experimental design rules out any confounding factor which might be related to the prosocial/strategic incentives among subjects and provides a lower bound on the amount of irrationality one might expect to observe in social interactions.

¹⁰For a literature review on (the failure of) rational expectations in other strategic settings see [Beard and Beil \(1994\)](#) and [Eyster \(2019\)](#).

part, the subject performs a task in isolation — without social interaction (*individual condition*). In the other part, she performs the same task with social interaction (*social condition*). As noted earlier, the order of these two parts is randomized, i.e., given a treatment, some subjects first see the individual condition and then proceed to the social condition, and others see them in reverse order (see Figure 1).

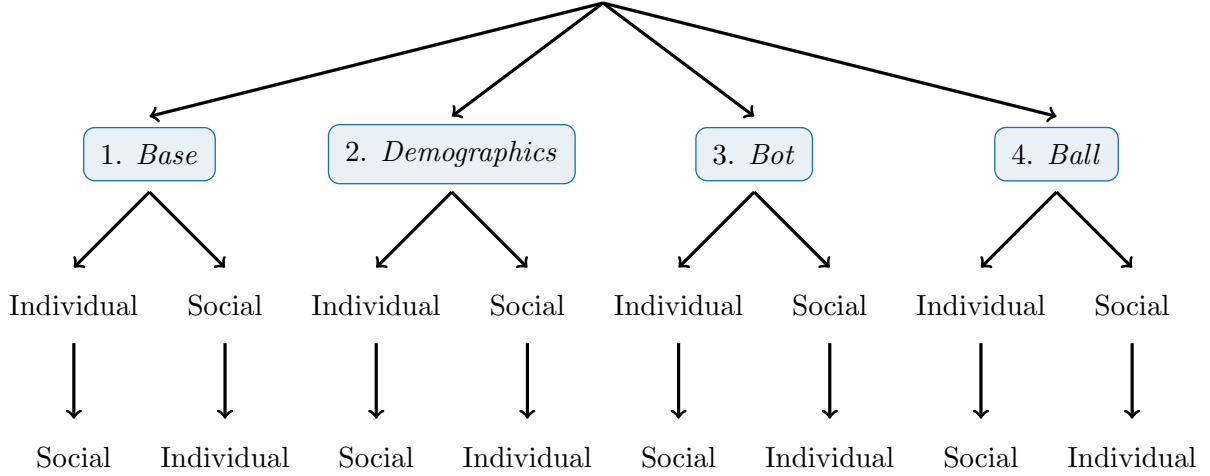


Figure 1: The experimental design

The individual condition is the same in all four treatments, but the social condition differs across treatments. The idea is to exogenously manipulate the subject’s knowledge about the participants with whom she is interacting across treatments. I will elaborate on the differences between treatments as I proceed in the following.

3.1 Individual Condition

The *individual condition* is a benchmark which measures the subjects’ behavior in the absence of social interaction. It consists of 21 rounds. In each round, two boxes are shown to the subject. Each box contains 10 balls of white or black color (see Figure 2 for an example). These boxes represent the possible states of the world, $\omega \in \{X, Y\}$. In the beginning of each round, a fair coin is anonymously flipped. If the coin is **Head**, the state is **X** and one ball is randomly drawn from

box X . If the coin is **Tail**, the state is Y and one ball is randomly drawn from box Y .¹¹ The subject does not observe the coin. She observes the ball, and then is asked to guess what the state is. The combination of white and black balls randomly changes over 21 rounds. Denote the fraction of white balls in box X by θ_X and the fraction of black balls in box Y by θ_Y . The combinations used in the experiment include a wide range of symmetric and asymmetric information structures: $\{(\theta_X, \theta_Y) \mid \theta_X, \theta_Y \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1\}, \theta_X \geq \theta_Y\}$.

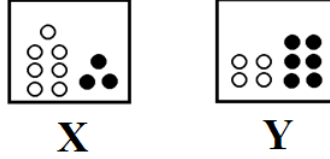


Figure 2: These two boxes represent possible states of the world (Here, for instance $\theta_X = 0.7$ and $\theta_Y = 0.6$)

Subjects are incentivized to make a correct guess:¹² it is best for a subject to pick a box with more black (white) balls when she observes a black (white) ball. In addition to collecting the subjects' choices, I add a survey question at the end of each round that asks for the subject's posterior belief. Specifically, the subject answers the following question after she reports her choice in each round: *with what probability do you think your guess is correct?*¹³ Unlike the choice, the survey question is not incentivized in the experiment for a few reasons. First, I found that the experiment lasts too long when subjects are required to go through an incentive compatible elicitation procedure for each of the posterior beliefs that they submit during the experiment. So, it may cause fatigue and contaminate the choice data that is vital for the main analysis. Second, I did use monetary incentives for posteriors in a pilot study using a revised version of Quadratic Scoring Rule (Brier, 1950).¹⁴ The pilot results suggested that the incentivized posteriors are not significantly different than the posteriors that are not incentivized. The prior literature has also shown that responses

¹¹This induces a prior probability of $\frac{1}{2}$ for each box. The language used in the actual experiment was slightly different: I used *box H* (head) and *box T* (tail) instead of box X and box Y to remind individuals about the randomization (see the appendix for experiment instructions).

¹²The idea is to randomly choose some rounds and pay the subject for each correct guess in those rounds. I explain the payment scheme in more details later.

¹³As I explain later, I am interested in knowing which state is more likely from the subject's perspective when she makes a choice. Effectively, I only need to know whether the subject chooses a state that she believes has a higher (>50%) or lower (<50%) chance of being correct.

¹⁴For the subjects who are incentivized for both the choice and the posterior, I randomly select one posterior and one choice for payment (the selection is independent).

to this type of survey questions, in the absence of incentives for honest revelation of expectations, do possess face validity when the questions concern well-defined events; see [Manski \(2004\)](#) for a detailed discussion. Forth, the main analysis in this study does not require posterior belief. So, the main findings do not rely on whether or not the survey question is incentivized. I kept the survey question very standard and easy to understand so that it is unlikely that the subject does not understand the survey question or incurs a cognitive cost to think about the answer ([Smith, 1991](#)). So, one can expect the reported posterior probabilities to be close to the subjective probabilities in the subject’s mind.¹⁵

3.2 Social Condition

The *social condition* is designed to study the subjects’ behavior in the presence of social interaction. The structure of the task is similar to that of the individual condition. The social condition consists of 21 rounds. In each round, the subject is randomly connected to another participant, called *neighbor* in the experiment, and receives information from one of the rounds in the neighbor’s individual condition. The subject observes the content of two boxes that has been shown to the neighbor. Her task is to guess what the state (selected box) is, based on the information that she obtains from the neighbor. As noted earlier, there are four treatments in the experiment and the transmitted information in the social condition is different across treatments. I explain the treatments in the following.

The first treatment is called *base*. In this treatment, the information coming through social interaction is the neighbor’s guess.¹⁶ Note that a neighbor here is a random subject who has previously participated in the experiment. The subject knows that the neighbor’s guess is incentivized and is based on a randomly drawn ball from the realized state (box). To summarize, in each round,

¹⁵I also did a robustness check at the end of my main experiment and incentivized all subjects according to Quadratic Scoring Rule. The elicited posteriors were very similar to those that were collected from the survey questions during the experiment. But I do not use these incentivized posteriors in my analysis because the incentive compatible elicitation of posteriors were always happening at the end of the experiment, after both the individual condition and the social condition had been finished. The fact that the incentivized posteriors were always collected after the end of the experiment might make the results inconclusive (the subjects were answering the same survey questions as they had observed during the experiment. The concern is that subjects might not think about the questions anymore because they had already seen the same questions before. Hence, elicitation mechanism might not have an impact on subjects’ posteriors.)

¹⁶By “guess” I mean the actual choice of the neighbor. The subject does not observe the neighbor’s posterior belief (the answer to the survey question).

the subject observes two boxes and the guess of a neighbor, but not the ball that the neighbor has observed. Then, the subject is asked to guess about the realized state. The experiment is designed such that the neighbor randomly changes in each round. Hence, the subject does not interact with the same neighbor over time and it is unlikely that the subject learns about a specific neighbor's behavior over the course of 21 rounds in the social condition.

The second treatment is called *demographics*. There is a slight difference between the social condition in this treatment and in the base treatment: on top of the neighbor's guess, the subject observes the neighbor's demographic information such as age, gender, years of education, and whether the neighbor has taken any probability/statistics courses. This treatment is designed to examine whether providing demographic information about the neighbor can alleviate the irrationality associated with the uncertainty of neighbor's behavior in the social interaction. If demographics provide additional information about the behavior of neighbor, the uncertainty might be lower in this treatment than the base treatment.

The third treatment is called *bot*. Everything in this treatment is the same as in the base treatment, except that the neighbor is a computer bot which is programmed to exhibit a specific behavior (i.e. rational). This means when the bot observes a white ball, it chooses the box with more white balls, and when it observes a black ball, it chooses the box with more black balls. The behavior of the bot is explained in details to the subjects in this treatment. Subjects see the guess of the bot in this treatment and then submit their own guesses about the realized state. The social interaction in this treatment is relatively more transparent than the earlier two treatments. So, one expects the irrationality associated with the uncertainty of neighbor's behavior to be significantly lower in this treatment than the base treatment.

The fourth treatment, which is called *ball*, is an augmented version of the base treatment in which the subject observes both the (human) neighbor's guess and the ball which has been shown to the neighbor. The uncertainty effect is expected to completely disappear in this treatment because the subject is provided with all the relevant information regarding her neighbor's choice.

3.3 Payment Scheme

Each subject receives \$6 show-up fee for participation. In addition to that, two rounds of the experiment are randomly selected and the subject wins \$12 for each correct guess in those two rounds.

4 Results

In this section, I first define the criteria for recognizing individual errors in the context of my experiment. I then analyse subjects' choices in the experiment to measure the frequency of these errors and examine the relation between them in the individual condition and in the social condition. The comparison between errors in the individual and in the social condition identifies the errors that are associated with the social interaction (e.g. the violation of rational expectations).

In the individual condition, a Bayesian rational subject should choose a box with more black balls when she observes a black ball, and a box with more white balls when she observes a white ball. Accordingly, I define an *individual irrationality* as an observation which deviates from this prediction.

Definition 1. *Individual Irrationality: A choice in the individual condition where the subject observes a white (black) ball, but chooses a box with more black (white) balls.*

In the social interaction, the conventional assumption in economics is that individuals have rational expectations about each other (and rationality is common knowledge). In the context of the current experiment, this implies that the subject should follow her neighbor's guess and choose the same box as the neighbor in the social condition. Accordingly, a *social irrationality* is defined as follows.

Definition 2. *Social Irrationality: A choice in the social condition where the subject chooses a box different from her neighbor's guess.*¹⁷

¹⁷The definition of social irrationality in the *Ball* treatment is a little bit different because the subject observes both the ball and the neighbor's guess when she is connected to the neighbor. In that case, I define social irrationality as a choice in which the subject chooses a box different from her neighbor's guess, given that the neighbor's guess is rational (i.e., does not contradict with the signal).

In the next section, I analyse the experimental data to measure the magnitude of individual irrationality and social irrationality in subjects' choices and to elaborate on the differences.

4.1 Data

The main experiment was conducted at Toronto Experimental Economics Laboratory (TEEL) in University of Toronto during December 2019. The experiment was programmed in oTree ([Chen et al., 2016](#)). In total, 151 subjects were recruited from the subject pool using Online Recruitment System for Economic Experiments ([Greiner, 2015](#)). The average payment across subjects was \$25.26.¹⁸ Table 1 provides evidence that individual characteristics are relatively balanced across the four treatments, confirming that the randomization was successful.

Table 1: Summary Statistics

	Treatment			
	Base	Demographics	Bot	Ball
Female (%)	77.5	73.5	78.9	61.5
Prob/Stat course (%)	75	64.7	68.4	69.2
Years of Education	15.0 (1.5)	14.85 (2.11)	14.73 (2.24)	14.48 (1.82)
Age	20.25 (1.81)	19.38 (1.39)	20.02 (2.04)	20.28 (1.88)
Number of Subjects	40	34	38	39

Note: Standard errors are presented in parentheses. The second row shows the percentage of subjects who have taken Probability/Statistics courses.

In the following, I exclude the cases in which both boxes have 5 black balls and 5 white balls, $\theta_X = \theta_Y = 0.5$, because theory does not have a prediction about the subject's behavior in those

¹⁸No subject participated in more than a single treatment. Subjects needed to be at least 18 years old to be eligible to participate in the experiment. The human neighbors in the social condition were 94 subjects who had participated in the experiment a few months before the main experiment.

cases. Subjects are expected to behave randomly in those rounds, a result that is supported by the data.¹⁹

4.2 How Do Errors Differ across the Individual and the Social Conditions?

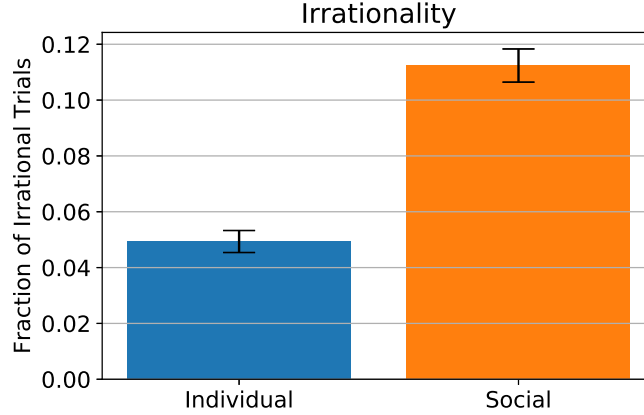


Figure 3: Individual irrationality and social irrationality (pooled data)

My first result examines the aggregate fraction of irrational choices in the individual condition and in the social condition. Figure 3 illustrates that the individual irrationality and the social irrationality are significantly greater than zero, even though subjects are incentivized for being correct. In the individual condition, subjects on average deviate from the theoretical prediction (Bayesian rational behavior) with a probability of 0.049 ($p\text{-value} < 0.001$).²⁰ In the social condition, even though subjects know that their neighbor's guess is incentivized with money, they on average do not follow the choice of their neighbor with a probability of 0.112 ($p\text{-value} < 0.001$). Surprisingly, the social irrationality is significantly higher than the individual irrationality ($p\text{-value} < 0.001$). This evidence suggests that subjects neglect the information more in the social condition than in the individual condition, a result that can be associated with the violation of rational expectations.²¹

¹⁹In the individual condition, when the two boxes have the same combination of balls (5 white and 5 black balls), subjects choose the left box with probability 0.44. Here, the null $H_0 : p = 0.5$ cannot be rejected at the 5% significance level ($p\text{-value} = 0.14$). Similarly, in the social condition, when both boxes have 5 white and 5 black balls, subjects do not follow their neighbor's guess with probability 0.48 ($p\text{-value} = 0.74$ for the null $H_0 : p = 0.5$).

²⁰This is lower than the error rate reported in prior literature for the individuals who arrive first in a standard social learning experiment (Anderson & Holt, 1997; Kübler & Weizsäcker, 2004). In Anderson and Holt (1997) 10% of the subjects whose information set was a private signal, did not follow their signal. In Kübler and Weizsäcker (2004), this behavior was observed in about 7% of all cases where first players saw only a private signal.

²¹Note that the comparisons in this section are within-subject, i.e., the same subjects are making on average more

Figure 4 presents the distribution of individual irrationality and social irrationality across subjects. The blue histogram shows that about 63% of subjects have no individual irrationality over the course of 21 rounds in the individual condition. In addition, 19.8% of subjects have exactly one individual irrationality, and the remaining 17.2% have more than one individual irrationality. So, the individual irrationality is not negligible for a considerable fraction of subjects.²² On the other hand, the orange histogram indicates that 54.3% of subjects have no social irrationality, 10.6% have exactly one social irrationality, and the remaining 35.1% have more than one. Comparing the two distributions, one can observe that the upper tail of the distribution is thicker in the social condition than in the individual condition. So, there is a clear shift in the error rate of subjects across the two conditions.

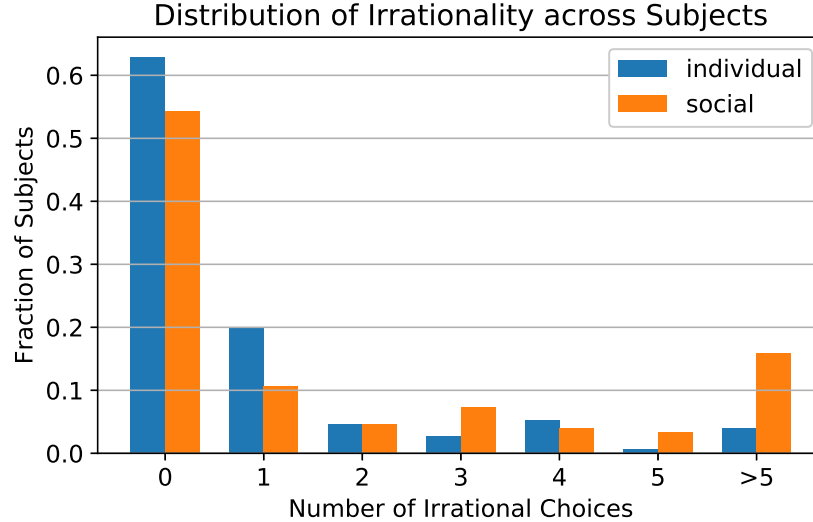


Figure 4: The distribution of individual irrationality and social irrationality across subjects

4.3 The Role of Uncertainty about Others' Behavior

In this section, I examine the mechanisms behind the additional neglect in the social condition.

The hypothesis is that the additional irrationality in the social condition is most likely due to the errors in the social condition than in the individual condition. Given my experimental design, it is also possible to do the analysis between-subject. The details of the between-subject analysis are provided in the appendix. The results are qualitatively similar there.

²²This result is consistent with [Ambuehl and Li \(2018\)](#) who report that 17% of their subjects made at least one irrational choice out of six trials.

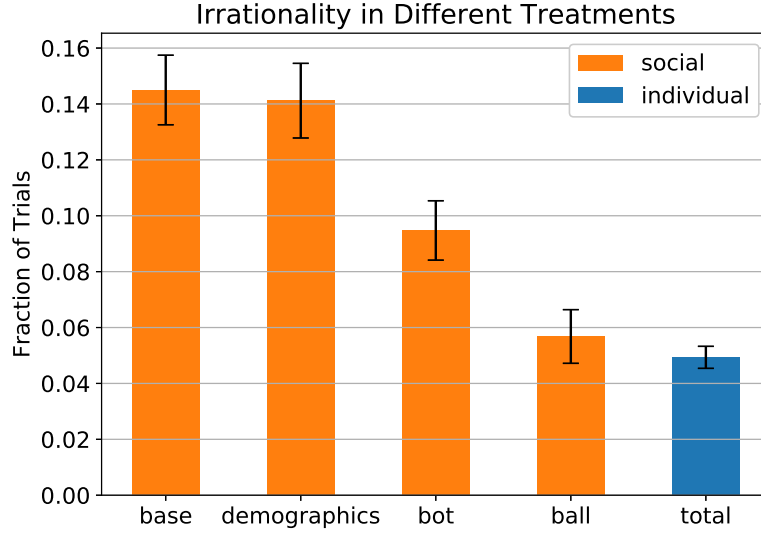


Figure 5: Social irrationality in different treatments

violation of rational expectations, which arises from the uncertainty about the neighbor’s behavior. In other words, because the neighbor’s decision-making process is ambiguous to the subjects, they cannot correctly extract the neighbor’s private information from observation of their choices, and thus violate the rational expectations. To test this idea, as noted earlier, I exogenously manipulate the subject’s knowledge about the neighbor across four treatments. Figure 5 illustrates the social irrationality in each treatment along with the aggregate individual irrationality.²³ The treatment “*base*” is a benchmark treatment in which subjects only observe the guess of their neighbor. The result in this treatment echos the earlier finding about the larger magnitude of neglect (irrationality) in the social condition than in the individual condition.

In the treatment “*demographics*”, subjects are provided with some demographic information about their neighbor (Age, Gender, Years of Education, Whether the neighbor took probability/statistics courses), on top of the neighbor’s guess. The result of this treatment shows that providing demographics slightly decreases the social irrationality compared to the base treatment, from 0.145 to 0.1411. But this effect is not statistically significant ($p\text{-value} = 0.83$).

In the treatment “*bot*”, subjects observe the guess of a computer bot. Here, the bot’s behavior

²³Recall that the individual condition is identical in all treatments. So, I do not break down the individual irrationality here and only report the aggregate individual irrationality for ease of exposition.

is known to subjects: it picks a box with more white balls when it observes a white ball, and picks a box with more black balls when it observes a black ball. The social irrationality in this treatment significantly drops to 0.094 compared to the base treatment ($p\text{-value} < 0.01$). This evidence is consistent with the hypothesis that the difference between the social irrationality and the individual irrationality is due to the uncertainty about the neighbor’s behavior.²⁴

Finally, in the treatment “*ball*”, subjects are provided with both the neighbor’s guess and the ball which was shown to the neighbor. Figure 5 shows that the social irrationality in this treatment is significantly lower than all other treatments ($p\text{-value} < 0.01$). Here, the difference between the social irrationality and the individual irrationality is no longer statistically significant. This result verifies that when there is no uncertainty about the neighbor’s behavior in the social condition, the magnitude of social irrationality is the same as the magnitude of individual irrationality. So, the additional neglect in the social condition disappears from the subject behavior.

4.4 The Observed Heterogeneity in Subjects’ Behavior

In this section, I run some regressions to examine the observed heterogeneity in the subjects’ behavior. My data contains demographic information about all the subjects and each of their neighbors. So, I can investigate how subjects’ characteristics and those of their neighbors explain the observed irrationality in the experiment.

First, I investigate the role of the subject’s characteristics. Specifically, I estimate the following regression,

$$Y_i \times 100 = \alpha + X_i\gamma + \epsilon \tag{1}$$

where Y_i is the fraction of irrational choices by subject i , X_i includes the subject’s observed characteristics: gender (dummy for female), years of education, age, and whether the subject has taken Probability/Statistics courses. The estimation results are provided in Table 2.

²⁴Note that although the social irrationality is alleviated in the bot treatment, it is still significantly higher than the individual irrationality. One natural question rises here: why is there a difference between the individual irrationality and the social irrationality in the bot treatment? Responses from an open survey question that was collected at the end of the experiment show that some of the subjects mistrust bots. This might explain why the social irrationality in the bot treatment remains significantly higher than the individual irrationality.

Table 2: The observed heterogeneity in the subject’s irrationality

DV	Data	
	Individual	Social
	(1)	(2)
	$Y_{ic} \times 100$	$Y_{ic} \times 100$
Gender (Female)	0.47 (1.84)	5.58* (2.88)
Education	1.00** (0.49)	0.11 (0.78)
Age	− 0.51 (0.54)	− 1.36 (0.85)
Prob/Stat course	− 4.02** (1.89)	− 6.85** (2.95)
Constant	2.81 (9.35)	37.19** (14.64)
Observations	151	151
R ²	0.063	0.097
Adjusted R ²	0.037	0.073
<i>Note:</i> * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$		

Columns (1) in Table 2 indicates that in the individual condition, all else equal, subjects who have taken Probability/Statistics courses make 4.02 percentage points less errors than subjects who have not. This result suggests that the individual irrationality is mainly driven by the lack of knowledge in probability and statistics. Column (2) illustrates that in the social condition, all else equal, the subjects who have taken Probability/Statistics courses make 6.85 percentage points less errors than subjects who have not. Our results are consistent with the previous findings in the literature. For example, [Armantier et al. \(2015\)](#) compare the inflation expectations reported by consumers in a survey with their behavior in a financially incentivized investment experiment and find evidence that respondents whose behavior cannot be rationalized by economic theory tend to have lower education and lower numeracy and financial literacy. In our experiment, however, the

subject’s observable characteristics cannot explain the additional irrationality that is associated with the uncertainty about the neighbor’s behavior in the social condition (i.e. violation of rational expectations).

Next, I examine the effect of the neighbor’s observable characteristics on the subject’s social irrationality. Note that only the subjects who are in the treatment *demographics* observe their neighbor’s characteristics. So, I need to restrict the data in this section to the choices made in the social condition of the demographics treatment. Here, the dependent variable is a binary choice. Hence, I estimate the following logistic regression,

$$Pr(D_{ij} \text{ is irrational}) = \frac{\exp (\alpha + X_i\gamma + X_j\delta)}{1 + \exp (\alpha + X_i\gamma + X_j\delta)} \quad (2)$$

where D_{ij} is the choice of subject i in round j , X_i includes the subject’s observable characteristics, and X_j includes the neighbor’s observable characteristics in round j (recall that the neighbor randomly changes in each round). As before, observable characteristics include gender, years of education, age, and whether the individual has taken Probability/Statistics course. The estimation results are shown in Table 3.

Columns (1) and (2) in Table 3 present the estimated coefficients for equation (2). The coefficients are insignificant in the first column. However, the second column shows that the neighbor’s observable characteristics have a statistically significant effect on the subject’s behavior: *ceteris paribus*, the subject is more likely to follow a neighbor whose age is higher, whose gender is female (versus male), and who has taken Probability/Statistics courses.

The coefficients of a logistic regression are not quantitatively interpretable. So, I report the average marginal effects in columns (3) and (4) of Table 3. The results in column (4) imply that, *ceteris paribus*, a subject is likely to follow a neighbor who has taken Probability/Statistics courses 5.8 percentage points more than a neighbor who has not. In addition, all else equal, a subject is 4.7 percentage points less likely to make an irrational guess when interacting with a female versus a male neighbor (4.7 percentage points more likely to follow the neighbor). The effect of the neighbor’s age is very small though, i.e., one year increase in the neighbor’s age increases the likelihood of being followed by the subject by 0.3 percentage points.

Table 3: The observed heterogeneity in the social condition of treatment “*demographics*” (Equation 2)

	Logit Coefficients		Average Marginal Effects	
	(1)	(2)	(3)	(4)
Subject’s Gender (Female)	0.52 (0.59)	0.48 (0.60)	0.062 (0.071)	0.056 (0.07)
Subject’s Education	−0.139 (0.17)	− 0.159 (0.17)	− 0.016 (0.02)	− 0.018 (0.02)
Subject’s Age	− 0.08 (0.23)	− 0.075 (0.23)	− 0.009 (0.0267)	− 0.008 (0.026)
Subject’s Prob/Stat course	− 0.44 (0.63)	− 0.46 (0.64)	− 0.052 (0.076)	− 0.052 (0.076)
Neighbor’s Gender (Female)		− 0.41** (0.20)		− 0.047* (0.025)
Neighbor’s Education		− 0.04 (0.032)		− 0.005 (0.003)
Neighbor’s Age		− 0.025** (0.01)		− 0.003*** (0.001)
Neighbor’s Prob/Stat course		− 0.51*** (0.196)		− 0.058** (0.025)
Constant	2.18 (3.2)	4.19 (3.51)		
Observations	680	680	680	680
Pseudo R ²	0.035	0.061	0.035	0.061

Note: Standard errors are clustered at the subject level.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4.5 Heterogeneity across Information Structures

A unique feature of the experiment is that subjects make decisions given several information structures. This allows me to examine the extent to which subjects' irrationality vary across signal precision. Figure 6 shows the total fractional of irrational choices conditional on each information structure. The general trend is that the irrationality decreases as signal precision increases. However, even in the extreme case of complete certainty ($\theta_X = \theta_Y = 1$) the irrationality is not zero. The notable insight is that numbers are significantly higher in the social condition (right figure) than the individual condition (left figure) which is consistent with our earlier findings.

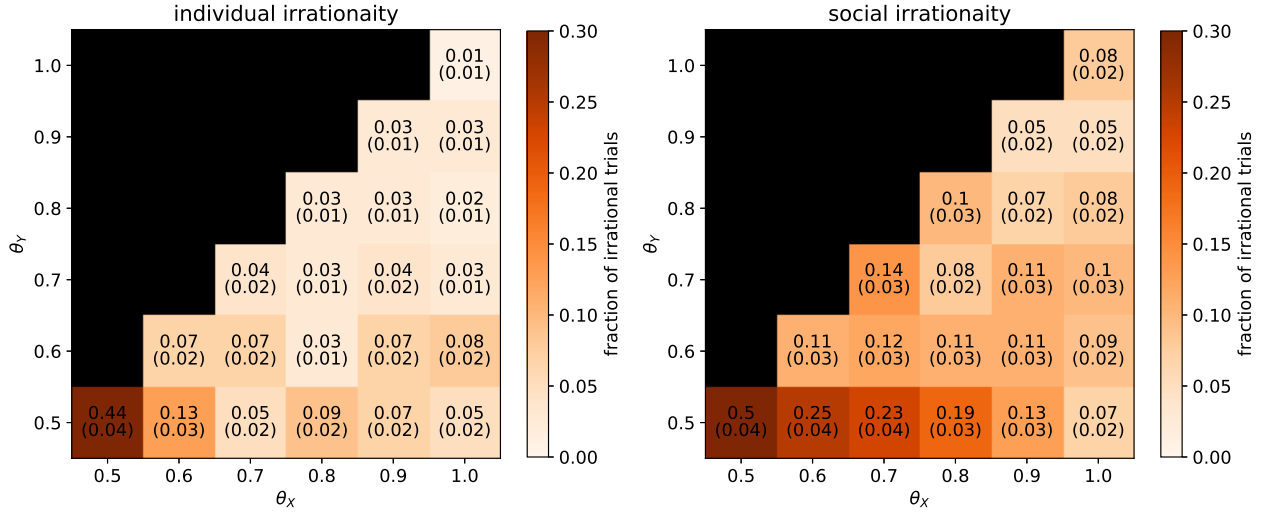


Figure 6: The fraction of irrational choices conditional on the information structure. Standard errors are shown in parentheses.

Note that even though each figure covers half of the possible combinations of θ_X and θ_Y , these results are equally applicable to the uncovered cases. This is because if one exchanges the contents of box X and box Y, the environment remains the same but the information structure changes from (θ_X, θ_Y) to (θ_Y, θ_X) . Hence, one can naturally extend these figure so that they cover the whole space of information structures.

Another way to examine subjects' behavior is to measure irrational choices conditional on both the information structure and the realized signal. This is important specially in cases where the

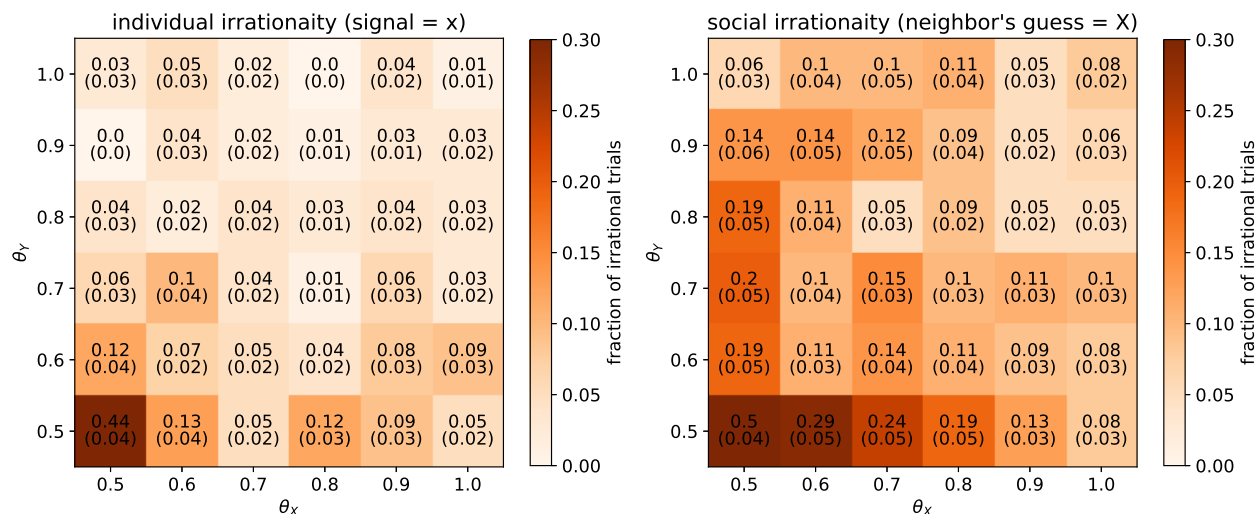


Figure 7: The fraction of irrational choices conditional on the information structure and the signal being x (in the individual condition) or a neighbor's guess of X (in the social condition). Standard errors are shown in parentheses.

information structure is asymmetric. As an example, consider a case in which $\theta_X = 1$ and $\theta_Y = 0.5$. In this case, box X has 10 white balls and box Y has 5 white and 5 black balls. Here, when the signal is a black ball, a Bayesian subject is certain that the state is box B. However, when the signal is a white ball, the Bayesian subject assigns a relatively higher likelihood to box X than box Y, but she would not be 100% sure about the state. So, one might expect to see a higher irrationality in the latter case than the former case.

Without loss of generality and for expositional purposes, I permute observations so that all the signals in the individual condition are a white ball, and all the signals in the social condition are a neighbor's guess of X . This is possible because of the nature of the experiment. For example, an observation in the individual condition where $(\theta_X, \theta_Y) = (0.5, 1)$ and the signal is y (black ball), can be permuted so that $(\theta_X, \theta_Y) = (1, 0.5)$ and the signal is x (white ball).²⁵ The results are presented in Figure 7.

Figure 7 demonstrates that errors are generally higher in cases where the uncertainty is higher.

²⁵This way, one only needs to plot one figure for each condition, and the figure shows irrationalities conditional on all information structures and signal " x ". If I do not permute observations, then I need to plot two figures for each condition: one figure would be conditioned on signal x (but covers half of the information space as in Figure 6) and the other figure would be conditioned on signal y .

For example, when the subject observes x conditional on $(\theta_X, \theta_Y) = (1, 0.5)$, the irrationality is higher than the mirror case in which $(\theta_X, \theta_Y) = (0.5, 1)$. As we discussed, this is the case because when a subject observes signal x in the former case, she is not certain about the state. However, observing the same signal in the latter case, the subjects would be almost certain about the state. Comparing the results across conditions, the numbers are generally higher in the social condition than the individual condition. So, even though subjects react to changes in the likelihood, the uncertainty about the neighbor’s behavior remains stable across information structures.

5 A Model of Decision-making

In this section, I introduce a framework to describe the individual decision-making process in the context of my experiment. I then provide two explanations for an observed error in the individual’s choice. Finally, I combine the subjects’ choice data with their responses to the survey questions, that measure their relative posterior beliefs in each round, to uncover the channel that is associated with the violation of rational expectations.

In the context of my experiment, the individual’s decision-making process can be modeled as a two stage procedure. Upon observing a signal, the subject first updates her belief and then picks one of the two possible states based on her posterior. This process is shown in Figure 8.

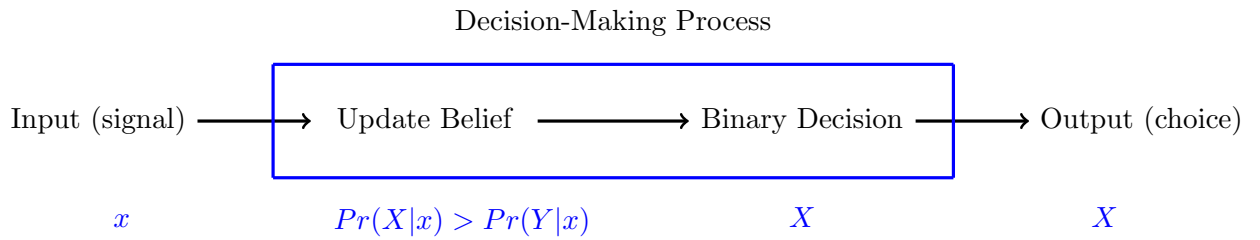


Figure 8: The individual decision-making process

This model is closely related to the deliberation process introduced in a more general decision framework by Walliser (1989). In the first stage, the subject treats the available data to form expectations on her environment (“cognitive rationality”). In the second stage, these expectations

are used to find out a selected action (“instrumental rationality”). As an example, suppose an individual observes signal x . In theory, the individual should update her belief in favor of state X in the first stage, $Pr(X|x) > Pr(Y|x)$, and then choose state X in the second stage. However, as I showed earlier, the individual’s choice (output) does not always comply with the signal (input). Individuals frequently make errors in this simple task. Now, consider a case in which a subject obtains signal x , but her final guess is Y (Figure 9). There are two explanations for this observation:

1. **Posterior Error:** It might be that the subject’s posterior is mistakenly in favor of state Y , $Pr(Y|x) > Pr(X|x)$, and this causes the subject to make an erroneous decision. This means that the subject’s posterior is in a wrong direction, but her choice is consistent with her incorrect posterior.²⁶
2. **Reasoning Error:** It might be that the subject’s posterior is correctly in favor of state X , $Pr(X|x) > Pr(Y|x)$, but she mistakenly chooses state Y .

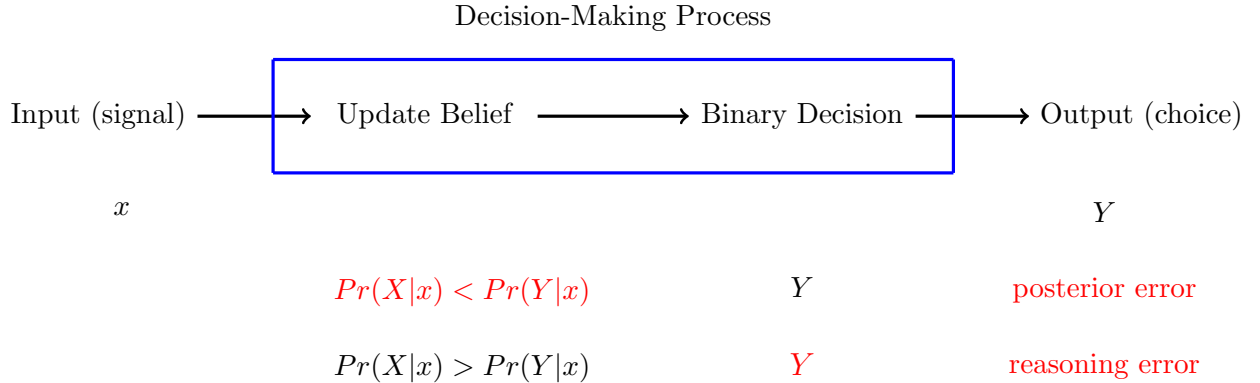


Figure 9: Two explanations for an observed error in the individual’s choice

In general, it is not possible to identify these two explanations by only observing the subject’s choice. This implies that to overcome the identification challenges, one need extra variation in the

²⁶It is important to notice that the posterior error is different from what is commonly known as “belief updating bias” in the literature (Kahneman & Tversky, 1972; Benjamin, 2019). A biased belief is not necessarily in a wrong direction, i.e., the biased belief and the Bayesian belief can both favor the same state while assigning different likelihoods to that state. For example, when the signal implies a 70% chance (in theory) to an event, a biased belief may assign a 60% chance to it (both are greater than 50%). However, a posterior error is the consequence of a severe bias that switches the direction of the posterior probability, i.e., an updated belief that is in a wrong direction (e.g. a belief of less than 50% in the earlier example).

data or additional information about subject’s behavior. As stated before, my experimental design solves this identification problem by collecting data on both subjects’ choices and their self-reported posteriors. So, I can distinguish between a posterior error and a reasoning error in the data.²⁷

Figure 10 illustrates the break down of the observed individual irrationality and social irrationality into posterior error and reasoning error. The figure shows that the probability of a reasoning error is equal to 0.018 in both the individual condition and the social condition. However, there is a statistically significant difference between the magnitude of posterior error across the two conditions; it is 0.031 in the individual condition, but 0.095 in the social condition ($p\text{-value} < 0.01$).

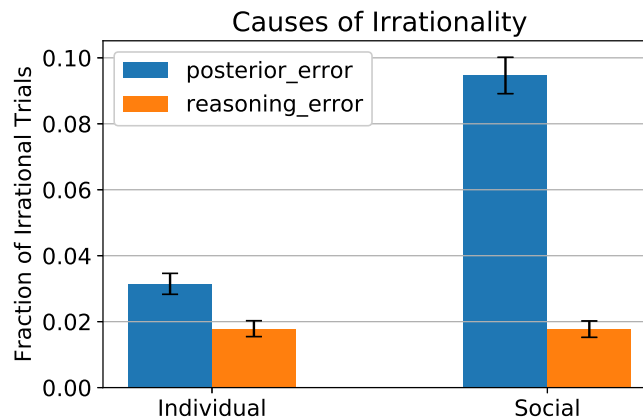


Figure 10: The break down of individual irrationality and social irrationality into posterior error and reasoning error

Figure 10 provides an important insight about the mechanism behind the violation of rational expectations in the social interaction. It suggests that the uncertainty about the neighbor’s behavior makes the belief updating more difficult in the social condition than in the individual condition. That is, subjects make more errors when they update their beliefs (e.g. the first stage of the decision-making process) in the social condition than in the individual condition. However, this uncertainty does not influence the second stage of the decision-making process, i.e., once posterior beliefs are formed, subjects follow the same reasoning procedure, and thus the magnitude of the reasoning error remains unchanged across the individual condition and the social condition.

Distinguishing between posterior errors and reasoning errors identifies a critical touchpoint in

²⁷The framework that is introduced here applies to both the individual condition and the social condition of my experiment. The only difference is that the signal (input) is a ball in the individual condition, while it is the guess of a neighbor (and any additional information coming along the neighbor’s choice) in the social condition.

the decision-making process and may provide solutions to nudge individuals towards making better decisions in social environments. It suggests that the violation of rational expectations may not necessarily result from the lack of math/Stats knowledge, but may be more about how uncertain subjects are about their neighbor’s behavior. In addition, failing to control for individual errors that are independent of the social environments (e.g. reasoning errors) may lead to unintended consequences. In appendix C, I estimate a unified model of individual behavior by borrowing techniques from the social learning literature (Grether, 1980; Anderson & Holt, 1997). There, I show that not accounting for different sources of error (posterior error versus reasoning error) can bias the estimated response precision parameter that intends to describe the individual behavior.²⁸ Orhun (2012) raises a similar point in a strategic setting and discusses econometric biases arising from a failure to account for the correlation between beliefs and preferences. My results highlight this insight in a situation without correlation between beliefs and preferences and where there is no strategic incentives between individuals.

6 Non-parametric Estimation of Decision-making Process

In this section, I develop and estimate a model that corresponds to the two-stage decision-making process introduced in Section 5. The estimation provides two non-parametric functions for each subject in each condition (individual versus social): one represents the belief updating process in stage 1, and the other represents the choice rule used in stage 2. For the first stage, I use the self-reported belief data, and for the second step, I use the combination of the choice and the belief data. The model is separately estimated for the individual condition and the social condition. Based on the earlier result in Figure 10, the identification of three types of errors introduced earlier is as follows: the comparison of belief functions in stage 1 across the individual and the social condition separates type A error from type C error, and the choice function in stage 2 identifies type B errors. Table 4 summarizes the identification insights from this section. Note that it was shown earlier that reasoning errors remain unchanged across the individual and social condition so

²⁸The predominant approach in the social learning literature is to derive predictions under the assumption that all other players obey a given model solution, for instance Perfect Bayesian Equilibrium or Quantal Response Equilibrium. But despite their undisputable usefulness, these solutions are often inaccurate descriptions of behavior and thus yield imperfect benchmarks Weizsäcker (2010).

that the comparison of stage 2 across conditions does not identify any new effect. If one believes there are other errors in the second stage which might be related to the social environment, then those errors can be also identified from this comparison.

	Stage 1 (belief)	Stage 2 (choice)
Individual	type A	type B
Social	type A + type C	type B

Table 4: The identification of three types of errors

6.1 Beliefs in Stage 1

Individual Condition. Denote subject i 's posterior belief about state X after observing signal s by $\mu_i(X|s; \theta)$, where $\theta = (\theta_X, \theta_Y)$ is the information structure and is common knowledge. I only assume $\mu_i(X|s; \theta) = 1 - \mu_i(Y|s; \theta)$ and do not impose any other assumption on μ_i so that the belief is individual-specific and as general as possible.²⁹ A unique aspect of the experiment is that it contains a wide variety of information structures so that there is enough variation in μ in the probability space and I am able to plot the distribution of beliefs across subjects at different points. Figure 11 presents the distribution of these beliefs (μ_i) across subjects against the Bayesian posterior.³⁰

If subjects were Bayesian, one would expect all points be exactly on the 45-degree line and the standard deviation be zero. The first fact is that posterior beliefs are positively correlated with Bayesian beliefs. So, subjects generally understand the changes in likelihood. However, one can see over-appreciation of probabilities in the left hand side and under-appreciation of probabilities in the right hand side. The average self-reported posterior follows a trend which in essence is similar to the well-known probability weighting transformation in Cumulative Prospect Theory (Tversky & Kahneman, 1992). This pattern is the result of what was called type A error in belief updating.

²⁹The well known model of Grether (1980) is a special case of my model in which μ is the same across subjects (no heterogeneity) and it has a specific parametric form: $\mu = \frac{Pr(s|X; \theta)^c}{Pr(s|X; \theta)^c + Pr(s|Y; \theta)^c}$.

³⁰Note that the shaded area shows one standard deviation above and below the mean. In the data, no subject reports probabilities under 0 and over 100. However, when one plots the mean values with one standard deviation around it, the shaded area may contain probabilities out of $[0, 100]$ range in corner points. This shall not mislead the reader.

Most importantly, there is substantial heterogeneity in beliefs so that the standard deviation is non-negligible and large specifically at boundaries (0 and 100). This is surprising because it implies some subjects make errors even in the extreme cases of certainty. So, to study the errors in the social interactions, one needs to carefully account for this type of heterogeneity.

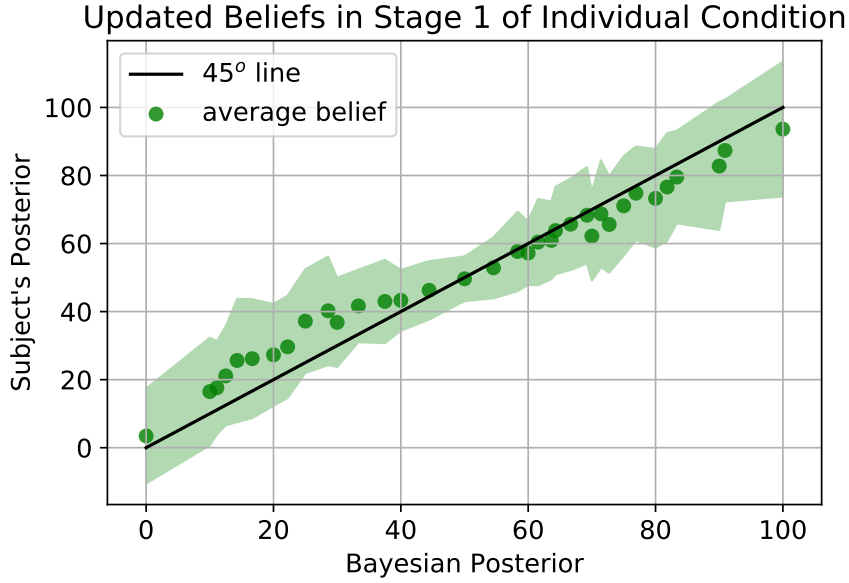


Figure 11: The belief distribution across subjects in the individual condition

Social Condition. Denote the subject's posterior belief about state X after observing a neighbor's guess of g by $\gamma(X|g; \theta)$, where $\theta = (\theta_X, \theta_Y)$ is common knowledge. I keep γ general and impose only one assumption on it: $\gamma(X|g; \theta) = 1 - \gamma(Y|g; \theta)$. Figure 12 presents the distribution of these beliefs (γ_i) across subjects against the Bayesian rational posterior. The overall trend here is similar to that of individual condition. There is over-appreciation in the left side and under-appreciation in the right side. One difference between Figures 11 and 12 is that the variance is higher in the social condition than the individual condition. This is probably due to higher uncertainty included in the social condition. Recall that beliefs in the social condition are impacted by type A and type C errors. So, higher variance would be a natural implication in the social condition.

To further separate type A and type C errors one needs to compare the average beliefs across conditions. One can see from Figures 11 and 12 that the average distance to 45-degree line seems to

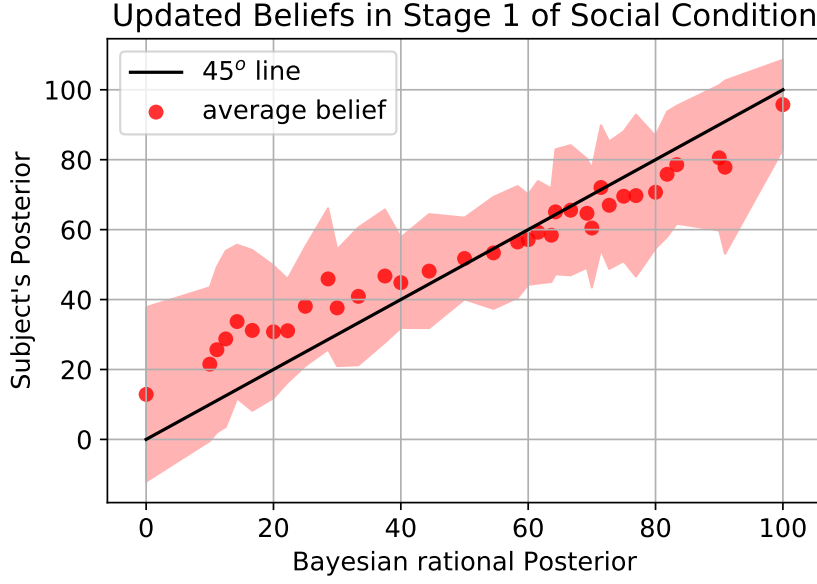


Figure 12: The belief distribution across subjects in the social condition

be higher in the social condition than the individual condition. To further clarify on this comparison, I separately and non-parametrically estimate a belief function for each condition. The results of the Nadaraya-Watson kernel regression (bandwidth = 15) is presented in Figure 13. As expected, the belief curve for the social condition exhibits a higher probability weighting bias compared to the individual condition. This effect is what was earlier named as type C error and is due to violation of rational expectations.

6.2 Choices in Stage 2

Individual Condition. In the second step of decision-making, conditional on the posterior belief, μ_i , the subject makes a binary decision. In the most general form, the subject's decision can be modeled as a mixed-strategy on X and Y : she chooses state X with probability $\alpha_i(\mu_i)$ and state Y with probability $1 - \alpha_i(\mu_i)$. Note that the standard Quantal Response Equilibrium (McKelvey & Palfrey, 1995) imposes a parametric assumption on this so that $\alpha_i(\mu_i)$ is a logistic function. Also, Perfect Bayesian Equilibrium is a special case where μ_i is the Bayesian posterior and α_i is degenerate on the state with higher posterior probability. I combine the choice data from the second stage of the experiment with the belief data from the first stage and plot the distribution

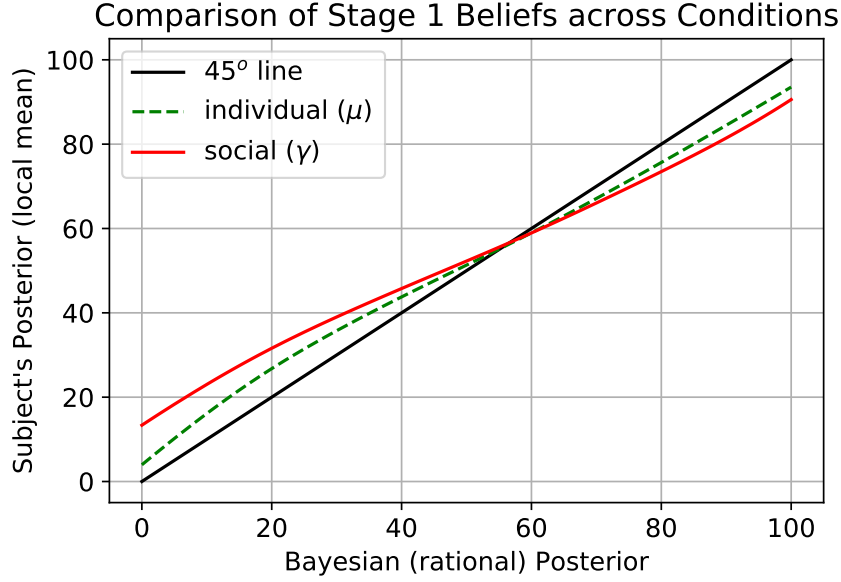


Figure 13: The non-parametric estimation of mean posterior belief in the individual and social conditions

of α_i at each point in Figure 14.

If subjects were fully rational, all the points with $\mu_i < 50$ would have been on $Choice = 0$ and all the points with $\mu_i > 50$ would be on $Choice = 1$, and the standard deviation would be zero at all given μ_i 's. The first general observation from Figure 14 is that subjects respond to changes in posteriors, meaning that they are generally more likely to choose an option that has a higher probability. However, the positive standard errors demonstrate that, as discussed earlier, subjects may sometimes make reasoning errors, i.e., they may pick an option that has a lower probability. This pattern can be consistent with a logistic function but one should take careful considerations regarding the substantial heterogeneity in choices. This analysis identifies the so called type B errors in subject's behavior.

Social Condition. In the second stage of the social condition, the subject makes a binary decision based on the neighbor's choice. The subject's decision is a mixing probability on X and Y : conditional on her posterior belief, γ_i , the subject chooses state X with probability $\beta_i(\gamma_i)$ and state Y with probability $1 - \beta_i(\gamma_i)$. The distribution of β_i is shown in Figure 15.

The general pattern in Figure 15 is similar to that of the individual condition (Figure 14). However, to further check for the differences, I separately and non-parametrically estimate a choice

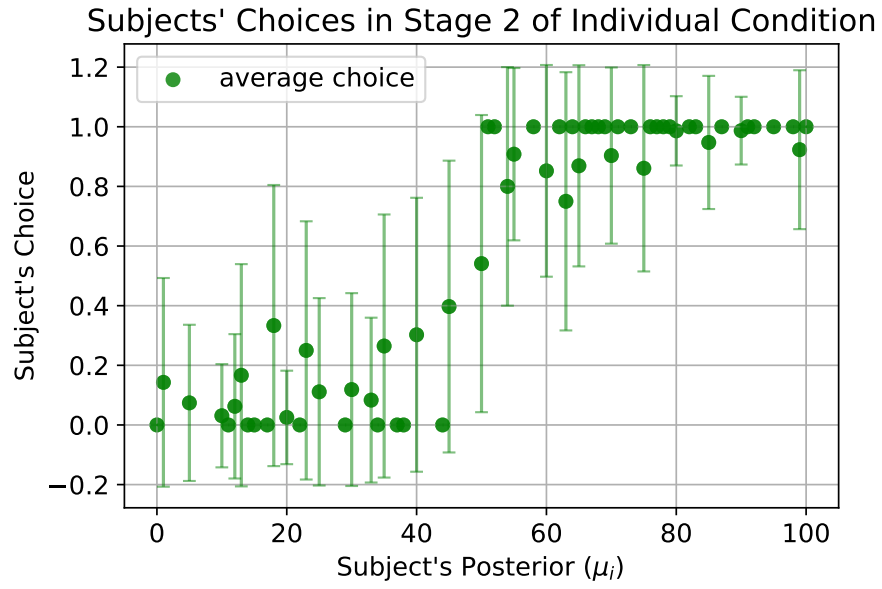


Figure 14: The choice distribution across subjects in the individual condition

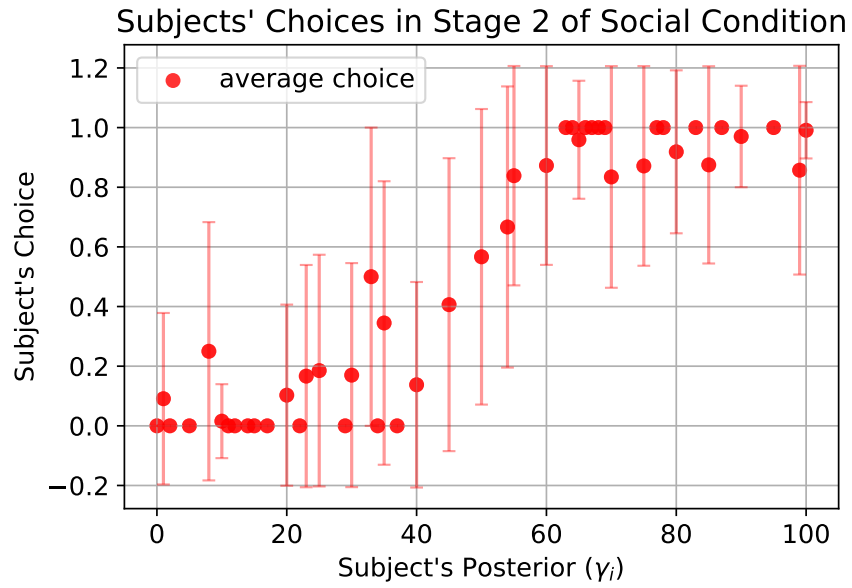


Figure 15: The choice distribution across subjects in the social condition

function for each condition. The results of the Nadaraya-Watson kernel regression (bandwidth = 15) is presented in Figure 16. As expected, the choice curve for the social condition exhibits a very similar pattern to the individual condition.³¹

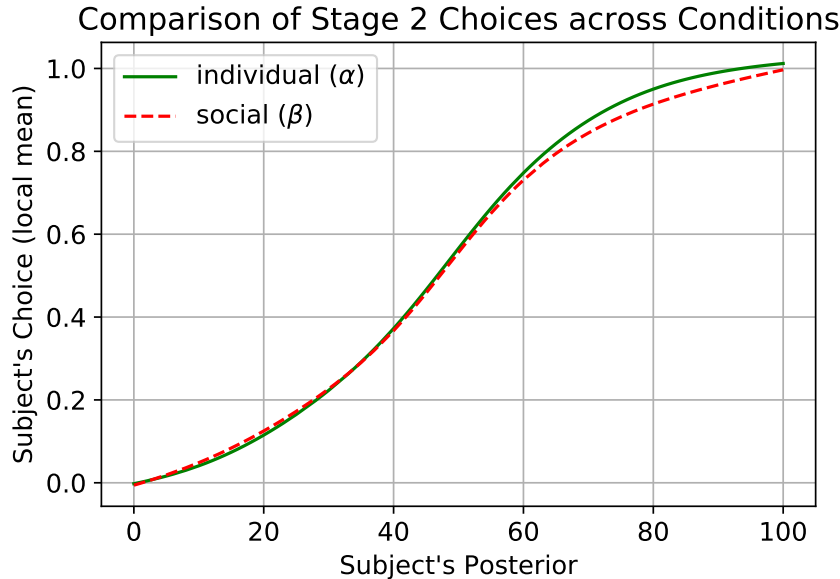


Figure 16: The non-parametric estimation of choice probabilities in the individual and social conditions

7 Conclusions and Implications

Many economic decisions, from the most mundane ones, like choosing a diner for lunch, to the most important ones, like adoption of a new technology by a firm, require making inferences about an underlying state of the world. In these situations social interaction is an important source of information. Economic agents often interact with each other via observation of choices. Such social interactions affect people's beliefs and can help them to make informed decisions. For example, traders in financial markets typically try to infer information about asset fundamental values from the order flow. Companies make investment decisions using consultants. Editors of scien-

³¹Note that, as I mentioned earlier, the experimental design is such that it can also identify errors related to SI that happen in the second stage of decision-making and are separate from the other three types of error. So, one might argue that the small difference between choice probabilities in the right hand side of Figure 16 is actually due to such errors. However, I believe this small difference is not due to a systematic error. Recall that in the previous section, I showed that the average reasoning error remains unchanged across conditions. So, I think the difference here might be just because of luck or may be related to the noise in the data (here, the analysis is at a more granular level than the previous section).

tific journals make their editorial decisions on the basis of referee reports. In a different context, [Bartels \(1988\)](#) discusses “cue-taking” in the U.S. presidential nomination campaigns in which an individual’s beliefs about a candidate is influenced by the decisions of others.

The conventional assumption in the literature is that decision makers have rational expectations about each other. However, individuals often do not observe other decision makers’ expectations. So, they may not always have precise predictions about other people’ behavior, and this may lead to a socially inefficient equilibrium when information is transmitted by observation.

In this study, I conduct a series of laboratory experiments to examine why individuals may fail to exhibit rational expectations in social interactions. I use a relatively simple and novel experimental setting to disentangle between individual errors that are independent of the social environment, and the errors that are caused by the violation of rational expectations. In a within-subject design, I compare subjects’ choices across an isolated condition and a social condition, and show that subjects make more errors in the presence than in the absence of social interaction, even when they receive informationally equivalent signals across the two conditions. That is, they neglect the provided information more when they interact with others than when they do not. Here, the additional neglect is associated with the violation of rational expectations.

To uncover the mechanism behind the additional neglect in the social condition, I design a series of treatment variations by exogenously manipulating the subject’s knowledge about her neighbor. I find that the unexpected irrationality in the social condition is mainly driven by the uncertainty about other people’s behavior: subject’s behave as if they lack knowledge of how others make decision based on their private signals. The implication of this result is that social interactions might not be as effective as one expects in theory. So, one should take careful considerations in examining the effects of social interactions using field data.

In econometric models of field data, researchers often impose the assumption of Bayesian rationality. While one should understand the need for these assumptions, to overcome identification challenges, it has been unclear whether and how one can identify different types of errors in decision makers’ behavior. By introducing a unique experimental design, this paper sheds lights on three categories of errors that often emerge in social interactions: errors associated with belief updating, errors due to incorrect reasoning, and errors related to the violation of rational expectations. I

show that measuring probabilistic beliefs is important because it provides extra data which can be utilized to identify new aspects of individual behavior. However, to non-parametrically estimate the complete decision-making process, one needs extra variations in the data. While the estimation is specific to the experiment context of this study, similar analysis can be used to separate various errors in decision-making using field data. Identification of non-equilibrium beliefs is an important topic that has recently attracted substantial attention ([Aguirregabiria & Magesan, 2020](#); [Aguirregabiria & Xie, 2021](#)). I believe my study extends this literature to situations where there is no strategic incentives between players and identifies the contribution of different errors to subjects' behavior. The results of this study can provide guidance to researchers and managers on how to benefit from measuring beliefs and incorporate them into their analysis.

A fruitful avenue for future research is to disentangle between errors shown in this study and other effects documented in the literature. For instance, [Orhun \(2012\)](#) and [Orhun and Urminsky \(2013\)](#) document the correlation between one's own evaluations and his estimates of other people's evaluations. This may need collecting new data and substantial variation in the data so that one can separately identify these new effects.

References

- Aguirregabiria, V., & Magesan, A. (2020). Identification and estimation of dynamic games when players' beliefs are not in equilibrium. *Review of Economic Studies*, 87(2), 582–625.
- Aguirregabiria, V., & Xie, E. (2021). Identification of non-equilibrium beliefs in games of incomplete information using experimental data. *Journal of Econometric Methods*, 10(1), 1-26.
- Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84, 488–500.
- Ambuehl, S., & Li, S. (2018). Belief updating and the demand for information. *Games and Economic Behavior*, 109, 21–39.
- Anderson, & Holt, C. A. (1997). Information cascades in the laboratory. *American Economic Review*, 87(5), 847–862.
- Angrisani, M., Guarino, A., Jehiel, P., & Kitagawa, T. (2020). Information redundancy neglect versus overconfidence: A social learning experiment. *American Economic Journal: Microeconomics*, forthcoming.
- Armantier, O., de Bruin, W. B., Topa, G., Klaauw, W. V. D., & Zafar, B. (2015). Inflation expectations and behavior: Do survey respondents act on their beliefs? *International Economic Review*, 56(2), 505-536.
- Bartels, L. M. (1988). Presidential primaries and the dynamics of public choice. *Princeton, N.J.: Princeton Univ. Press*.
- Beard, T. R., & Beil, R. O. (1994). Do people rely on the self-interested maximization of others? an experimental test. *Management Science*, 40, 252-262.
- Benetton, M., & Compiani, G. (2021). Investors' beliefs and cryptocurrency prices. *working paper*.
- Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations*, 2, 69-186.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5), 992-1026.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78, 1–3.

- Camerer, C., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3), 861-898.
- Charness, G., & Levin, D. (2009). The origin of the winner's curse: A laboratory study. *American Economic Journal: Microeconomics*, 1 (1), 207-236.
- Chen, D., Schonger, M., & Wickens, C. (2016). otree - an open-source platform for laboratory, online and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88-97.
- Ching, A. T., Hossain, T., Tehrani, S. S., & Zhao, C. (2021). How do people update beliefs? evidence from the laboratory. *working paper*.
- Cyert, R., & DeGroot, M. (1974). Rational expectations and bayesian analysis. *Journal of Political Economy*, 82, 521-536.
- Dominitz, J., & Hung, A. A. (2009). Empirical models of discrete choice and belief updating in observational learning experiments. *Journal of Economic Behavior and Organisation*, 69, 94-109.
- Draganska, M., Misra, S., Aguirregabiria, V., Bajari, P., Einav, L., Ellickson, P., ... Zhu, T. (2008). Discrete choice models of firms' strategic decisions. *Marketing Letters*, 19(3), 399-416.
- Edwards, W. (1968). Conservatism in human information processing. in (B. Kleinmuntz, ed.), *Formal Representation of Human Judgment*, 1st edition, New York: Wiley.
- Enke, B. (2020). What you see is all there is. *Quarterly Journal of Economics*, 135(3), 1363-1398.
- Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *Review of Economic Studies*, 86, 313-332.
- Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1), 1-20.
- Eyster, E. (2019). Errors in strategic reasoning. *Handbook of Behavioral Economics: Applications and Foundations*, 2, 187-259.
- Filippis, R. D., Guarino, A., Jehiel, P., & Kitagawa, T. (2021). Non-bayesian updating in a social learning experiment. *Journal of Economic Theory*, forthcoming.
- Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., ... Verleghe, P. (2005). The firm's management of social interactions. *Marketing letters*, 16(3), 415-428.
- Goeree, J. K., Palfrey, T. R., & Rogers, B. W. (2007). Self-correcting information cascades. *Review*

- of Economic Studies*, 74 (3), 733-762.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, 1, 114-125.
- Grether, D. (1980). Bayes rule as a descriptive model: the representatives heuristic. *The Quarterly Journal of Economics*, 95(3), 537-557.
- Hartmann, W. R., Manchanda, P., Nair, H., Bothner, M., Dodds, P., Godes, D., ... Tucker, C. (2008). Modeling social interactions: Identification, empirical methods and policy implications. *Marketing letters*, 19(3), 287-304.
- Huck, S., & Weizsäcker, G. (2002). Do players correctly estimate what others do? evidence of conservatism in beliefs. *Journal of Economic Behavior & Organization*, 47, 71-85.
- Jindal, P., & Aribarg, A. (2021). The importance of price beliefs in consumer search. *Journal of Marketing Research*, forthcoming.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cognitive Psychology*, 3 (3), 430-454.
- Kalai, E., & Lehrer, E. (1993). Rational learning leads to nash equilibrium. *Econometrica*, 61, 1019-1045.
- Kübler, D., & Weizsäcker, G. (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *Review of Economic Studies*, 71, 425-441.
- Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, 14 (3), 115-136.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5), 1329-1376.
- March, C., & Ziegelmeyer, A. (2018). Excessive herding in the laboratory: The role of intuitive judgments. *Cesifo working paper series no. 6855, CESifo, Munich..*
- McKelvey, R., & Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10, 6-38.
- Nöth, M., & Weber, M. (2003). Information aggregation with random ordering: Cascades and overconfidence. *The Economic Journal*, 113(484), 166-189.
- Orhun, Y. (2012). Systematic differences in beliefs about others in strategic interactions. *working paper*.

- Orhun, Y., & Urminsky, O. (2013). Conditional projection: How own evaluations influence beliefs about others whose choices are known. *Journal of Marketing Research*, 50(1), 111-124.
- Phillips, L., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3), 346-354.
- Schelling, T. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, 143-86.
- Smith, V. L. (1991). Rational choice: The contrast between economics and psychology. *Journal of Political Economy*, 99(4), 877- 897.
- Stahl, D., & Wilson, P. (1995). On players' model of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10, 218-254.
- Stone, D. F., & Zafar, B. (2014). Do we follow others when we should outside the lab? evidence from the ap top 25. *Journal of Risk and Uncertainty*, 49(1), 73-102.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323.
- Walliser, B. (1989). Instrumental rationality and cognitive rationality. *Theory and Decision*, 27, 7-36.
- Weizsäcker, G. (2003). Ignoring the rationality of others: Evidence from experimental normal form games. *Games and Economic Behavior*, 44, 145-171.
- Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, 100 (5), 2340-2360.

Appendix A1: Experiment Instructions

Welcome and thank you for participating in this experiment. Please read this instruction carefully. All participants in this experiment are recruited in the same way as you and read the same instructions as you do. It is important that you do not discuss the details of this experiment with anyone else after the experiment.

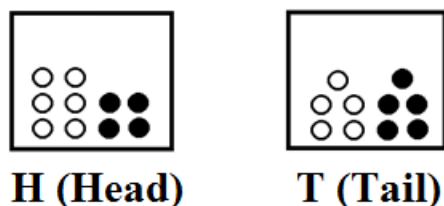
You will receive \$6 for participating in this experiment. During the course of the experiment you can earn more. Your earning will depend on your performance during the experiment. All your decisions and your earning will be treated confidentially.

This session is part of an experiment about how humans make decision under uncertainty. The experiment consists of multiple parts. You need to read the following instruction to understand what will happen in each part of the experiment.

FIRST PART:

In the first part of the experiment, you'll face 21 rounds of decision-making. Each round proceeds as follows:

Two boxes are shown to you. Each box contains 10 balls. The combination of BLACK and WHITE balls can be different in each round. You can see an example below.



Then, a fair coin is flipped. If the coin is **Tail**, one ball is randomly drawn from **box T (Tail)**. If the coin is **Head**, one ball is randomly drawn from **box H (Head)**. You do not observe the coin, but you see the ball. For example if the ball is black, you see the following:



You then guess which box the ball has been drawn from. Specifically, you will answer the following question:

Please guess which box the ball has been drawn from. Enter your guess below:

☐ Box H (Head)

☐ Box T (Tail)

Remember that your answers during the experiment determine your payment at the end of the experiment. So, if you guess the correct box, you will likely receive more money at the end of the experiment. At the end of the experiment, two of your guesses will be selected at random and you will receive \$12 for each correct guess.

To summarize part 1 of the experiment, you will face 21 rounds. In each round, you experience the following steps:

1. A fair coin is flipped. If the coin is Tail, one ball is randomly drawn from box T (Tail). If the coin is Head, one ball is randomly drawn from box H (Head).
2. You see the ball.
3. You guess which box the ball has been drawn from.

Note that you should consider each round as an independent experiment. There is no relation between different rounds.

SECOND PART: (Treatment *Base*)

In the second part of the experiment, you again go through 21 rounds. Each round proceeds as follows:

You are randomly connected to another participant, who is called your “NEIGHBOR”. Your neighbor is an individual who has participated in this experiment before. He/She is not present in the lab right now. You then receive information about one of the rounds that your neighbor has gone through. The setting of your neighbor’s experiment was the same as what you saw in the previous part. In each round, the neighbor was observing one ball, randomly drawn from one of the two available boxes, and then he/she was submitting a guess.

Your task is as follows: you first see the two boxes that has been shown to your neighbor. You then observe your neighbor’s guess. Finally, you guess what the correct box is in your neighbor’s experiment. Remember that your neighbor’s guesses were for money. It means that a correct guess by your neighbor was increasing his/her chances of getting money.

To summarize the second part of the experiment, you will face 21 rounds. In each round, you are randomly connected to another participant, who is called your “NEIGHBOR”. Then, you proceed as follows:

1. You see two boxes which were shown to your neighbor.
2. You see your neighbor’s guess (your neighbor’s guess was based on a ball randomly drawn from one of the boxes).
3. You guess which box is the correct choice in your neighbor’s experiment.

Please keep in mind that you perform this task in multiple rounds and you might be connected to a different person in each round. So, you are not necessarily interacting with the same person in all rounds.

SECOND PART: (Treatment *Demographics*)

In the second part of the experiment, you again go through 21 rounds. Each round proceeds as follows:

You are randomly connected to another participant, who is called your “NEIGHBOR”. Your neighbor is an individual who has participated in this experiment before. He/She is not present in the lab right now. You then receive information about one of the rounds that your neighbor has gone

through. The setting of your neighbor's experiment was the same as what you saw in the previous part. In each round, the neighbor was observing one ball, randomly drawn from one of the two available boxes, and then he/she was submitting a guess.

Your task is as follows: you first see the two boxes that has been shown to your neighbor. You then observe your neighbor's guess (you will also see some demographic information about your neighbor such as: gender, age, years of education, and whether he/she has taken any probability/Statistics course). Finally, you guess what the correct box is in your neighbor's experiment. Remember that your neighbor's guesses were for money. It means that a correct guess by your neighbor was increasing his/her chances of getting money.

To summarize the second part of the experiment, you will face 21 rounds. In each round, you are randomly connected to another participant, who is called your "NEIGHBOR". Then, you proceed as follows:

1. You see two boxes which were shown to your neighbor.
2. You see your neighbor's guess (your neighbor's guess was based on a ball randomly drawn from one of the boxes). You also observe some demographic information about your neighbor.
3. You guess which box is the correct choice in your neighbor's experiment.

Please keep in mind that you perform this task in multiple rounds and you might be connected to a different person in each round. So, you are not necessarily interacting with the same person in all rounds.

SECOND PART: (Treatment *Bot*)

In the second part of the experiment, you again go through 21 rounds. Each round proceeds as follows:

You are connected to a bot, which is called your "NEIGHBOR". The setting of the experiment is similar to what you saw in the previous part. Two boxes are shown to you and the bot. Then, a ball is randomly drawn from one of the boxes. We show the ball to the bot (not you) and let the bot guess which box the ball is drawn from.

The bot is programmed such that when it observes a white ball, it picks the box with more white balls, and when it observes a black ball, it picks the box with more black balls. You then see the bot's guess. Finally, you guess which box the ball has been drawn from.

To summarize the second part of the experiment, you will face 21 rounds. In each round, you are connected to a bot that is called your "NEIGHBOR". Then, you proceed as follows:

1. You and your neighbor (BOT) see two boxes.
2. You see your neighbor's guess (your neighbor's guess is based on a ball randomly drawn from one of the boxes).
3. You guess which box is the correct choice.

NOTE: The bot is programmed such that it picks the box with more black balls when it observes a black ball, and it picks the box with more white balls when it observes a white ball.

SECOND PART: (Treatment *Ball*)

In the second part of the experiment, you again go through 21 rounds. Each round proceeds as follows:

You are randomly connected to another participant, who is called your "NEIGHBOR". Your neighbor is an individual who has participated in this experiment before. He/She is not present in the lab right now. You then receive information about one of the rounds that your neighbor has gone through. The setting of your neighbor's experiment was the same as what you saw in the previous part. In each round, the neighbor was observing one ball, randomly drawn from one of the two available boxes, and then he/she was submitting a guess.

Your task is as follows: you first see the two boxes that has been shown to your neighbor. You then observe your neighbor's guess and the ball that has been shown to him/her. Finally, you guess what the correct box is in your neighbor's experiment. Remember that your neighbor's guesses were for money. It means that a correct guess by your neighbor was increasing his/her chances of getting money.

To summarize the second part of the experiment, you will face 21 rounds. In each round, you are

randomly connected to another participant, who is called your “NEIGHBOR”. Then, you proceed as follows:

1. You see two boxes which were shown to your neighbor.
2. You see your neighbor’s guess as well as the ball that had been shown to your neighbor.
3. You guess which box is the correct choice in your neighbor’s experiment.

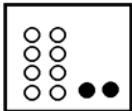
Please keep in mind that you perform this task in multiple rounds and you might be connected to a different person in each round. So, you are not necessarily interacting with the same person in all rounds.

PAYMENT:

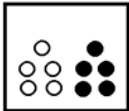
You will receive \$6 for showing up. In addition to that, two of your guesses during the experiment will be randomly selected and you get an extra \$12 for each correct guess.

Note: Your payment will be determined at the end of the experiment session after you finish all parts of the experiment. You will not know how much you earn in each part during the experiment.

Appendix A2: Experiment Interface (oTree)




H (Head)



T (Tail)

One box was selected. A ball is now randomly drawn from the selected box:



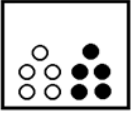
Please guess which box the ball has been drawn from. Enter your guess below:

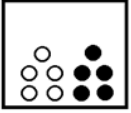
☐ Box H (Head)

☐ Box T (Tail)

With what probability do you think your answer is correct?

Figure 17: Individual Condition


H (Head)


T (Tail)

Your Neighbor's guess was Box H (Head)


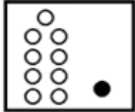
Please guess which box is the correct choice. Enter your guess below:

☐ Box H (Head)

☐ Box T (Tail)

With what probability do you think your answer is correct?

Figure 18: Social Condition (Base)



H (Head)**T (Tail)**

Your Neighbor's guess was Box H (Head)

And here is some information about your neighbor:

Age: 44

Years of Education: 12

Gender: Male

Has taken any probability/statistics course: NO

Please guess which box is the correct choice. Enter your guess below

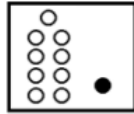
☐ Box H (Head)

☐ Box T (Tail)

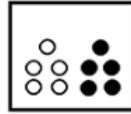
With what probability do you think your answer is correct?

Figure 19: Social Condition (Demographics)

You are now connected to your neighbor (a bot).



H (Head)



T (Tail)

Your Neighbor's guess is Box T (Tail)

Please guess which box is the correct choice. Enter your guess below:

- ☐ Box H (Head)
- ☐ Box T (Tail)

With what probability do you think your answer is correct?

Figure 20: Social Condition (Bot)

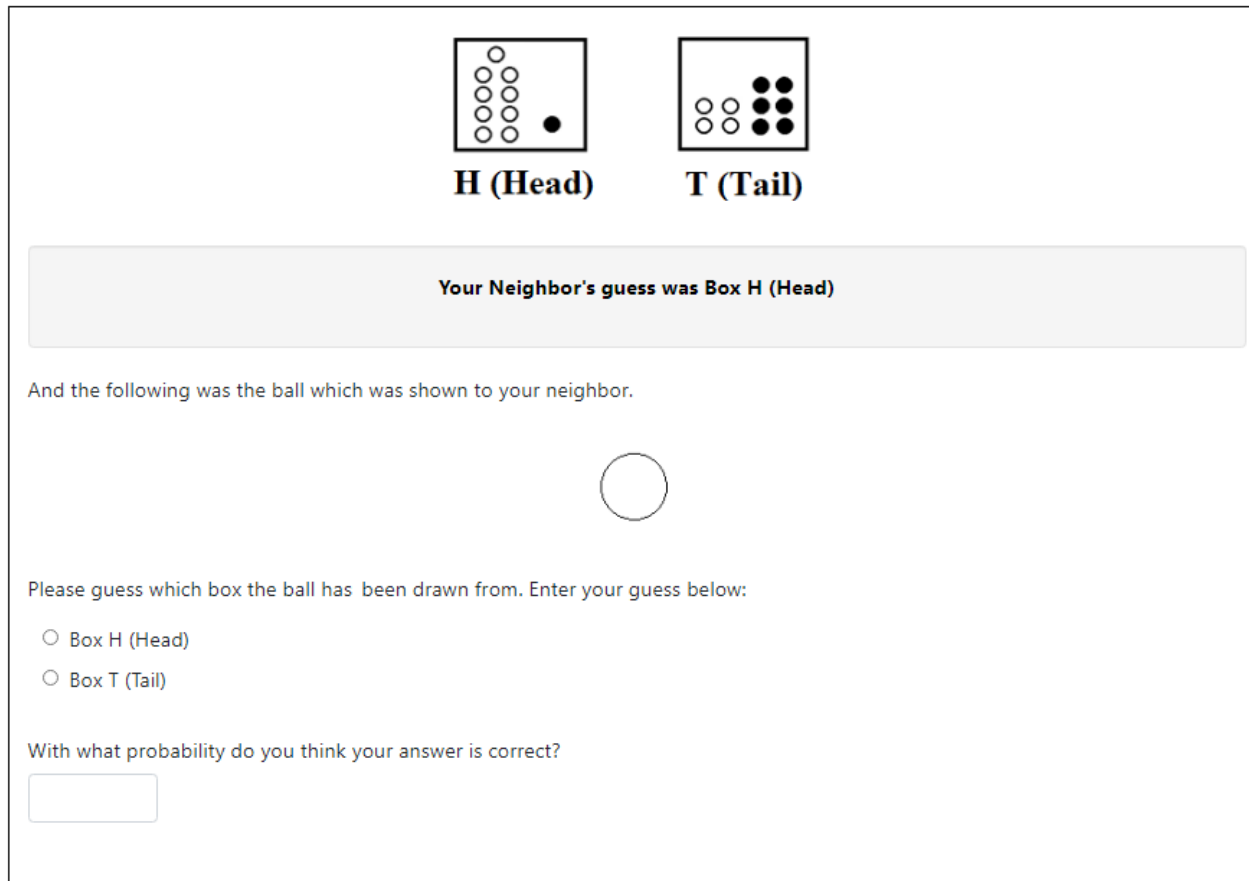


Figure 21: Social Condition (Ball)

Appendix B: Between-Subject Analysis

In the previous sections, the analysis was done on a pooled data, meaning that the data for each subject consists of decisions in both the individual and the social conditions. This means the earlier results were based on a “within-subject” analysis. In this section, I do a robustness check and examine whether the results hold if the analysis is done “between-subject”. To achieve this, I need to compare the individual irrationality of subjects who see the individual condition first, with the social irrationality of subjects who see the social condition first. The results of this between-subject analysis is shown in Figure 22. This figure verifies that the main finding holds even if the analysis is done *between-subject*.

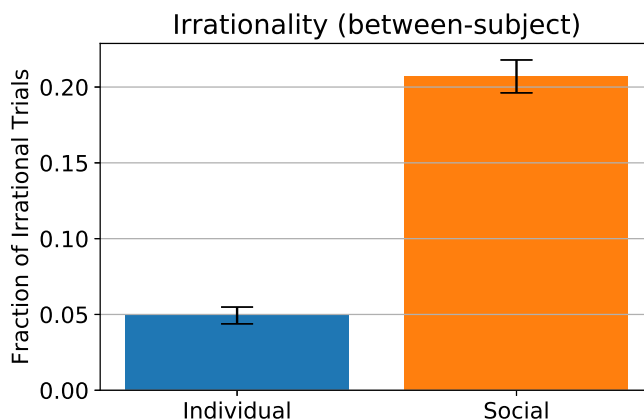


Figure 22: Individual Irrationality and Social Irrationality (between-subject analysis)

I can also redo the underlying analysis of Figure 5 at a between-subject level. The results, which are shown in Figure 23, are qualitatively similar to those presented in the body of the article.

Appendix C: A Unified Model of Individual Behavior

In this section, I present and estimate a behavioral model that corresponds to the two-stage decision-making process introduced in Section 5. The model combines two frameworks from the existing literature: for the first stage of the decision-making process, the model adapts a standard framework that was introduced in [Grether \(1980\)](#). For the second stage of the decision-making process, the model uses logistic response functions to determine choice probabilities ([McKelvey & Palfrey, 1995](#);

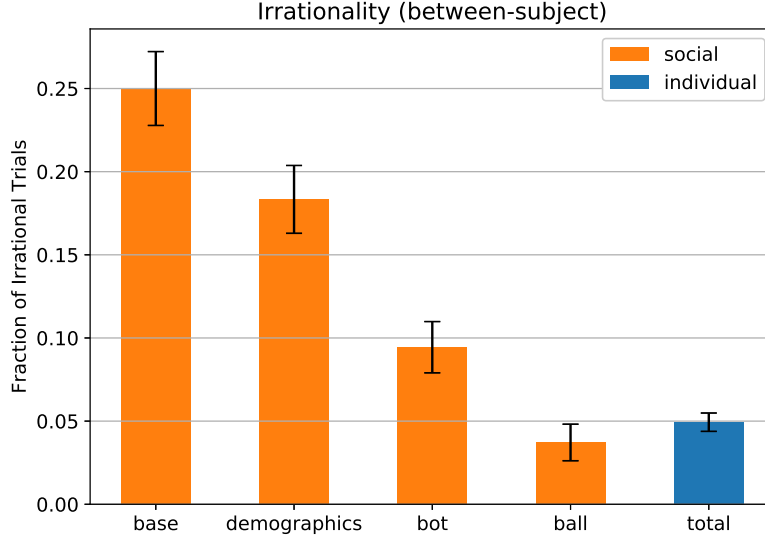


Figure 23: Social irrationality in different treatments (between-subject analysis)

Anderson & Holt, 1997). The model parameters are estimated by a two-step method. In the first step, I use the data from the individual condition to estimate the individual errors that are independent of the social condition. These estimates are then used in a second step to estimate the subject's belief about the neighbor's behavior in the social condition. This two step method uncovers the impact of the uncertainty about the neighbor's behavior on the subject's error rate in the social condition.

The Behavioral Model

My goal in here is to build a comprehensive model that incorporates the two stages of the decision-making process shown in Figure 8. For the first stage, I use the traditional framework of Grether (1980). In this framework, when an individual obtains signal s , she updates her belief as follows:

$$\pi(X|s) = \frac{p(s|X)^c p(X)}{p(s|X)^c p(X) + p(s|Y)^c p(Y)} \quad (3)$$

where $s \in \{x, y\}$, $p(\cdot)$ is the true probability, and c is a measure of bias in the posterior. Bayes rule is a special case of this equation with $c = 1$. In my experiment, priors are $p(X) = p(Y) = 0.5$. So,

the above equation can be simplified as follows,

$$\pi(X|s) = \frac{p(s|X)^c}{p(s|X)^c + p(s|Y)^c} \quad (4)$$

Here, $c < 1$ is associated with underinference, i.e., as if the signals provide less information about the state than they actually do. This implies that the more the posterior errors, the lower the c . Alternatively, $c > 1$ corresponds to overinference, signals provide more information than they actually do.

For the second stage of the decision-making process, I use a logit function (Anderson & Holt, 1997). Here, conditional on the posterior belief, $\pi(.|s)$, an individual makes a binary choice, D , with the following probability,

$$Pr(D = X|s) = \frac{1}{1 + e^{-\beta(\pi[X|s] - \pi[Y|s])U}} = \frac{1}{1 + e^{-\beta(2\pi(X|s) - 1)U}} \quad (5)$$

where β is a measure of response precision and U is the bonus of a correct guess (\$12). In this framework, the reasoning error is inversely related to β : the reasoning error diminishes when $\beta \rightarrow \infty$, and it increases as $\beta \rightarrow 0$. Note that there is a notable difference between this model and that of Anderson and Holt (1997). In Anderson and Holt (1997), the subjects are assumed to update their beliefs using the Bayes rule, but here the posterior belief, $\pi(.|s)$, is not necessarily Bayesian (e.g. it can be biased).

Figure 24 summarizes the behavioral model. In the first stage, the subject updates her belief according to equation (4). In the second step, conditional on the updated belief, the subject makes a binary choice according to equation (5).

Estimation: First Step

The first step of the estimation employs data from the individual condition of my experiment. In the individual condition, s is the color of a ball randomly drawn from the realized state. So, the objective probabilities $p(s|X)$ and $p(s|Y)$ can be easily computed from the content of each box. In addition, the subject's choice, D , and her posterior, $\pi(X|s)$, are directly observed in the data. This

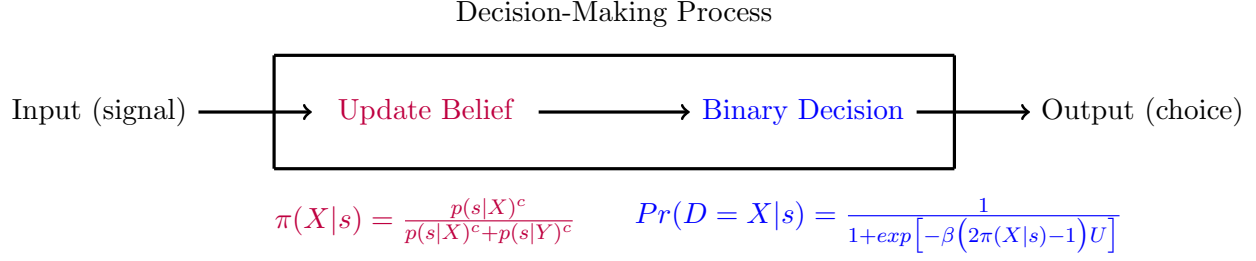


Figure 24: The behavioral model

implies that parameters c in equation (4) and β in equation (5) can be directly estimated from the data in the individual condition.³²

The estimation results are reported in Table 5. Column (1) shows that the estimated value for parameter c is significantly lower than 1,³³ meaning that in the individual condition, subjects infer less from the signal than they should. This evidence is consistent with the fact that subjects make posterior errors in the first stage of the decision-making process. Column (2) presents the estimated value for β . The small value of this parameter speaks to the fact that subjects make reasoning errors in the second stage of the decision-making process.

Table 5: The First Step Estimation Results Using the Pooled Data of the Individual condition

	OLS	Logit:True Posterior	Logit:Bayes Posterior
	(1)	(2)	(3)
c	0.888*** (0.008)		
β		0.472*** (0.018)	0.373*** (0.012)
Observations	3171	3171	3171
R ²	0.778		
Pseudo R ²		0.519	0.647

*** $p < 0.01$

³²For the parameter c , I first rearrange equation (4) as a linear regression, $\ln\left(\frac{\pi(X|s)}{\pi(Y|s)}\right) = c \ln\left(\frac{p(s|X)}{p(s|Y)}\right)$, and then use OLS for the estimation.

³³The 95% confidence interval for c is (0.872,0.903)

Column (3) in Table 2 presents the estimated value for β conditional on the assumption that posteriors in equation (5) are Bayesian. This is the conventional assumption in the social learning experiments (Anderson & Holt, 1997; Kübler & Weizsäcker, 2004). Here, the estimate for β is 0.373, significantly lower than the estimated value in the second column. This comparison suggests that the posterior error is an important factor in the individual decision-making process and not accounting for posterior errors, biased beliefs in the logit function, can lead to biased estimates for β (overestimation of the reasoning error).

Estimation: Second Step

The goal in the second step of the estimation is to use the estimates from the first step, $\hat{c} = 0.888$ and $\hat{\beta} = 0.472$, to estimate the subject's belief about the neighbor's error rate in the social condition. Here, the estimation only employs data from the social condition of my experiment.³⁴

Note that the decision-making process in the social condition still follows the framework in Figure 24, with one exception: here, the subject observes the guess of a neighbor as signal. Denote the neighbor's guess by $'s' \in \{'x', 'y'\}$ and the neighbor's private signal by $s \in \{x, y\}$. The subject's posterior after observing the neighbor's guess can be derived as follows,

$$\pi(X|'s') = \frac{p('s'|X)^c}{p('s'|X)^c + p('s'|Y)^c} = \frac{p('s'|X)^{0.888}}{p('s'|X)^{0.888} + p('s'|Y)^{0.888}} \quad (6)$$

The parameter c in this equation is a structural parameter and is set to 0.888, the estimate from the first step estimation. However, $p('s'|X)$ and $p('s'|Y)$ in equation (6) cannot be directly computed by observing the content of boxes. These probabilities depend on the subject's belief about the behavior of the neighbor. Let's denote the subject's belief about her neighbor's error

³⁴I exclude the data for the *Ball* treatment in this section. As noted earlier, the subject observes both the neighbor's guess and the ball that the neighbor has observed. So, the subject's belief about the neighbor's error rate is irrelevant in that treatment.

rates by a pair $(\tilde{\beta}, \tilde{c})$. Then, it is straightforward to derive the followings:

$$p('x'|X) = Pr(D_n = X|x) p(x|X) + Pr(D_n = X|y) p(y|X) \quad (7)$$

$$p('y'|X) = [1 - Pr(D_n = X|x)] p(x|X) + [1 - Pr(D_n = X|y)] p(y|X) \quad (8)$$

$$p('x'|Y) = Pr(D_n = X|x) p(x|Y) + Pr(D_n = X|y) p(y|Y) \quad (9)$$

$$p('y'|Y) = [1 - Pr(D_n = X|x)] p(x|Y) + [1 - Pr(D_n = X|y)] p(y|Y) \quad (10)$$

where $Pr(D_n = X|s)$, $s \in \{x, y\}$ is the subject's belief about the neighbor's choice probability:

$$Pr(D_n = X|s) = \frac{1}{1 + e^{-\tilde{\beta}(2\tilde{\pi}(X|s)-1)U}} \quad (11)$$

and $\tilde{\pi}(X|s) = \frac{p(s|X)^{\tilde{c}}}{p(s|X)^{\tilde{c}} + p(s|Y)^{\tilde{c}}}$. Substituting equations (7)-(10) in (6), one can derive the following expressions,

$$\ln\left(\frac{\pi(X|'x')}{1 - \pi(X|'x')}\right) = 0.888 \times \ln\left(\frac{Pr(D_n = X|x) p(x|X) + Pr(D_n = X|y) p(y|X)}{Pr(D_n = X|x) p(x|Y) + Pr(D_n = X|y) p(y|Y)}\right) \quad (12)$$

$$\ln\left(\frac{\pi(X|'y')}{1 - \pi(X|'y')}\right) = 0.888 \times \ln\left(\frac{[1 - Pr(D_n = X|x)] p(x|X) + [1 - Pr(D_n = X|y)] p(y|X)}{[1 - Pr(D_n = X|x)] p(x|Y) + [1 - Pr(D_n = X|y)] p(y|Y)}\right) \quad (13)$$

where $\pi(X|.)$ is the subject's posterior after observing the neighbor's guess. To estimate the subject's belief about the neighbor's behavior, I assume $\tilde{c} = 0.888$, i.e., the subject's belief about the neighbor's posterior bias is correct. Then, I estimate $\tilde{\beta}$ by using a non-linear least square method to fit the data to equations (12)-(13). The non-linear least square estimate for $\tilde{\beta}$ is 0.038 (*s.d.* = 0.009), which is significantly lower than the corresponding parameter for the individual reasoning error estimated in the first step, $\hat{\beta} = 0.472$ (*s.d.* = 0.018). This result suggests that the uncertainty in the social interaction causes the subjects to assign a lower response precision to their neighbor than their own. That is, subjects think their neighbors make more errors than they actually do.