Mixing Propensity and Strategic Decision Making*

Duk Gyoo Kim[†] Hee Chun Kim [‡]

September 12, 2018

Abstract

This paper examines a link between an individual's strategic thinking in beauty contest games and (possibly non-rational) decision-making patterns in a non-strategic setting. Experimental evidence shows that subjects' strategic behavior, which used to be understood as a result of (possibly limited) cognitive iterations, is closely related to non-strategic decision-making patterns. We claim that such a relationship partially explains conflicts in previous reports on the strategic behaviors observed in the laboratory. The relationship requires attention because the assumption that individuals are rational in the decision-theoretic sense can create a sizable misinterpretation of strategic behavior.

1 Introduction

A growing number of studies in economics and political science consider bounded rationality both in a non-strategic environment,[1] where a single player makes a decision under an uncertain state and in a strategic environment, where she responds to the other agents' unknown intentions and behavior. When making a decision in the non-strategic environment, individuals are often cognitively limited: They may not recognize or understand all the aspects that affect their payoffs, or they may lack the cognitive

*We thank Colin Camerer and Brian Rogers for their encouragement, and we thank PJ Healy, Kirby Nielsen, Sangwon Park, Henrik Orzen, Jörg Oechssler, and participants in the 2016 Western Economic Association International Conference, 2018 HeiKaMaX Workshop, 2018 Foundation of Risk and Uncertainty Conference, and 2018 International Conference on Game Theory for their comments. Hee Chun Kim greatly acknowledges the financial support from Economic Theory Center at Washington University in St. Louis and Korea Information Society Development Institute.

[†]Department of Economics, University of Mannheim, *d.kim@uni-mannheim.de*

[‡]Korea Information Society Development Institute, *heechunkim@kisdi.re.kr*

[1]In the sense that a decision maker may use some strategies to deal with uncertainty, the non-strategic environment does not necessarily mean it involves no strategy at all. Rather, it may be understood as a single-player game.

ability to draw an ideal decision as much as they needed. Observations from the strategic environment also seem to be inconsistent with the theoretical predictions attained under the assumption of full rationality, not only because their rationality is bounded, but also because their belief about other individuals' bounded rationality varies.

The primary goal of this paper is to examine how individuals' non-strategic—and possibly non-rational—decision-making patterns over probabilistic events are related to their strategic ones. To analyze strategic observations, the main body of the literature has implicitly assumed that "individuals are rational in the decision-theoretic sense of choosing strategies that are best responses to consistent beliefs" (Crawford, 2016), which hereinafter we call decision-theoretic rationality. However, experimental work shows that when subjects are asked to make repetitive decisions under uncertainty, a significant number (more than 40%) of subjects do not make decisions that maximize their expected payoff; instead, they *match* their decision frequencies to the probability of events, which is called *probability matching* (Rubinstein, 2002; Neimark and Shuford, 1959). For example, when people are asked to play ten rounds of Matching Pennies (MP) games, and they are informed that in each game a coin will be tossed independently and the coin will land heads with a 70% chance, some of them choose Heads for seven out of the ten rounds and Tails for three out of the ten, in order to match their choice frequencies with the probability of events. To maximize the expected payoff, they should have chosen Heads for all the rounds. We introduce a broader notion of probability matching because other (seemingly non-rational) decision-making patterns could exist. For example, even if they choose Heads for all ten rounds in one set, we cannot rule out the possibility that they will choose Heads for all the rounds in seven out of ten repetitive sets and choose Tails for all the rounds in three of the other ten sets. Alternatively, even if they know that choosing Heads all the time would maximize the expected payoff, they still might want to consistently choose an outside option when one is available, believing that this would serve as a hedge. We call this broad notion of an individual's tendency in repetitive decision makings a *mixing propensity* because this tendency results in mixed choices in the same environment.[2]

We claim that when the mixing propensity of an individual is not considered, it is challenging to map the individual's strategic behaviors to her underlying belief. The

[2]Many studies provide models of preferences for randomization. Dwenger et al. (2012) provide a theory of responsibility aversion, which implies a demand for randomization. Levitt (2016) finds that randomization (coin toss) on major life decisions positively affects happiness, which might also reflect the responsibility aversion. Machina (1985) and Cerreia-Vioglio et al. (2018) consider convex preferences to account for the affinity towards randomization among equally preferred options. Although investigating why some people have such a preference for randomizing their choices is worthwhile, we want to clarify here that the primary purpose of this study is not to rationalize the mixing propensity. Rather, we take their choice patterns as a given and investigate further what we can find.

beauty contest game or its modified versions have been widely used to estimate individuals' cognitive levels and their underlying beliefs about the population. One typical beauty contest game proceeds as follows: "Each of those who participate in this game simultaneously submits a number between 0 and 100, and a huge prize goes to the person who submits the closest number that is two-thirds of the average of all the submitted numbers." The unique Nash equilibrium strategy is for everyone to submit 0, but this requires sufficiently many, if not infinite, steps of iterated dominance. If everyone randomly chooses one number between 0 and 100, the average of the submitted number would be 50; that is, the winner would be someone who submits a number close to 33. If everyone who follows the same logic submits 33, then 22 would be the winning number. If everyone picks 22, then 15 would be the winning number, and so on. That is, to reach the equilibrium strategy, one must assume that everyone could follow these logical iterations as many times as is needed, which is not always accurate because an individual's cognitive ability is often bounded. We build upon two leading theories that formalize bounded rationality in strategic thinking: the Level-k (Lk) model (Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006) and the cognitive hierarchy (CH) model (Camerer et al., 2004). Both models assume that individuals use only finite (=k) steps of iterated dominance, and such k varies by individual. One notable difference is that the Lk model assumes that individuals believe that others' cognition levels are homogeneous at Lk-*1*, while the CH model assumes that individuals believe the cognition levels are distributed over L*0* to Lk-*1*, where the population distribution is assumed to be stable. To analyze experimental observations, previous studies implicitly share an assumption that every subject is equipped with *decision-theoretic rationality*. In other words, the main body of the literature has assumed that no subject has any sort of mixing propensity, which may create sizable misinterpretation: An individual who has a certain type of mixing propensity may show homogeneous choice patterns even when she has a heterogeneous belief, while an individual who has another mixing propensity may exhibit heterogeneous choice patterns that fully reflect her heterogeneous belief, even when the best response to the belief is a probabilistic mixture of many choices. Questions that naturally followed include whether the failure to consider mixing propensity in previous studies has compromised their elicitation of the structure of beliefs, and if so, how severe it is.

To address our questions, we conducted two sets of within-subjects laboratory experiments: An ordinary decision-making (ODM) experiment that uses a modified matching pennies game, and a strategic decision-making (SDM) experiment that uses a modified beauty contest game. In a nutshell, from the ODM experiment, we identify the mixing propensity of individuals and categorize it into one of four types. Within such a mixing propensity type, observations from the SDM experiment can be analyzed more clearly to better describe the belief distribution.

Our observations are summarized as follows: First, in the ODM experiment, about half of the subjects were classified as a Rational Optimizer (RO), each of who chooses rational decisions consistently; a third of the subjects were classified as a Probability Mather (PM), each of whom exhibits a probability matching propensity; and the rest of them were as a Hedging Mather (HM), each of whom consistently prefers a hedging option whose payoff is adjusted downward by their own risk preferences. There was no Uniform Matcher (UM), or one who matches the frequencies of the choice bundles to the probability of events. Second, the overall cognitive iteration level of the PM-type subject was lower than that of the RO-type subject, and the PM-type made more varied strategic decisions than the RO-type. Third, the HM-type subjects, similar to the PM-type subjects, carried over a smaller number of cognitive iterations to make a strategic decision than did the RO-type ones, but their decisions were less varied than those of PM-type subjects. Together these observations suggest that in the previous studies, the estimated population distribution of cognitive ability to carry over strategic thinking was distorted *downward* because a substantial fraction of population (the PM-type and the HM-type subjects) systematically deviate from decision-theoretic rationality. We also claim the PM-type subjects can represent their entire underlying belief structure more accurately, while the cognition level of HM-type and RO-type subjects is likely to be underestimated.

The rest of this paper is organized as follows. In the following subsection, we review related studies. Section 2 describes details of experimental design. The statistical method for inference is described in Section 3. Section 4 shows the results of the experiment and discusses its implications. Section 5 concludes.

1.1 Related Literature

This study is grounded in empirical and theoretical findings of bounded rationality in strategic behavior. We mainly focus on two leading behavioral models: the Level-k model developed by Costa-Gomes and Crawford (2006) and the Cognitive Hierarchy model developed by Camerer et al. (2004). Both models share two assumptions: (1) individuals are rational in the decision-theoretic respect that they choose strategies that are the best responses to consistent beliefs; and (2) individuals play strategies of a finite level of iterated dominance. The models differ in their assumptions about subjects' beliefs regarding the strategic behavior of other players. The Level-k (Lk) model assumes that individuals uniformly believe that all their opponents play the same level of iterated dominant strategy. For example, the $L2$ subject assumes that all his/her opponents play a one-time iterated dominant (or $L1$) strategy. From that assumption, Lk subjects are supposed to play a certain strategy that is the best response to their uniform beliefs. In Costa-Gomes and Crawford (2006), about 55% of subjects show a level of play that indicates adoption of the Level-k model. On the other hand, some subjects explicitly mix two or more different strategies, each of which represents a different level of iterated dominance. Such a systematic pattern does not coincide with the assumption of uniform belief. Costa-Gomes and Crawford (2006) claim such a mixing propensity may be the results of learning. That is, even among individuals who start from the initial uniform belief, the experience leads subjects to shift to the higher level of iterated dominance while retaining the uniform belief structure. However, for some subjects, such mixing occurs irrespective of the time horizon. These observations demand an alternative model that explains this behavioral pattern. The Cognitive Hierarchy (CH) model allows individuals to have a heterogeneous belief structure. For example, the L2 subject assumes that his or her opponent plays both the L1 strategy and the L0 strategy; the latter is a uniform random strategy. Depending on his or her belief regarding the proportion of those who use the two different strategies, each subject may find a different best response. To explicitly estimate the structure of belief, Camerer et al. (2004) use observations from previous studies as well as their own experimental observations. However, even though their model allows for a heterogeneous belief structure, Camerer et al. (2004) cannot fully explain the observations of mixed choices. That is, if a subject has a heterogeneous belief about the other players, consistently choosing the interim choice which is the best response to the heterogeneous belief as a whole can be strictly better than mixing several choices, each of these respectively corresponds to the best response to a part of the heterogeneous belief.

In the sense that we try to understand higher order rationality better, our goal is consistent with that of Kneeland (2015), who proposes a more explicit design of experiments to identify the higher order rationality. Rather than adopting Kneeland's ring games of many (more than three) players, we stick to the two-person guessing games. Because our primary objective is to find relationships between the decision-making patterns in non-strategic environments (a player vs. random events) and choices in strategic environments (a player vs. another player), we design the two experiments to be as structurally similar to one another as possible. The 11-20 game introduced by Arad and Rubinstein (2012) is also an excellent tool for eliciting higher order rationality. We did not conduct the 11-20 game because it might not capture the behavior of HM-type players who prefer suboptimal but safer options.

We posit that individuals may show different responses to the same belief, and this difference in decision-theoretic rationality may lead to the apparent puzzle that mixes different strategies. Examples abound. In Rubinstein (2002), about half of the undergraduate subjects matched their frequency of choices to the probability of events for repetitive decision-making tasks. Thaler (2016) reports a similar result among MBA students at a top university. Though the contexts varied, the fundamental question that the authors asked subjects to perform was the independent repetition of the MP game described above.[3] Likewise, many studies in the psychology literature find a significant propensity for mixing different strategies. Neimark and Shuford (1959) and Vulkan (2000) also provide lab-experiment observations that support the existence of probability matching behavior. If we regard this mixing propensity as preferences for the randomization of choices, the experimental evidence expands. Agranov and Ortoleva (2017) found that a large majority of experiment participants exhibit stochastic choice when they are asked to answer the same questions several times in a row. Dwenger et al. (2012) reported that university applications in Germany exhibit a choice pattern that is consistent with a preference for randomness. If similar probability matching behavior can also occur in strategic situations, then the underlying belief structure about the other players' cognition levels could be better revealed by the mixing strategies of different levels.

Although probability matching has been well documented in the literature, few experimental studies have explicitly considered these behavioral patterns in the optimization process for identifying underlying belief structure in the strategic decision-making environment. Georganas et al. (2015) examined whether individuals show similar levels of iterated dominance in different forms of the game. Georganas et al. (2015) had several individuals play different games: four separate non-strategic tests and a strategic decision-making session. In the strategic decision-making session, subjects played the 'undercutting game' and the 'beauty-contest game' for four and five times, respectively. While the undercutting game only allowed discrete choices, the beauty contest allowed some interim choices that do not represent any level of iterated dominance. Even in the two games that share a similar structure (requiring players to exploit an iterated dominant strategy), individuals showed almost no correlation between the levels of iterated dominance. Moreover, there was no significant connection between an individual's traits, such as IQ, and his or her level of iterated dominance. Georganas et al. (2015) attempted to find consistency in the strategic process in different environments but did not examine

[3] Rubinstein (2002) performed the "catch the messenger" game in which a detective's task is to determine the location of a video camera each day and identify as many unknown messengers as possible while knowing the probability of catching the messenger at each location. The video camera should have been installed all the time at the location where the probability is the highest, but only a small portion of students always played the stochastic dominant action. Thaler (2016) asked MBA students to make a streak of 5 matching-pennies choices. Each choice was either Heads or Tails, and a fair coin was tossed five times after they made choices. The payoff of matching at Tails was 1.5 times higher. They should have chosen Tails all the time, which is the stochastically dominant action, but the most common observation was three Tails and two Heads, matching the ratio of the payoffs.

the process regarding individual optimization patterns.

2 Experimental Design

We use a within-subject design. The same subjects participated in two different experiments: In the ordinary decision-making (ODM) experiment, subjects made a streak of decisions in which payoffs depend on unknown but realized events. In the strategic decision-making (SDM) experiment, subjects made a streak of decisions in which payoffs depend on the randomly matched subject's decisions.

2.1 Overview

In the ODM experiment, as illustrated in Table 1, subjects repeatedly play by themselves modified Matching Pennies games with unknown events. The subject's options, U, M, and D in this example, are listed in the first column. The first row shows events and probabilities; in this example, (L=3/4, R=1/4) means that event L will be realized with probability 3/4, and event R with probability 1/4. The matrix shows the subject's payoff. For example, if she chooses M and event L is randomly drawn, she earns $(3-v)/4$ points, where v is an individually-measured parameter that makes M as favorable as certainty equivalence. Their risk preferences were individually measured prior to the experiment.

Game 1	L= 3/4	R= 1/4
U	1	0
M	$3(1-v)/4$	$(1-v)/4$
D	0	1

Table 1: An Example of the Payoff Matrix in the ODM Experiment

The ODM experiment consists of four separate games, and each game consists of four sets. Each set also consists of four rounds. Therefore, subjects make decisions for a total of 64 rounds. A new event is drawn from the known probability distribution at the beginning of each set (four rounds). Subjects are informed that an event (either L or R in the example) is realized, and that event will not be changed within a set, but they do not know which event is realized. In other words, subjects face the same but unknown event for four rounds. After that, a new event is drawn, and they face another unknown event for another four rounds, and so on. Based on the subjects' choice patterns from four different games, we categorize their mixing propensities.

The SDM experiment was conducted with the same subjects who participated in the ODM experiment. In the two-player beauty contest game in Costa-Gomes and Crawford (2006), subjects earn more when they guess the match's action more accurately. This idea is maintained in our SDM experiment. Both player 1 (P1) and player 2 (P2) know the choice intervals and target parameters of P1 and P2. P1's goal is to submit a number within P1's choice interval, and P1's payoff becomes larger as the submitted number is closer to P2's number times P1's target parameter. Similarly, P2's goal is to choose a number within P2's choice interval, and P2 earns more if the number is closer to P1's number times P2's target parameter. For example, if both players choose a number between 0 and 100, and if the goal of the game is to submit a closer number of two-thirds of the other player's number, P1's range is [0,100], P1's target parameter is 2/3, P2's range is [0,100], and P2's target parameter is 2/3. There are three distinct differences between ours and the previously-conducted two-player beauty contest games. First, subjects play eight rounds of beauty contest games in five sets. In each set, they play with a new match. This setup allows us to fully examine the relationship between the individuals' strategic choice patterns and their mixing propensity type. Second, the payoff function is deliberately designed to distinguish a player's deterministic choice from a naïve random choice within an interval. Third, we introduce a calculation panel that tracks the subjects' exact thought process.

2.2 The ODM Experiment

We design the ODM experiment to identify an individual's mixing propensity type. The entire ODM experiment consists of four different Matching Pennies games, and subjects play each game repetitively. Each game consists of four sets, and each set consists of four rounds. That is, each Matching Pennies game is repeated for 16 rounds. Subjects are told that a new event is randomly drawn from the known probability distribution per each set. Subjects face the same event for four rounds within a set, and as the set changes, they will face another event for another four rounds. Since there are four different games, each subject plays 64 rounds (4 games × 4 sets × 4 rounds) during the entire experiment.

To prevent subjects from learning about the event from previous outcomes, we did not disclose the outcome of the game to the subjects during the experiment. They were informed of the realized outcome at the end of the experiment and got paid privately by the outcome. Moreover, the game with the events (or the action played by the computer player) allows us to prevent subjects from concerning about the other-regarding preferences, such as inequity aversion (Fehr and Schmidt, 1999).

Table 2 describes four Matching Pennies games used in the ODM experiment. In every round, subjects choose one of the entities listed in the first column. The event is drawn from the first row with the probability associated with each event. For example, subjects can choose one among U, M, and D in Game 1, and an event is either L with probability 3/4 or R with probability 1/4. The subject's payoff is described in the payoff matrices.

The structure of each Matching Pennies game is varied by (1) the existence of dominant actions, (2) the number of choices, and (3) the highest expected payoffs subjects can earn. Table 3 summarizes the structure.

Note that M in Games 1 and 2 and B in Games 3 and 4 are choices that subjects can earn nonzero payoffs for any event, which we call a hedging action. The discount of payoffs for the hedging action, v_i for subject i, is adopted to restrict the subject's bias toward the hedging action because of his/her risk aversion. Since the hedging action gives the same expected payoff from each single action choice and always guarantees a positive amount of payoff, risk-averse subjects may consider the hedging action to be the dominant choice as long as the payoff is greater than the certainty equivalence of the game. To exclude this concern, we measured their risk preferences before the beginning of the ODM experiment and discounted their payoff of the hedging action accordingly. Thus, the payoff of the hedging action equals the certainty equivalence of the game.[4]

From the ODM experiment, we categorize subjects into one of the four possible types, based on their choice patterns. The Rational Optimizer (RO) plays an optimal action that maximizes the expected payoff for all rounds for all sets. The Probability Matcher (PM) mixes his/her action to match the given probability within each set, and this mixing proportion is equal across the sets. The Uniform Matcher (UM) plays the same action within each set, but the proportion of sets with a single action will be equal to the given probability. The Hedging Matcher (HM) plays an 'intermediate' action for all rounds in all sets. We use the Maximum Likelihood estimation for categorization. Their behavioral patterns distinguish these four types.

Theoretical Benchmark 1. *(Mixing Propensity) Individuals with different mixing propensity show different decision-making patterns;*

1. *An RO-type player always chooses the action that maximizes the expected payoff.*

2. *A PM-type player mixes different actions within each set, and the proportion of mixing follows the probability distribution of the corresponding states.*

3. *A UM-type player chooses a single action within each set but changes the action across the sets. The proportion of the mixing follows the probability distribution of the corresponding states.*

[4]We follow Holt and Laury (2002) to measure the risk preferences.

Game 1	L = 3/4	R = 1/4
U	1	0
M	$3(1-v_i)/4$	$(1-v_i)/4$
D	0	1

Game 2	L = 1/2	C = 1/4	R = 1/4
U	1	0	0
M	0	1	0
D	0	0	1
B	$(1-v_i)/2$	$(1-v_i)/4$	$(1-v_i)/4$

Game 3	L =1/2	R = 1/2
U	1	0
M	$(1-v_i)/2$	$(1-v_i)/2$
D	0	1

Game 4	L = 1/4	LC = 1/4	RC = 1/4	R = 1/4
U	1	0	0	0
MU	0	1	0	0
MD	0	0	1	0
D	0	0	0	1
B	$(1-v_i)/4$	$(1-v_i)/4$	$(1-v_i)/4$	$(1-v_i)/4$

Table 2: Matching Pennies Games in the ODM

Each subject plays all four games in random order. The discount for a hedging behavior, v_i (M in Games 1 and 2 and B in Games 3 and 4), varies according to the subject's risk preferences, which are measured by a survey beforehand.

	Existence of Dom. actions?	The number of states	The highest expected payoff
Game 1	Y	2	3/4
Game 2	N	2	1/2
Game 3	Y	3	1/2
Game 4	N	4	1/4

Table 3: Comparisons of Four Matching Pennies Games

4. *An HM-type player always chooses a hedging action that provides a positive payoff in all cases.*

Table 4 shows possible choice patterns of each type in Game 1.

Game 1	Set 1	Set 2	Set 3	Set 4
RO	U4	U4	U4	U4
PM	U3D1	U3D1	U3D1	U3D1
UM	U4	D4	U4	U4
HM	M4	M4	M4	M4

Table 4: Predicted Behavior of Four Types in Game 1

This table shows how a player with a certain type of mixing propensity will choose actions. When a player is expected to choose an action A∈ {U,M,D} for n times, it is denoted by An. The RO-type subject will play U, the choice that gives the largest expected payoff, all of the time. The PM-type subject will match the frequency of her choices with the probability of events. Thus, in each set of four rounds, she will mix three Us and one D, up to permutation. The UM-type subject will play the same action within a set, but she will match the frequency of her choice blocks with the probability of events. Three sets of Us and one set of Ds will be chosen, up to permutation. The HM-type subject will choose H all of the time.

The RO-type subject is expected to play action U all the time because U maximizes the expected payoff. The PM-type subject is expected to play action U three times in each set because the PM-type is expected to mix his/her play to match the frequency of each choice with the given probability within each set. The UM-type subject is expected to play the same action, either U or D, within each set, but the proportion of sets that each action is played will be matched to the given probability. The HM-type subject is expected to play the intermediate action M all of the time.

2.3 The SDM Experiment

We design the SDM experiment to identify individual strategic decision-making patterns. The entire experiment consists of eight sets, and each set consists of five rounds of the two-player beauty contest game. In each set, two subjects are randomly and anonymously matched and play a game for a whole set of five rounds with the partner. That is, they repeat playing one beauty contest game for five rounds within a set. As the set changes, each subject is randomly and anonymously re-matched to another partner, with whom they play a new beauty contest game. The eight games have different structures concerning the choice intervals and target parameters.

As in the ODM experiment, the realized outcome from the subjects' choices was not disclosed during the experiment. That is, subjects played the game without feedback, and the outcome of their choices was revealed only at the end of the completed SDM experiment. This restriction prevented subjects from learning retrospectively or through experience. In each round, subjects earned payoffs according to a "payoff function,"[5] and the monetary compensation was paid based on the sum of payoffs at the end of the experiment. Subjects were informed about every detail of the SDM experiment.

We used eight different two-player beauty contest games, in which choice intervals and target parameters varied. The form of each game is described by four factors: a choice interval and a target parameter for both player 1 and player 2. For notational simplicity, α denotes a choice interval $[100,500]$, β denotes $[100,900]$, δ denotes $[300,900]$, and γ denotes $[300,500]$. A target parameter is denoted as $n_1 = 0.5$, $n_2 = 0.7$, $n_3 = 1.1$, and $n_4 = 1.5$ respectively. Moreover, we designed the games to vary regarding the number of iterations needed to arrive at the Nash equilibrium choice, the pattern of iterated strategies, and the location of the Nash equilibrium choice. Thus, always choosing the largest or the smallest number in the choice interval does not maximize a subject's payoffs, and subjects were informed of this at the instruction stage. Combined with the no-feedback policy, such variations prevented subjects from routinizing their strategic/behavioral choice patterns. Thus, we can expect that subjects would concentrate on their own strategy and belief to maximize their payoff in each game. Table 5 summarizes the details of the structure of games.

We designed the eight games to be paired into four pairs so that each subject can be assigned to play the two different positions in the same game. For example, the game labeled $\alpha n_2 \beta n_4$ and the game labeled $\beta n_4 \alpha n_2$ are paired, so player 1 in game $\alpha n_2 \beta n_4$ plays the exactly same role as player 2 in game $\beta n_4 \alpha n_2$. Similarly, player 1 in game $\beta n_4 \alpha n_2$ plays exactly the same role as player 2 in game $\alpha n_2 \beta n_4$. Subjects were also

[5]To minimize concerns about miscalculation and heterogeneity in comprehensibility, we did not give subjects the exact functional form. Details are provided in Section 2.3.1.

informed about this feature of the games. To minimize experience-based learning, we shuffled the order of the eight games without allowing two paired games to be played consecutively.

Game	Target Structure	#Iterations	Pattern of Iterations	End with Dominance
$\alpha n_2 \beta n_4$	Mix/High	17	A	Y
$\beta n_4 \alpha n_2$	Mix/High	18	A	N
$\delta n_3 \beta n_1$	Mix/Low	4	A	Y
$\beta n_1 \delta n_3$	Mix/Low	5	A	N
$\beta n_1 \beta n_2$	Low	4	S	Y
$\beta n_2 \beta n_1$	Low	4	S	Y
$\delta n_3 \gamma n_3$	High	2	A	N
$\gamma n_3 \delta n_3$	High	2	A	Y

Table 5: Forms of the Beauty Contest Games

The form of each game is described by four factors: a choice interval and a target number for both player 1 and player 2. α denotes a choice interval [100,500], β denotes [100,900], δ denotes [300,900], and γ denotes [300,500]. A target number is denoted as $n_1 = 0.5$, $n_2 = 0.7$, $n_3 = 1.1$, and $n_4 = 1.5$. Target Structure describes the pair of target numbers. 'High' (respectively, 'Low') describes the case in which both target numbers (for example, n_2 and n_4 in $\alpha n_2 \beta n_4$) are greater (resp., smaller) than 1. 'Mix/High' (respectively, 'Mix/Low') describes the case in which one number is greater than 1, and the other one is smaller than 1, and the multiplication of two numbers is greater (resp. smaller) than 1. #Iterations describes how many steps of the iterated dominance are required to arrive at the Nash equilibrium. Pattern of Iterations describes whether the number that corresponds to each step of the iterated dominance is monotone increasing/decreasing (denoted as 'S') or oscillates (denoted as 'A'). When the Nash equilibrium strategy is to choose the boundary of the choice interval, the strategy coincides with the iterated dominance.

The eight games differ in various respects. The main difference derives from the target structure, which describes the pair of target numbers. In two games, both target numbers are greater than 1. In another two games, both numbers are smaller than 1. In the remaining four games, one target number is greater than 1, and the other one is smaller than 1. In two games (resp., the other two games) among those four, the multiplication of two numbers is greater (resp., smaller) than 1. The games also differ in terms of the required number of steps using the iterated dominance to arrive at the Nash equilibrium choice. In two games, the best response is monotone increasing/decreasing in the level of the iteration of dominance, while in the other six games the best responses oscillate. In five games, the Nash equilibrium choices coincide with the boundary of the choice interval, while the Nash equilibrium choices are interior for the other three games.

2.3.1 Calculation Panel

When interpreting observations from two-player beauty contest games, the possibility of miscalculation is one of the main concerns. Previous studies consider a symmetric payoff function wherein payoff decreases with the absolute difference between the actual choice and the ideal guess (the matched player's number times the target number). Although the payoff structure is claimed to be simple, it could be problematic if the calculation ability of subjects varies. To avoid potential issues that could arise from miscalculation, we provide a calculation panel rather than asking subjects to calculate their optimal choices based on their beliefs.

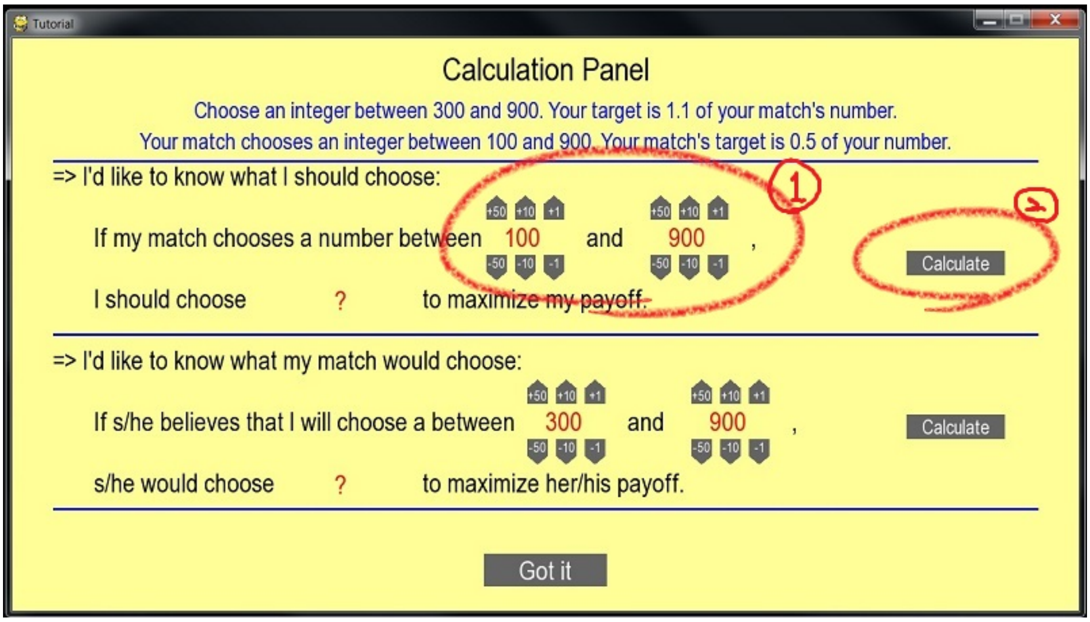


Figure 1: A Screen-shot of Calculation Panel

This calculation panel has two main purposes: to avoid potential miscalculations and to obtain richer data that will provide a better understanding of how a subject's thought process works. In the module above, a subject can calculate his/her optimal choice based on his/her belief about the opponent's choice. In the module below, a subject knows what the opponent will choose within his/her own beliefs.

A subject can use the calculation panel to find the exact number that corresponds to his best response to his prediction of his partner's choice. The calculation panel contains two modules. The module ("module A") in the upper half gives the best response to the player's prediction of his opponent's choice. The module ("module B") in the lower half provides the best response for the opponent with respect to her prediction of the opponent's prediction of the player's choice. For each module, subjects can choose an interval, with a singleton allowed. In module A, subjects can input the lower bound and the upper bound of an interval within which the opponent's choice might be placed. After that, clicking the 'Calculate' button generates the red-colored number that maximizes the player's own payoffs. The distribution over the range is fixed as the uniform distribution over the selected range, about which subjects were informed. Figure 1 considers player 1 in the game $\delta n_3 \beta n_1$. If player 1 wants to find the L1 strategy based on the belief that his/her partner's choice could be any number from 100 to 900, he/she may enter 100 for the lower bound and 900 for the upper bound and click the Calculate button. Module A will generate the result of 678 as the closest integer approximate of the best response 678.3. Subjects are informed that their use of the panel and its result will not affect their monetary outcome at the end of the session.

This approach has at least two notable merits. First, the calculation panel allows subjects to find the exact number that corresponds to the higher order of iterated dominance (Lk with $k > 1$) with no possibility of miscalculation. To know the L3 choice, player 1 needs to know in advance his/her partner's L2 choice, which would be the best response to player 1's L1 choice. Thus, player 1 can input the lower bound 100 and the upper bound 900, which correspond to his/her own range of choice to have his/her own L1 choice. After that, that player inputs the L1 choice to the range of module B (by inputting the same number in both the lower and upper bounds) in order to calculate the L2 choice, which is the opponent's best response. For example, 678 in game $\delta n_3 \beta n_1$ is the choice that player 1 will make if player 1 is of the L1 type. From this result, player 1 can input 678 in module B to obtain the result of the L2 choice of the opponent, which is 339. Then, player 1's best response can be obtained by inputting 339 in module A. Second, the usage record of the calculation panel allows us to track the individual subject's decision process. That is, having tractable records for the calculation process at each step of iterated dominance helps us identify the paths that subjects followed to reach the final decision. Even if their actual decision differs from the last calculation result, we are still able to understand how the subjects use their calculation process when making the strategic decision.

We acknowledge that the introduction of the calculation panel has its drawbacks. First, the calculation panel implicitly pushes subjects to assume either uniformly randomizing $L0$ behaviors or anchoring $L0$ behaviors. It is known that the estimation of the belief structure is sensitive to the assumption of the $L0$ behaviors, yet this calculation panel does not accommodate the odd possibility that some subjects with higher-order rationality believe that $L0$ players would choose one of the discrete choices–say, a lower bound and an upper bound. Second, introducing the calculation panel still does not perfectly reveal whether subjects make random choices. If subjects choose any number without using the calculation panel for a set, we can infer that they make random choices, but no subjects made such a purely random choice. Subjects can also make any choice even after using the calculation panel, but we cannot tell whether that choice is randomly made or is derived from their own thought processes after taking into account the results of the calculation panel. However, we believe that the merits outweigh the drawbacks.

Given this structure, Table 6 shows that player 1's choices correspond to each level of the iterated dominance, the Nash equilibrium, and the remaining intervals correspond to each round of the iterated deletion. L1 is the first level of the iterated dominance when the subject believes his/her opponent chooses numbers drawn from a uniform distribution over the entire interval. $L2$ (resp. $L3$) is the second (resp. the third) level of iterated dominance at which the subject considers their opponent to play the $L1$ (resp. $L2$) strategy. NE is the Nash equilibrium choice of the game. For example, consider game $\alpha n_2 \beta n_4$. The L1 strategy is to choose 419. Then the L2 strategy (the optimal strategy when the opponent chooses his/her L1 strategy) is to choose 361. Similarly, the L3 strategy is to choose 440. The remaining interval strictly undominated in the first round of the iterated dominance is [100,450].

Game	L1	L2	L3	NE	1st Round	2nd Round	3rd Round	4th Round
$\alpha n_2 \beta n_4$	419	360	440	500	100, 450	105, 500	105, 472.5	110.25, 500
$\beta n_4 \alpha n_2$	515	629	540	750	150, 750	150, 675	157.5, 750	157.5, 708.75
$\delta n_3 \beta n_1$	678	363	373	300	300, 900	300, 495	300, 495	300, 300
$\beta n_1 \delta n_3$	330	339	181	150	150, 450	150, 450	150, 247.5	150, 247.5
$\beta n_1 \beta n_2$	303	209	106	100	100, 450	100, 315	100, 157.5	100, 110.25
$\beta n_2 \beta n_1$	419	212	146	100	100, 630	100, 315	100, 220.5	100, 110.25
$\delta n_3 \gamma n_3$	463	550	550	550	330, 550	363, 550	399.3, 550	439.3, 550
$\gamma n_3 \delta n_3$	500	500	500	500	330, 500	363, 500	393.5, 500	439.3, 500

Table 6: Player 1's Strategic Choices with respect to the Iterated Dominance
This table shows theoretical predictions regarding various contingencies. The first column headed 'L1' shows the best response of player 1 when the opponent chooses any random number in the interval. The second through fourth columns, headed 'L2', 'L3', and 'NE,' respectively, show the corresponding best responses of player 1 when the opponent's cognitive iteration level varies accordingly. The fifth to last columns show the strictly undominated intervals after the corresponding rounds of the iterated dominance.

3 Results

Six sessions of laboratory experiments were conducted with 86 participants in the Missouri Social Science Experimental Lab (MISSEL) at the Washington University in St. Louis. From these participants, we obtained 62 effective subjects from those who passed screening tests in both experiments. We estimated their type by using the data from the

ODM experiment. Using the MLE method[6], we found that in all effective samples, the proportions of RO-, PM-, and HM-type subjects are 45.2%, 25.8%, and 29%, respectively (Table 7). Since no single subject exhibits choice patterns consistent with the UM type, we restrict our attention to RO-, PM-, and HM-types.

Type	Count	%
RO	28	45.2
PM	16	25.8
HM	18	29.0
UM	0	0
Sum	62	100

Table 7: Overall Distribution of Mixing Propensity in the ODM experiment
Based on the choice patterns in the ODM experiment, we categorize subjects into one of four possible types.

We relate this categorization of subjects to their behavioral pattern in the SDM experiment. Our primary hypothesis is that subjects' behavioral patterns in the ODM experiment are positively associated with behavioral patterns in the SDM experiment. We find that the following two observations are consistent with the predictions of our primary hypothesis.

1. RO-type subjects tend to show a higher level of cognitive iterations with a smaller variance of the choice distributions than PM-type subjects.

2. HM-type subjects are less likely to diversify their choices than PM-type subjects, and the variance of their choice distributions is similar to that of RO-type subjects.

In short, RO-type subjects seem to be more reflective than HM- or PM-type subjects, and PM-type subjects change their decisions according to their belief distribution regarding their match's cognitive ability. For example, if a PM-type subject believes that he plays the game with L0 player with a 70% chance and L1 player with a 30% chance, then he behaves as if he is the L1 player for 70% of all the games, and the L2 player for the remaining 30% of the games. That creates more choice variation. Meanwhile, HM-type subjects exhibit the lowest cognitive iterations of the three groups, and they seem to consistently make 'belief-weighted' choices in keeping with their own beliefs.

Result 1. RO-type subjects tend to show a higher level of cognitive iterations with a smaller variance of the choice distributions than PM-type subjects.

[6] For detailed statistical analysis, see Appendix A. Roughly speaking, provided that the subject has a specific type, we calculate each subject's "error" and find the type that yields the smallest error.

To develop a simple understanding of how many cognitive iterations each subject executed, we aggregated the subjects' choice data from the SDM experiment by their ODM type. Roughly speaking,[7] we did the following. Each individual's choices were coded as one integer between 1 and 4, where number k corresponds to k-th cognitive iterations. For the sake of simplicity we call such codes 'cognition levels.' Note that no decisions were coded as L0 because every subject used the calculation panel at least once for each set. If the choices were made for more than four cognitive iterations, including the Nash Equilibrium choices, we coded them as 4. In this manner, we found each subject's "distribution" of revealed cognition levels. Collecting individual distributions according to the ODM type, we identified the distribution of variances of cognition levels for each group.

	$E[\mu_i]$	$E[\sigma_i^2]$
RO	3.17	0.61
PM	2.71	1.03
HM	2.32	0.65

Table 8: Mean of Individual Means and Variances in Revealed Cognition Level by Type
This table shows how the choice patterns in the SDM experiment differ by the type categorized in the ODM experiment. $E[\mu_i]$ refers to the mean of individual means of revealed cognitive iterations for making decisions. $E[\sigma_i^2]$ refers to the mean of individual variances of the revealed cognitive iterations.

Table 8 shows two summary statistics for the individual distributions of cognitive iterations for each type. On average, RO-type subjects showed the cognition level of 3.71, while the mean of individual variances of the cognition level was 0.61. PM-type subjects showed the average cognition level of 2.71, with an average variance of 1.03. HM-type subjects showed the average cognition level of 2.32, with an average variance of 0.65. RO-type and HM-type subjects showed a relatively lower variance than PM-type subjects. In contrast, RO-type subjects showed a relatively higher average cognition level than other subject types. This result implies that the types of subjects categorized by the observations in the ODM experiment can be used to describe their behavioral patterns in the SDM experiment.

To gain a more detailed understanding of the difference, we conducted a comparison test of distributions for each of the pair groups.

Table 9 summarizes how decision patterns are statistically different from one another. We tested the null hypothesis that the mean and variance of the two distributions are from the same population.[8] The test results show that RO-type subjects, on average, have a higher cognition level than both PM-type and HM-type subjects, and these

[7]For a detailed description of the procedure used to obtain the aggregate observations, see Appendix A.

[8]We used the Fisher method (F-test) to measure the statistical significance of the difference between the two distributions.

	Mean Tests			Variance Tests		
Types	RO&PM	RO&HM	PM&HM	RO&PM	RO&HM	PM&HM
P-Values (one-sided)	0.012	0.004	0.11	0.044	0.439	0.024

Table 9: Test Results and p-values Between Distributions

differences are statistically significant at the 5% level (PM-type) and the 1% level (HM-type), respectively. PM-type subjects have a weakly higher cognition level than HM-type subjects, but this difference is statistically insignificant. On average, in the SDM experiment, PM-type subjects show a larger variation in their cognition level than did either RO-type subjects or HM-type subjects, and these differences are statistically significant at the 5% level. There is no statistically significant difference in the variance of the cognition level of RO-type subjects and HM-type subjects. Altogether, we found that, on average, RO-type subjects are more likely to show a higher cognition level with less variation than PM-type subjects.

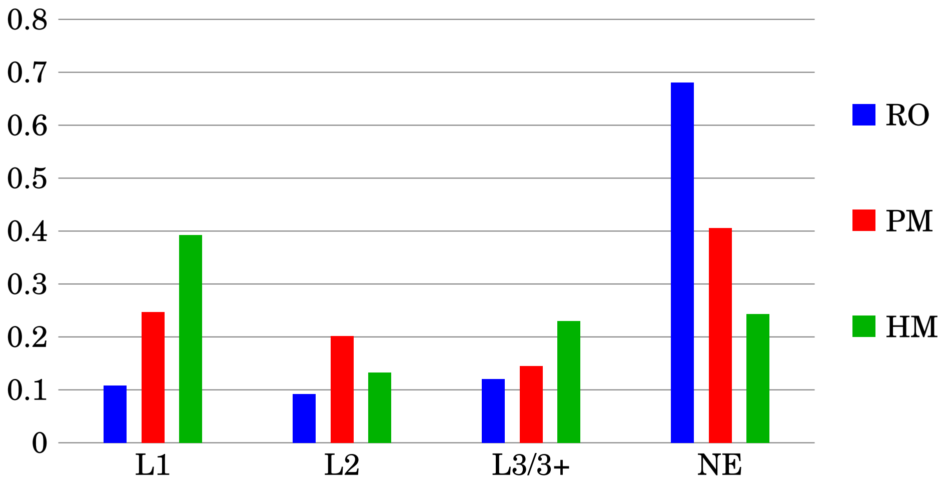


Figure 2: Proportions of Total Choices in Each Type

Figure 2 shows a proportion of choices from subjects in each type. More than 65% of choices of RO-type subjects exhibit a higher (4 or more) cognition level. The cognition levels of PM-type subjects are more dispersed than those of the RO-type. In summary, this observation supports our presumption that in the SDM experiment, PM- and RO-types inherit their behavioral pattern from the ODM experiment. RO-type subjects, who consistently chose actions that maximized the expected payoff in the ODM experiment, also showed consistent choice behavior at a certain level of cognition. PM-type subjects, who in the ODM experiment matched the frequency of their actions to a given probability of events, also spread their actions to several different cognition levels based on their beliefs about the cognitive level distributions of their opponents.

We also observe from Figure 2 that cognition levels of HM-type subjects are as dispersed as those of PM-type subjects. Our second observation is that the source of the variation is different.

Result 2. HM-type subjects are less likely to diversify their choices than PM-type subjects, and the variance of their choice distributions is similar to that of RO-type subjects.

Our second observation is that HM-type subjects are distinguished from other types in the respect that they choose a certain 'intermediate' value that may be reflected in their belief distribution regarding their match's level of cognition. In the ODM experiment, HM-type subjects chose actions that give positive, but depreciated, payoffs in all cases. We interpret their behavior as a subjective optimization that is supposed to minimize the risk of the wrong prediction. Thus, we presume that HM-type subjects may choose some belief-weighted value that could be calculated by a weighted average of the optimal choice at each cognition level. For example, suppose that a subject forms a belief about her opponent's cognition level distribution as L0: L1: L2 = 20: 30: 50 (%). In this case, although the L3 choice may give her a maximized expected payoff, she may choose the $0.2*L1+0.3*L2+0.5*L3$ choice consistently.

To acquire statistical evidence, we test whether the choice patterns of HM-type subjects differ from those of other-type subjects. Because the aggregate choice distribution of RO-type subjects is single-picked (while the other two distributions are not), we focus here on the variances of individual subject' choices. In Table 8, the average of the individual variances of RO-type subjects is 0.61, while that of HM-type subjects is 0.65. We tested the null hypothesis that the two sample distributions are from a population with the same variance. The test result (p-value of 0.439) cannot reject the null hypothesis. We also tested a hypothesis that these distributions have the same variance as the PM-type subjects. The test result rejects the hypothesis that the RO- and the HM-type group has the same mean of the variance of individual choices as the PM type at the 5% of significance level (Table 9).

In other words, both RO-type subjects and HM-type subjects, regardless of whether the choice is or is not optimal, have a tendency not to change their decisions within the same set of the SDM experiment. HM-type subjects are likely to choose an intermediate (belief-weighted) level of cognition consistently, which makes a variance of choices similar to that of RO-type subjects. Their behavioral pattern is sharply distinguished from that of PM-type subjects, who tend to change the choices within the same set of the SDM

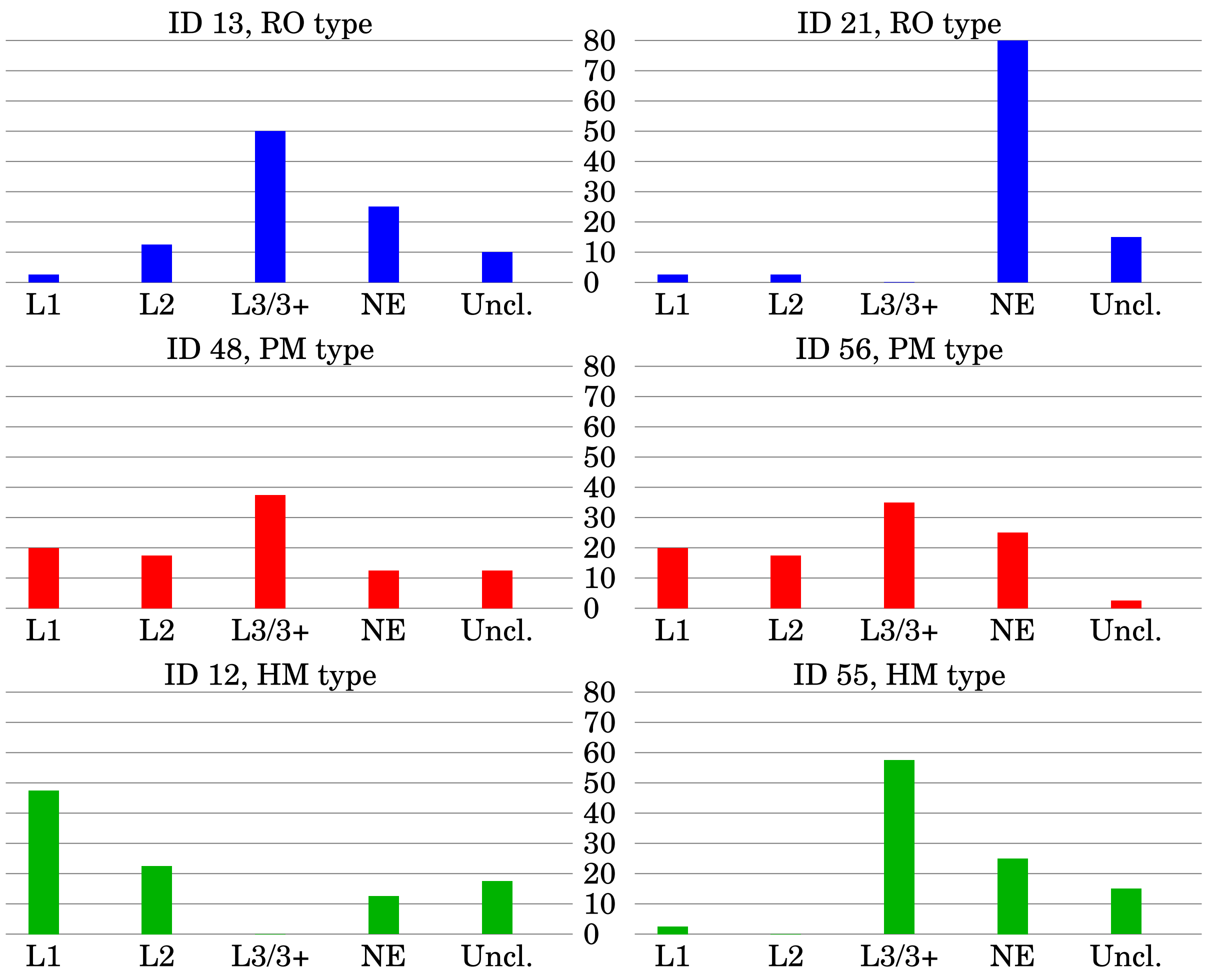


Figure 3: Examples of Individual Subjects' Choices by Type

experiment. The observed cognition level of the HM-type subjects is lower than that of the RO-type subjects. Such a behavioral pattern is consistent with our prediction regarding the HM type. Figure 3 shows the choices of individual subjects by type. We picked from each type two subjects whose choices clearly illustrate the results. As we discussed earlier, subjects 13 and 21, the RO type, mostly made decisions at the higher order rationality, and they were less diversifying the choices than PM-type subjects. Subjects 48 and 56, the PM type, diversified their choices more than the other two types. Subjects 12 and 55, the HM type, tended to stick to their decisions, but each subject's cognition level varied. Also of interest is the fact that the unclassified choices of the HM type subjects were close to the average of two cognition level choices: Subject 55 consistently chose 161 in set 6, which is 0.5*$L2$+0.5*$L3$, and subject 12 frequently chose 509, which is close to 0.5*$L1$+0.5*$L2$.

3.1 Recovery of Belief Structure

We investigate the belief structure of the PM type. From the results described above, we claimed that the behavioral patterns observed in the ODM experiment have some predictive power for the SDM experiment. This result implies that the choice distribution of the RO-type subjects, which is mostly bunched at the Nash equilibrium action, cannot fully reveal the underlying belief structure of these subjects. The only inference we can draw from this observation is that most of RO-type subjects believe that their match is most likely to choose the Nash Equilibrium action. However, PM-type subjects tend to diversify their responses in the SDM experiment, and their responses can reveal their underlying belief structure. Thus, to recover their underlying belief structure, we consider the actual responses of 16 PM-type subjects.

Level	L1	L2	L3/3+	NE	Unclassified	SUM
Count	97	79	57	159	248	640
w/ Unclassified (%)	15.2	12.3	8.9	24.8	38.8	100
w/o Unclassified (%)	24.7	20.2	14.5	40.6	-	100

Table 10: Overall Distribution of Cognition Level for the PM Type in the SDM

Table 10 and Figure 4 summarize how PM-type subjects distributed their responses at almost every level of cognition. Even though the Nash Equilibrium (NE) responses are the most frequently chosen actions, they allocated nontrivial proportions of their choices to different levels of cognition. The sum of proportions of L1 and L2 actions is larger than that of NE actions. This result implies that most of the PM-type subjects respond

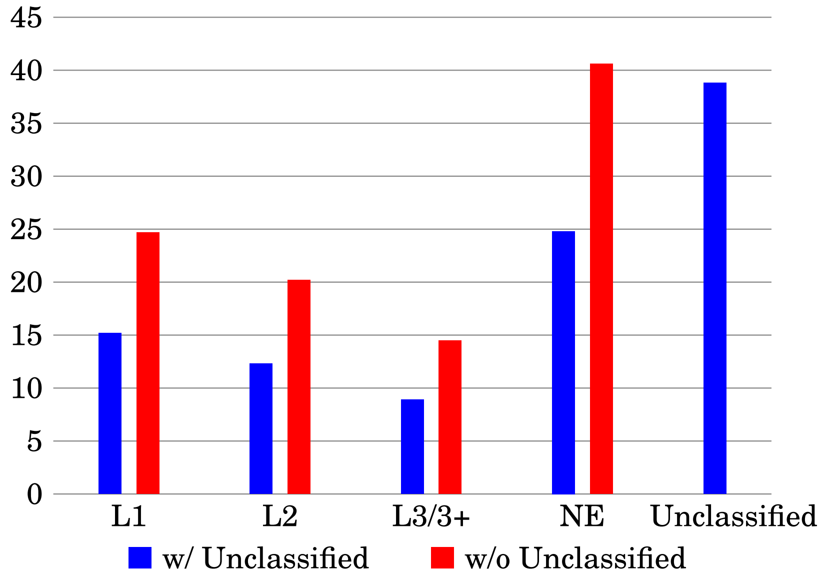


Figure 4: Proportions of Total Choices from the PM Type

to the existence of L0 or L1 subjects.

Another interesting observation is that as the cognition level increases, the proportion of the level decreases. From Table 10, proportions of L1, L2, and L3/3+ levels are 24.7%, 20.2%, and 14.5%, respectively. When we consider unclassified responses as a random-like behavior (which correspond to L0 behavior), the trend remains consistent.

The last observation is that the proportion jumps to 40.6% at the NE action. This proportion of PM-type subjects is smaller than that of RO-type subjects, but at the same time, it is too large to conclude that PM-type subjects cannot carry on necessary cognitive iterations to reach the NE action. A combination of the following could interpret this observation. First, PM-type subjects may believe that a smaller proportion of the population is equipped with a higher cognition level. Thus, they may strategically assign a smaller proportion to the higher level of cognition than RO-type subjects. Second, subjects may consider the NE action as a focal point and therefore put the highest weight. These two interpretations jointly imply that the ability of cognitive iteration of PM-type subjects is as high as that of RO-type subjects.

4 Concluding remarks

In this study, we examine how an individual's (possibly non-rational) choice patterns are related to their strategic decision-making patterns. We consider that each individual who faces a probabilistic event has a different way of making decisions, and we categorize these into three different types: the Rational Optimizer (RO), the Probability Matcher (PM), and the Hedging Matcher (HM). We found more than a half of our subjects show choice patterns other than rational optimization. Our main observation is that when asked to make strategic decisions, each type shows different decision-making patterns. While RO-type subjects focused more on the Nash equilibrium action, PM- and HM-type subjects choose their actions in response to lower cognition levels. PM-type subjects, in particular, diversify their actions to multiple levels of cognition in the SDM experiment as they diversify their decisions in the ODM experiment.

Assuming that PM-type subjects match the frequency of strategic choices to their belief distribution about the other players' cognition levels, as they did in the ODM experiment, we can recover details of their underlying belief structure. We observe that PM-type subjects played different actions within a set and assigned a lesser proportion of actions to higher levels of cognition (L1: 15.2%, L2: 12.3%, L3/3+: 8.9%). This result suggests that subjects strategically assigned their actions by their underlying beliefs. Moreover, we observe that PM-type subjects also assigned a higher proportion (24.8%) to the Nash equilibrium action. This finding supports the results of Camerer et al. (2004), who claim that subjects regard the Nash equilibrium as a focal point. Summarizing these observations, we conclude that subjects do have heterogeneous belief structure, and it consists of several different levels of cognition, but we still consider the Nash equilibrium as a plausible focal point.

The relationship between the decision-making patterns in the ODM experiment and the SDM experiment suggest that the literature may have underestimated the belief of bounded rationality. If every subject were the PM type, the belief structure estimated by the Level-k theory must be downward biased. If every subject were either the RO type or the HM type, the belief structure estimated by the Cognitive Hierarchy model would underestimate the variance of the distribution.

References

Agranov, Marina and Pietro Ortoleva, "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 2017, *125* (1), 40–68.

Arad, Ayala and Ariel Rubinstein, "The 11-20 Money Request Game: A Level-k Reasoning Study," *American Economic Review*, 2012, *102* (7), 3561–3573.

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong, "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

Cerreia-Vioglio, Simone, David Dillenberger, Pietro Ortoleva, and Gil Riella, "Deliberately Stochastic," Working Paper, National Bureau of Economic Research 2018.

Costa-Gomes, Miguel A. and Vincent P. Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 2006, *96* (5), 1737–1768.

Costa-Gomes, Miguel, Vincent P Crawford, and Bruno Broseta, "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 2001, *69* (5), 1193–1235.

Crawford, Vincent P., "New Directions for Modelling Strategic Behavior: Game-Theoretic Models of Communication, Coordination, and Cooperation in Economic Relationships," *Journal of Economic Perspectives*, November 2016, *30* (4), 131–150.

Dwenger, Nadja, Dorothea Kübler, and Georg Weizsäcker, "Flipping a Coin: Theory and Evidence," 2012. Mimeo, Available at SSRN: http://ssrn.com/abstract=2192422.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 1999, pp. 817–868.

Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber, "On the persistence of strategic sophistication," *Journal of Economic Theory*, 2015, *159* (Part A), 369–400.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, December 2002, *92* (5), 1644–1655.

Kneeland, Terri, "Identifying Higher-Order Rationality," *Econometrica*, 2015, *83* (5), 2065–2079.

Levitt, Steven D., "Heads or Tails: The Impact of a Coin Toss on Major Life Decisions and Subsequent Happiness," Working Paper 22487, National Bureau of Economic Research August 2016.

Machina, Mark J., "Stochastic Choice Functions Generated From Deterministic Preferences Over Lotteries," *The Economic Journal*, 1985, *95* (379), 575–594.

Neimark, Edith D. and Emir H. Shuford, "Comparison of predictions and estimates in a probability learning situation.," *Journal of Experimental Psychology*, 1959, *Vol 57(5)*, 294–298.

Rubinstein, Ariel, "Irrational diversification in multiple decision problems," *European Economic Review*, 2002, *46* (8), 1369–1378.

Thaler, Richard H., "Behavioral Economics: Past, Present, and Future," *American Economic Review*, July 2016, *106* (7), 1577–1600.

Vulkan, Nir, "An Economist's Perspective on Probability Matching," *Journal of Economic Surveys*, 2000, *14* (1), 101–118.

Appendix A. Statistical Model Specification

A.1. ODM Model Specification

In the ODM experiment, subjects make one choice in each round, yielding 64 choices in total. According to the choices, we categorize subjects' choice patterns into four types: RO (the Rational Optimizer), PM (the Probability Matcher), UM (the Uniform Matcher), and HM (the Hedging Matcher). We adopt the maximum likelihood method.

Let $x^{i,k}_{g,j} \in \{0,1,2,3,4\}$ denote the number of subject i's decisions that are identical to the type k subject's decision in game g of set j. We define $k \in K = \{RO, PM, UM, HM\}$ and $g, j = 1,2,3,4$. We similarly define a vector $x^{ik}_{g} = (x^{i,k}_{g,1}, \cdots, x^{i,k}_{g,4})$.

We define $\epsilon^k \in [0,1]$ as the type-specific rate of random choice that is independently and identically distributed. c_g is the number of actions that each subject has in game g. That is, $c_1 = c_2 = 3$, $c_3 = 4$, and $c_4 = 5$. Since ϵ^k is assumed to be type-specific and identically distributed over all the choices, we formulate the probability that the subject of the type k makes some predicted decisions in game g as $1-\epsilon^k+\epsilon^k/c_g = 1-(c_g-1)\cdot\epsilon^k/c_g$.[9] Then, $L^{i,k}_g(\epsilon^k \,|\, x^{i,k}_g)$ is the probability of observing $x^{i,k}_g$ when the subject i is of type k:

$$L^{i,k}_g(\epsilon^k|x^{i,k}_g) = \prod_{j=1}^{4}\left[1-(c_g-1)\cdot\epsilon^k/c_g\right]^{x^{i,k}_{g,j}} \times \left[\epsilon^k/c_g\right]^{4-x^{i,k}_{g,j}}.$$

Similarly, we define $\hat{x}^{i,k}_g$ as the number of sets of subject i's decision that equals type k's decision in game g. That is, $\hat{x}^{i,k}_g$ counts the number of sets in each vector $x^{i,k}_g$ such that each set has exactly the same number of the type k subject's decision. For example, consider a RO-type subject who chooses the same action in 3 sets (say, set 1, 2, and 3) and mixes two actions in another set (set 4). Then, $\hat{x}^{i,k}_g = 3$ since the number of sets that equals to the RO-type subject's decision is 3 (set 1, 2, and 3). With the similar notion,

[9]For example, suppose that a subject is an RO. With $\epsilon^{RO} = 0$, or if she does not make any mistakes, she will choose the action that maximizes the expected payoff with a probability of one. If $\epsilon^{RO} = 1$, that is, if she makes a choice in a completely random manner, then the probability of the optimal choice is $1/c_g$.

we define a vector $\hat{x}^{i,k} = (\hat{x}_1^{i,k}, \cdots, \hat{x}_4^{i,k})$. Then, we define $\hat{L}^{i,k}(\epsilon^k | \hat{x}^{i,k})$ as the probability of observing $\hat{x}^{i,k}$ when the subject i is of type k:

$$\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k}) = \prod_{g=1}^{4}\left[L_g^{i,k}(\epsilon^k|x_g^{i,k})\right]^{\hat{x}_g^{i,k}} \times \left[1-L_g^{i,k}(\epsilon^k|x_g^{i,k})\right]^{4-\hat{x}_g^{i,k}}.$$

Next, we define $z^{i,k}$ as a type indicator for subject i, where $z^{i,k} = 1$ if subject i is of type k and $\sum_{k\in K} z^{i,k} = 1$. From $\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k})$, subject i's maximum likelihood function can be calculated:

$$\begin{aligned} L^i(\epsilon, z^i|x^i) &= \prod_{k\in K}\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k})^{z^{i,k}} \\ &= \prod_{k\in K}\left[\prod_{g=1}^{4}\left\{L_g^{i,k}(\epsilon^k|x_g^{i,k})\right\}^{\hat{x}_g^{i,k}} \times \left\{1-L_g^{i,k}(\epsilon^k|x_g^{i,k})\right\}^{4-\hat{x}_g^{i,k}}\right]^{z^{i,k}}, \end{aligned}$$

where $\epsilon = (\epsilon^k)_{k\in K}$, $z^i = (z^{i,k})_{k\in K}$, and $x^i = (x_g^{i,k})_{k\in K}^{g=1,\cdots,4}$.

As a result, we can estimate the distribution of $z^i = (z^{i,k})_{k\in K}$, which allows us to categorize subject i's individual mixing propensity. We categorize subject i into the type (one of four) that has the highest $z^{i,k}$. The details follow.

A.1.1. Type Categorization

Each subject with a different mixing-propensity type may have a different pattern for each game. Suppose that subject i played (UUUD, UUUM, UUDD, UUUU) in game 1. Each entry of the vector corresponds to four-round actions played in each set. If she were to be the RO-type subject, then $x_1^{iRO} = (x_{1,1}^{i,RO}, \cdots, x_{1,4}^{i,RO}) = (3, 3, 2, 4)$ and $\hat{x}_1^{i,RO} = 1$. That is, any non-U actions are considered to be "wrong" actions for an RO-type subject. If she were to be the PM-type subject, $x_1^{i,PM} = (4, 3, 3, 3)$ and $\hat{x}_1^{i,RO} = 1$. We count any permutation of three Us and one D within a set as "right" actions for the PM type and deviations from it as "wrong" actions. In set 2, the subject played M instead of D (which was supposed to be chosen as the PM type), and we count three U actions as the matched actions for the PM type and M as a mismatched action. Similarly, in set 3 we count one D action as the mismatched action because the subject played D more than one time. If she were to be the UM-type subject, then she is supposed to play either four Us or four Ds in each set, so $x_1^{i,UM} = (3,3,2,4)$ gives the highest matching entries. In sets 1, 2, and 4, the action U is interpreted as a dominant action and in set 3, D is interpreted as the dominant action, so $\hat{x}_1^{i,UM} = 1$. Lastly, if she were to be the HM-type subject, $x_1^{i,HM} = (0,1,0,0)$ and $\hat{x}_1^{i,HM} = 0$. Table 11 shows examples of choice patterns of the RO-type subject and the PM-type subject.

Rational Optimizer				
Set 1-4	Game1	Game2	Game3	Game4
R1	U	U	ANY	ANY
R2	U	U	ANY	ANY
R3	U	U	ANY	ANY
R4	U	U	ANY	ANY

Probability Matcher				
Set 1-4	Game1	Game2	Game3	Game4
R1	U	U	U	U
R2	U	D	U	MU
R3	U	U	M	MD
R4	D	D	D	D

Table 11: Examples of the RO-type (left) and the PM-type (right) subject's choice patterns

In Games 3 and 4, any choice yields the same expected payoff. Thus, the rational optimizer would mix any of the choices. The choice patterns for the probability matcher are up to permutation.

In Games 1 and 2, the RO-type subject is supposed to play U for every round of every set. In Games 3 and 4, any choice is regarded as the optimal choice because all actions are expected to give the exactly the same payoff. We identify the RO-type from the other types by the pattern of plays exhibited in Games 1 and 2.

We can distinguish the PM-type subject from the other type by observing how each subject mixes the proportion of plays: The proportion must be kept across the sets and matched to the given distribution of the events. In Game 1, the PM-type subject is supposed to play U three times and D once. The order of plays is irrelevant as long as the frequency is kept to 3 Us and 1 D in each set. In Game 2, U and D are each supposed to be played two times in each set. This proportion of mixing actions is matched to the given distribution of the events ($Prob(\mathrm{L}) = Prob(\mathrm{R}) = 1/2$). In Game 3, to match the given distribution of the events, action U is supposed to be chosen two times, and M and D are supposed to be chosen one time each in every set. In Game 4, all actions (U, MU, MD, and D) are supposed to be chosen one time in each set.

Uniform Matcher				
	Game1	Game2	Game3	Game4
Set 1	U4	U4	U4	U4
Set 2	U4	U4	U4	MU4
Set 3	U4	D4	M4	MD4
Set 4	D4	D4	D4	D4

Hedging Matcher				
	Game1	Game2	Game3	Game4
Set 1	M4	M4	B4	B4
Set 2	M4	M4	B4	B4
Set 3	M4	M4	B4	B4
Set 4	M4	M4	B4	B4

Table 12: Examples of the UM-type (left) and the HM-type (right) subjects' choice patterns

Table 12 shows examples of choice patterns of the UM-type subject and the HM-type subject. The UM-type subject is supposed to play the same action within each set. Unlike the RO type, the UM-type subject may change his actions in each set so that he matches the frequency of choices to the probability distribution of the corresponding events. For example, in Game 1, the UM-type subject may choose four Us in three sets and four Ds in the other set. In Game 2, the UM-type subject may choose four Us in two sets and four Ds in the other two sets. In Game 3, U would be played in two sets and M and D would be chosen in one set each. In Game 4, one action would be played in each set. By observing such choice patterns, we can distinguish the UM-type subject from the other type.

The HM-type subject is supposed to play the hedging actions that always provide some positive payoff. In each game, such hedging action (M in Games 1 and 2 and B in Games 3 and 4) would provide a discounted payoff according to their own risk preferences, so it cannot be more beneficial than playing the rational action. Since the HM-type subject act in a homogeneous manner, we can distinguish the HM-type easily from the other types.

A.2. SDM Model Specification

A.2.1 Information and Payoff

In each round, subjects face the set of information $\{[a^1,b^1],p^1;[a^2,b^2],p^2\}$, where $[a^i,b^i]$ is the range within which player $i \in \{1,2\}$ can choose a number, and p^i is the target parameter for player i. That is, each subject completely knows his own and his partner's strategic environments. In every game, subjects are notified that they play the role of player 1 and their partner plays the role of player 2. We denote $x^i \in [a^i,b^i]$ as a choice of player i. In each round, player 1 earns payoffs that depend on his choice x^1 and player 2's choice x^2. We set the payoff function $P(x^1|x^2;[a^1,b^1],p^1,[a^2,b^2],p^2)$ as follows:

$$P(x^1|x^2;[a^1,b^1],p^1,[a^2,b^2],p^2) = 100\left[1-\frac{\left|\ln\left(\frac{x^1-p^1\cdot x^2}{|a^1-p^1\cdot b^2|}+1\right)\right|}{\ln\left(b^1-a^1+p^1(b^2-a^2)\right)}\right]$$
$$\equiv 100\left[1-\frac{\left|\ln\left(\frac{x^1-p^1\cdot x^2}{|\underline{e}^1|}+1\right)\right|}{\ln\left(\overline{e}^1-\underline{e}^1\right)}\right],$$

where $\overline{e}^1$ and $\underline{e}^1$ denote the largest and smallest possible differences between x^1 and $p^1\cdot x^2$, respectively. That is, $\overline{e}^1 \equiv b^1-p^1\cdot a^2$ and $\underline{e}_1 \equiv a^1-p^1\cdot b^2$. For notational simplicity, we denote $P(x^1|x^2;[a^1,b^1],p^1,[a^2,b^2],p^2) \equiv P^1(x^1|x^2;e^1)$, where $e^1=(\overline{e}^1,\underline{e}^1)$.

Player 1 can maximize his own payoff by minimizing the difference between x^1 and $p^1\cdot x^2$ with respect to his own prediction of x^2. Suppose that player 1 believes that his partner chooses a certain action x^2 that is uniformly drawn from an interval $[\underline{x}^2, \overline{x}^2] \subseteq [a^2, b^2]$. We denote the belief as a probability distribution $f^1(x^2|\underline{x}^2, \overline{x}^2)$. Then, the expected payoff from player 1's choice x^{1*} is given by

$$E[P^1(x^{1*}|x^2;e^1)] = \int_{\underline{x}^2}^{\overline{x}^2} \left[100 \left(1 - \frac{\left| \ln(\frac{x^{1*} - p^1 \cdot x^2}{|\underline{e}^1|} + 1) \right|}{\ln(\overline{e}^1 - \underline{e}^1)} \right) \right] f^1(x^2|\underline{x}^2, \overline{x}^2) dx^2.$$

Then, the optimal choice x^{1*} that maximizes $P^1(x^{1*}|x^2;e^1)$ satisfies the following equation:

$$(x^{1*} - p^1 \cdot \underline{x}^2 + |\underline{e}^1|)(x^{1*} - p^1 \cdot \overline{x}^2 + |\underline{e}^1|) = (|\underline{e}^1|)^2$$

The payoff function has two notable features. First, the payoff decreases in a concave manner when $|x^1 - p^1 \cdot x^2| > 0$. Second, the slope of the function when $x^1 - p^1 \cdot x^2 > 0$ is different when $x^1 - p^1 \cdot x^2 < 0$. Due to the asymmetric concavity of the payoff function, a point-wise prediction is distinguished from an interval-wise prediction. That is, the subject who uses a point-wise prediction would choose differently from someone who uses the interval prediction of which mean happens to be the point-wise prediction. For example, suppose that player 1 has $p^1 = 1.5$ and $[a^1, b^1] = [100, 900]$. Consider two cases in which (1) player 1 believes that player 2 plays exactly 300 and (2) player 1 has a belief that player 2's choice is uniformly drawn from an interval $[100, 500]$. In the former case, player 1 may choose 450 to capture $p^1 \cdot x^1 = 1.5 \times 300$. In the latter case, player 1 may choose $50(\sqrt{205} - 4) \simeq 515.89$ to best-respond to his own belief. Costa-Gomes and Crawford (2006) (henceforth CGC) adopted a kinked linear function that imposes different linear slopes at two different intervals. Even though they avoided a simple linear function, they could not distinguish the choice made by the point-wise prediction from that made by the interval-wise prediction. In CGC, two different predictions lead to the same choice of 450.[10] Imposing an asymmetric concavity on the payoff function could be useful means of avoiding those two belief structures. Second, we normalize the payoff function by using the difference between the largest possible prediction $\overline{e}_1$ and the smallest possible prediction $\underline{e}_1$. This normalization helps subjects attain similar amounts of payoffs in every game. Since each game has a different choice interval, the extent of the prediction error also can vary across games. The normalization adjusts the unit of

[10]Distinguishing the point-wise prediction from the interval-wise prediction is a sensitive issue as far as the iterated dominance model is considered. In both the Lk and the CH models, the L1 strategy could be based on the interval-wise prediction if we assume that players who adopt the L1 strategy believe that the L0 player plays randomly over the interval. On the other hand, the point-wise prediction is different in the respect that the arbitrary belief anchors on a certain point in the interval. Consequently, separating these two cases provides evidence that usefully confirms that individuals develop their predictions based on the belief that L0 players exist.

payoffs so that the extent of payoff loss from the prediction error will be measured in a less heterogeneous manner.

A.2.2 Statistical Model Specification

In the SDM experiment, we focus on the identification of the individual strategy with respect to the mixing propensity, which requires the estimation of an individual strategy and the mixing propensity type, respectively. For this reason, the estimation was conducted through a two-layer process.

In the first layer, we fix the individual mixing propensity k among four types (RO, PM, UM, and HM). Then, given the fixed individual type, we guess a type-specific strategy $s^i(k)$ of subject i with respect to such a fixed type. We allow multiple type-specific strategies. In the next subsection we discuss how we can guess the type-specific strategies from the actual data. Having obtained such a set of strategies $S^i(k)$, we then use the maximum likelihood method to estimate the probability distribution of likelihood for each strategy. From the result of the estimation, we pick the most probable type-specific strategy $s^{i*}(k)$ for type k.

In the second layer of the estimation, we collect the most probable type-specific strategies for each type k and define such set of four type-specified strategies $S^{i*} \equiv \{s^{i*}(RO), s^{i*}(PM), s^{i*}(UM)$ as a set of types for the second layer. Among those four types of S^i, we use the maximum likelihood method to estimate the most probable type.

Let $a^i_{j,l}$ and $b^i_{j,l}$ denote subject i's lower and upper bounds in round l of set j respectively. $x^i_{j,l}$ is subject i's unadjusted guess in round l of set j. Considering that $x^i_{j,l}$ may be constrained by these bounds, we define an adjusted guess $R(x^i_{j,l}) \equiv \min\{b^i_{j,l}, \max\{a^i_{j,l}, x^i_{j,l}\}\}$ that restricts the actual choice into the interval of the bounds in each round. We define a target guess for an individual with type s in round l of set j as $t^s_{j,l}$. Since our experiments allow subjects to choose integer values only, $T^{i,s}_{j,l} \equiv [t^s_{j,l} - 0.5, t^s_{j,l} + 0.5] \cap [a^i_{j,l}, b^i_{j,l}]$ is a target bound for the adjusted guess of individual i with type s in round l of set j. That is, $T^{i,s}_{j,l}$ restricts the choosable bound around the exact target $t^s_{j,l}$ with respect to a small possibility of error (± 0.5). Since we assume that all subjects can find the correct guess by using the calculation panel, we only allow a very narrow range around the exact target guess.

$\epsilon^s \in [0,1]$ is a type-specific error rate of the adjusted guess and $d^s(R(x^i_{j,l}), \lambda)$ is an error density of an individual with type s, with a precision level λ for the adjusted guess in round l of set j. For precision level λ, we assume that λ is the same across the sets and rounds. We also assume that ϵ^s is identically and independently distributed over all the rounds.

$P^i_{j,l}(x|y)$ is subject i's payoff from his own guess x, given his partner's guess y in round

l of set j. From this payoff, we define the expected payoff of an individual with type s in round l of set j as $P^{i,s}_{j,l}(x)$:

$$P^{i,s}_{j,l}(x) \equiv \int_{a^i_{j,l}}^{b^i_{j,l}} P^i_{j,l}(x|y) f^s_{j,l}(y) dy$$

where $f^s_{j,l}(y)$ is a density of y distributed according to the belief of type s.

We assume a "spike-logit" shape of error.[11] Given this assumption, $d^s(R(x^i_{j,l}), \lambda)$ is defined as :

$$d^s(R(x^i_{j,l}), \lambda) = \begin{cases} \frac{\exp[\lambda P^{i,s}_{j,l}(R(x^i_{j,l}))]}{\int_{[a^i_{j,l}, b_{j,l}] \setminus T^{i,s}_{j,l}} \exp[\lambda P^{i,s}_{j,l}(z)] dz} & \text{for } R(x^i_{j,l}) \in [a^i_{j,l}, b_{j,l}] \setminus T^{i,s}_{j,l} \\ 0 & \text{for } R(x^i_{j,l}) \in T^{i,s}_{j,l}. \end{cases}$$

We denote $n^{i,s}_j$ as the number of rounds that subject i with type s plays the exact guess of type s in set j. $N^{i,s}_j$ is a collection of such rounds in set j. We denote vectors $x^i_j \equiv (x^i_{j,1}, x^i_{j,2}, \cdots, x^i_{j,5})$ as the guesses of subject i and $R(x^i_j) \equiv (R(x^i_{j,1}), R(x^i_{j,2}), \cdots, R(x^i_{j,5}))$ as the adjusted guesses of subject i in set j.

Considering that individual i with type s chooses the (adjusted) guess $R(x^i_{j,l})$ with probability $1 - \epsilon^s$, we have a sample density for $R(x^i_{j,l})$ in set j, denoted by $d^s(R(x^i_j), \epsilon^s, \lambda)$ as follows:

$$d^s(R(x^i_j), \epsilon^s, \lambda) \equiv (1 - \epsilon^s)^{n^{i,s}_j} (\epsilon^k)^{5 - n^{i,s}_j} \prod_{l \notin N^{i,s}_j} d^s(R(x^i_{j,l}), \lambda).$$

Similarly, we define $R(x^i) \equiv (R(x^i_1), R(x^i_2), \cdots, R(x^i_8))$ as subject i's adjusted guesses for the entire experiment and $d^s(R(x^i), \epsilon^s, \lambda)$ as a sample density function for the entire experiment :

$$d^s(R(x^i), \epsilon^s, \lambda) \equiv \prod_{j=1}^{8} d^s(R(x^i_j), \epsilon^s, \lambda).$$

Now we define $z^{i,s}_j$ as an indicator of type s for subject i, where $z^{i,s}_j = 1$ if the subject is of type s and $\sum_{s \in S} z^{i,s}_j = 1$. $\epsilon \equiv (\epsilon^s)_{s \in S}$ is a vector of error rates for all types and $z^i_j \equiv (z^{i,s}_j)_{s \in S}$ is a vector of type indicators in set j. From these definitions, we have subject i's log-likelihood function $L(z^i_j, \epsilon, \lambda | R(x^i_{j,l}))$ as follows:

[11] We share the assumption of Costa-Gomes and Crawford (2006) regarding the distribution of error: Each subject is assumed to have an exact choice within her target bound (i.e., $T^{i,s}_{j,l}$) with probability $1 - \epsilon$. Outside her target bound, the distribution of error follows a logistic distribution with a precision level λ. As a result, the shape of her error rate will be like a spike (or a plateau) with a peak within her target bound. The use of the calculation panel reduces the possibility of error that purely comes from miscalculation. Moreover, through rounding, we allowed the range of exact choice to include the closest integer.

$$L(z_j^i, \epsilon, \lambda | R(x_{j,l}^i)) \equiv \sum_{s \in S} z_j^{i,s} \ln \left[d^s(R(x^i), \epsilon^s, \lambda) \right].$$

A.2.3. Strategy Elicitation by Type

We use actual choice data from the SDM experiment to guess the possible set of specific strategies per type from the ODM experiment. Unlike the ODM experiment, the SDM experiment does not provide explicit avenues of belief formation that individuals are expected to follow. The RO-type subjects, who might use a single action for the whole experiment, are relatively easy to identify. On the other hand, identifying the other types—the PM, UM, and HM types—requires that we find not only strategies that subjects use but also the mixing proportion among those strategies. This consideration requires, in theory, that we try an infinite number of different mixing proportions with different strategies. For example, a PM-type subject who adopts the L1 strategy and the L2 strategy with a mixing proportion 0.75 and 0.25 and another PM-type subject who adopts the L1 and L2 strategies with a mixing proportion 0.50 and 0.50 should be classified as different types. To avoid this difficulty, we restrict our attention to a limited set of strategies. To this end, we exploit the actual choice observations to guess the probable strategies and the mixing proportion among them by individuals.

- Rational Optimizer (RO) type

RO-type subjects always have a fixed set of type-specific strategies: RO-L1-type subjects are expected to always choose the L1 strategy, RO-L2-type subjects are expected to always choose the L2 strategy, and so on. Since we restrict our attention to only four strategies (L1, L2, L3 and NE), this assumption restricts the set of the RO-type subjects' strategies. That is, we construct the strategy set for RO-type subject i as $S^i(RO) = \{RO-L1, RO-L2, RO-L3, RO-NE\}$. From four strategies, we find the specific strategy $s^{i*}(RO) \in S^i(RO)$ that maximizes the likelihood.

In each round of the sets, RO-type subjects are supposed to play a certain action that corresponds to their optimal strategy. For example, the RO-NE-type subject would play the NE strategy in each round. While most rounds allow different strategies, Games $\delta n_3 \gamma n_3$ and $\gamma n_3 \delta n_3$ share the same numbers for different strategies. In Game $\delta n_3 \gamma n_3$, the L2 player and the NE player can play the same choice 550, and in Game $\gamma n_3 \delta n_3$, the strategy type for a subject who chooses 500 would not be distinguished. To distinguish them, we need to rely on the records from the calculation panel. In Game $\delta n_3 \gamma n_3$, an L2 player would use the calculation panel to calculate an L1 partner's choice, 500, and then, by putting 500 into his own calculation panel, he would learn that 550 is the best-responding choice. An NE player, to get 550, may start by calculating his own choice

Game	RO–L1	RO–L2	RO–L3	RO–NE
$\alpha n_2 \beta n_4$	419.4	361.1	440.3	500
$\beta n_4 \alpha n_2$	515.9	629	541.8	750
$\delta n_3 \beta n_1$	678.3	363.9	373.1	300
$\beta n_1 \delta n_3$	330.8	339.2	181.9	150
$\beta n_1 \beta n_2$	350	173.9	122.5	100
$\beta n_2 \beta n_1$	347.8	245	121.75	100
$\delta n_3 \gamma n_3$	300	550	363	550
$\gamma n_3 \delta n_3$	500	330	500	500

Table 13: Predicted patterns of play in the SDM experiment for RO-type subjects

for the L1 strategy and have 300 from the initial calculation. Also, an NE player may repeatedly use (more than 3 times) the calculation panel to arrive at the NE strategy.

(2) Probability Matcher (PM) type

To identify the strategies of PM-type subjects, we need to specify not only strategies but also the mixing proportion among the strategies. In the case of the RO type, picking a certain strategy from the fixed set is enough to allow identification of the type-specific strategy. However, PM-type subjects may use more than one strategy with respect to their own belief structure. For this reason, we need to specify both the multiple strategies they can adopt and the frequency with which those strategies are used. We do so as follows: To identify (pure) strategies, we restrict our focus to 11 different combinations of strategies. Since we assume that subjects use only four pure strategies (L1, L2, L3, and NE), PM-type subjects could have: (i) 6 different combinations of 2-strategy cases: L1 + L2, L1+ L3, L1 + NE, L2 + L3, L2 + NE, and L3 + NE; (ii) 4 different combinations of 3-strategy cases: L1 + L2 + L2, L1 + L2 + NE, L2 + L3 + NE, L1 + L3 + NE; and (iii) one combination of 4-strategy cases. For the purpose of illustration, we consider a L1 + L2 case described in Table 14. In this example, we describe hypothetical choices of a PM-type subject who uses L1 and L2 strategies and mixes them with proportion L1 : L2 = 3 : 2. Each two columns describe the actions of each round and corresponding strategies. The first column shows choices of the PM-type subject of L1 + L2 strategy. The subject can use either a pure L1 strategy or a pure L2 strategy. The next column shows the corresponding strategies for each choice. For example, a subject's action 419.4 (in the first column) in Game $\alpha n_2 \beta n_4$ of round 1 corresponds to L1 strategy (in the second column).

Probability Matcher with L1(60%)+L2(40%)										
Game	Round1	Stg.	Round2	Stg.	Round3	Stg.	Round4	Stg.	Round5	Stg.
$\alpha n_2 \beta n_4$	419.4	L1	361.1	L2	419.4	L1	361.1	L2	419.4	L1
$\beta n_4 \alpha n_2$	515.9	L1	515.9	L1	515.9	L1	521.7	L2	521.7	L2
$\delta n_3 \beta n_1$	678.3	L1	363.9	L2	363.9	L2	678.3	L1	678.3	L1
$\beta n_1 \delta n_3$	339.2	L2	330.8	L1	330.8	L1	330.8	L1	339.2	L2
$\beta n_1 \beta n_2$	350	L1	173.9	L2	173.9	L2	350	L1	350	L1
$\beta n_2 \beta n_1$	245	L2	347.8	L1	245	L2	347.8	L1	347.8	L1
$\delta n_3 \gamma n_3$	550	L2	300	L1	300	L1	300	L1	550	L2
$\gamma n_3 \delta n_3$	330	L2	500	L1	500	L1	330	L2	500	L1

Table 14: An example of a PM-type subject's pattern of play in the SDM experiment

We first identify the strategies used in each set and then find an average proportion among them. During this process, we focus on choices that correspond to L1, L2, L3, and NE strategies. Once we have an average proportion among these strategies, we round it up/down to be fitted with 5-round setting. In this way, a PM-type subject who uses different collections of strategies with different mixing proportions can also be classified. This process is grounded in the assumption that PM-type subjects would keep the same mixing proportion across the sets for the same collection of pure strategies. This assumption allows us to guess the mixing proportion from an observed average proportion of choices. From a vector $x^i \equiv (x^i_1, x^i_2, \cdots, x^i_8)$, we can find the frequency of each choice that corresponds to each strategy. Then, we use the observed frequency to guess the mixing proportion.

Specifically, we find at most eight candidates for strategies specified within each set for PM-type subjects. For example, consider subject i's choices in set 1, which show the proportion among each strategy as L1 : L2 : L3: NE = 2 : 1 : 1 : 1. We name a type-specified strategy that shows the mixing proportion L1 : L2 : L3 : NE = 2 : 1 : 1 : 1 as 'PM-1.' Then we compare the actual choices in other sets (set 2–set 8) with this PM-1 strategy. Similarly, suppose the actual choices in set 2 shows L1 : L2 : L3 : NE = 1 : 2 : 1 : 1. Then, we have another type-specified strategy named 'PM-2' that has the mixing proportion L1 : L2 : L3 : NE = 1 : 2 : 1 : 1, and so on. When we apply this approach for each result in every set, we have at most eight different candidates of mixing strategies specified for PM-type subjects.

After we construct all of the subject's possible candidates for mixing strategies, we calculated the candidate that is most likely to explain the all (at most) eight mixing strategies candidates. For example, consider the case we labeled 'PM-1' as a candidate that explains the strategies in set 2. The only difference between the choices in set 1 and those in set 2 is that the subject played L2 one additional time; this would be an L1 play if he was the PM-1 subject, which is counted as deviation from the PM-1 strategy. By repeating this, we could estimate which type-specific strategy yields the “best fit" for each PM-type subject.

(3) Uniform Matcher (UM) type

To guess the type-specific strategy for UM-type subjects, we consider the observations from the entire experiment.[12] In contrast to what we did for the PM type, we assume that UM-type subjects play the same strategy within each set and the strategy changes across sets. We can easily guess the strategies used in each set. To guess the mixing proportion, we count the frequency of the strategies from the entire experiment. That is, we only count all choices played in the experiment (from set 1 to set 5, excluding all non-RO-type actions). Naturally, the actual mixing proportion is the guess for the mixing proportion. For example, consider a UM-type subject with strategies L1+L2+NE with the mixing proportion L1 : L2 : NE = 4: 2: 2. In Table 15, the UM-type subject uniformly played the L1 strategy in sets 1, 2, 6, and 8, the L2 strategy in sets 4 and 5, and the NE strategy in set 3 and 7. From this observation, we can guess the collection of strategies adopted by the subject as L1, L2 and NE. The mixing proportion among the adopted strategies can be guessed by counting the mixing proportion in the entire experiment.

Uniform Matcher with L1(50%)+L2(25%)+NE(25%)										
Game	Round1	Stg.	Round2	Stg.	Round3	Stg.	Round4	Stg.	Round5	Stg.
$\alpha n_2 \beta n_4$	419.4	L1	419.4	L1	419.4	L1	419.4	L1	419.4	L1
$\beta n_4 \alpha n_2$	515.9	L1	515.9	L1	515.9	L1	515.9	L1	515.9	L1
$\delta n_3 \beta n_1$	300	NE	300	NE	300	NE	300	NE	300	NE
$\beta n_1 \delta n_3$	339.2	L2	339.2	L2	339.2	L2	339.2	L2	339.2	L2
$\beta n_1 \beta n_2$	173.9	L2	173.9	L2	173.9	L2	173.9	L2	173.9	L2
$\beta n_2 \beta n_1$	347.8	L1	347.8	L1	347.8	L1	347.8	L1	347.8	L1
$\delta n_3 \gamma n_3$	550	NE	550	NE	550	NE	550	NE	550	NE
$\gamma n_3 \delta n_3$	500	L1	500	L1	500	L1	500	L1	500	L1

Table 15: An example of a UM-type subject's pattern of play in the SDM experiment

For the actual guess of the possible set of specific strategies for the UM type, consider

[12] Since no subjects were categorized as the UM type, we did not employ the approach described in this section. We write this in order to complete the description.

the example of a UM-type subject with observed strategies L1 : L2 : L3 : NE = 15 : 10 : 13 : 2. From this observation, we can find the relative proportion of the strategies. However, we can imagine situations in which the observed frequency of the strategies may not clearly fit the 8-set setting or in which multiple strategies appear within the same set. For the former case, we rounded up or down for the proportions that are multiples of 1/8. For example, when we round up or down the relative proportion L1 : L2 : L3 : NE = 15 : 10 : 13 : 2 = 0.375 : 0.25 : 0.325 : 0.05, we could have several different approximations. That is, 0.375 : 0.25 : 0.325 : 0.05 can be rounded up to 0.375 : 0.25 : 0.375 : 0.125, or to 0.375 : 0.25 : 0.375 : 0. When we have such several different approximations, we consider the most frequently used strategy in each set to be the major strategy of that set. For the case that rounds up 0.325 to 0.375 = 3/8 for L1, three sets are required to uniformly play L1. Then we find three sets in which L1 appears mostly. In these sets, we count non-L1 choices as deviations from the L1 strategy. Applying this process to all the other strategies we derive the sample density $d^{UM}(R(x^i), \epsilon^{UM}, \lambda)$.

(4) Hedging Matcher (HM) type

We allow HM-type subjects to choose non-Lk choices, and this relaxation leads us to another difficulty: whether to consider such choices as a strategic hedging behavior or not. To address this concern, we exploit two assumptions: (1) Any hedging behavior should be based on the belief that is comprised of multiple Lk strategies, and (2) any hedging behavior based on a certain belief should be bounded by the interval formed from the belief. For example, consider a HM-type subject in game $\alpha n_2 \beta n_4$ who has a belief that her partner may play either the L1 or the L2 strategy with some mixing proportion. Then, her belief may form a boundary for her hedging choice, and that boundary may depend on the L1 (419.4) and L2 (361.1) strategies. Given her belief of about L1 and L2, choosing any numbers outside of the interval formed by 419.4 and 361.1 (in this case, [361.1, 419.4]) is always a weakly dominated strategy by any number in the interior of the interval. From these assumptions, we can infer that any HM-type subjects may choose the number within the interval that is bounded by Lk strategies based on her belief. This inference, although it allows a range that is broader than that for either UM- or PM-type subjects, provides grounds for determining whether the subject shows consistent HM-type behavior. For example, consider an HM-type subject who has in mind the L1 and L2 strategies. Table 16 shows that all of the choices taken by the HM-type subject are consistently located in the interval between the L1 and L2 strategies. This pattern of play can distinguish the HM-type subject from random-playing subjects.

For HM-type subjects, specifying the adopted strategies would be sufficient. In contrast to what we did for UM- or PM-type subjects, we focus on identifying HM-type sub-

Hedging Matcher with L1+L2							
Game	Round1	Round2	Round3	Round4	Round5	L1	L2
$\alpha n_2 \alpha n_4$	380	400	415	375	365	419.4	361.1
$\alpha n_4 \alpha n_2$	516	580	600	550	629	515.9	629
$\alpha n_4 \beta n_1$	400	450	650	550	600	678.3	363.9
$\beta n_1 \alpha n_4$	332	333	333	338	335	330.8	339.2
$\beta n_1 \beta n_2$	350	180	250	200	300	350	173.9
$\beta n_2 \beta n_1$	250	300	333	325	280	347.8	245
$\delta n_3 \gamma n_3$	300	550	350	500	400	300	550
$\gamma n_3 \delta n_3$	350	400	500	450	400	500	330

Table 16: An example of a HM-type subject's pattern of play in the SDM experiment

jects by itself. We consider 6 different combinations of two strategies that result in six different intervals: L1 + L2, L1 + L3, L1 + NE, L2 + L3, L2 + NE, and L3 + NE. Even though the range of target guesses is wide, we can identify the consistent choice patterns made by a HM-type subject. First, every combination of strategies has a unique formation that depends on the game structure. For example, L1 + NE may have the broadest range in most of games, but not in games 1 and 8. Therefore, even if a subject who behaves randomly may seem to consistently choose the number within the interval in the L1 and NE strategies, that subject should show consistent deviation in every round of games 1 and 8. Moreover, the shape of the interval in the game also changes. In games 1, 2 and 7, the choices from the L1 strategy are lower than those from the NE strategy, while in games 3, 4, 5, and 6 the opposite is true. From this structure, the HM-type subjects with a certain belief needs to know the exact range that will bound his/her optimal hedging choice. Moreover, the ranges covered by these combinations at most are equal to or are smaller than half of the entire choice interval. For this reason, the probability that a random-behaving subject consistently chooses numbers in a range bounded by pertinent strategies is negligible.

Appendix B. Experiment Instructions

B.1. The ODM Experiment

Preliminary Survey:

Before you start the actual task, we'd like to ask you to answer three survey questions. Consider a hypothetical situation where I give you either a fixed amount of money, X, or a simple lottery ticket. You can choose only one. The lottery ticket, denoted by (W, p), gives you money W with probability p, or nothing with probability $1-p$. For example, $(\$100, 0.4)$ is a lottery ticket that gives you \$100 with a probability 0.4, or nothing with a probability 0.6. We want you to compare lottery tickets with fixed amounts of money.

[Three questions asking certainty equivalences of (\$10,0.5), (\$1,0.3), and (\$1,0.7), respectively, are followed.]

Important Preliminary: "... randomly drawn from a probability distribution??"

We want you to understand what we mean by "an event is randomly drawn from a probability distribution." A probability distribution is a description of possible events and their chances. For example, if you toss a fair coin, with a 50% of chance it will land heads (H) and with another 50% of chance it will land tails (T). Here the possible events are the faces of a coin, H and T, and the corresponding chances are 0.5 each. Then the probability distribution of coin tossing can be described as (H,0.5; T,0.5). When we say "an event (here, the face of a coin) is randomly drawn from (H,0.5; T,0.5)," wee mean that we toss a coin and either H or T is realized, but we will not tell you what the actual realization is.

Here is another example. If we say "an event is randomly drawn from (L,0.2; C,0.5; R,0.3)," then it means three things. First, an event L (,C or R) will be drawn with a 20% (, 50% or 30%) of chance, respectively. Second, one event among L, C and R is realized according to their chances. Third, we will not tell you what the realization is.

During this experiment, you will frequently read "an event is randomly drawn from a probability distribution" in various contexts. We will assume now that you completely understand the meaning of the sentence. Please raise your hand if you need further explanation.

Description of Games:

Your task is to make choices to earn points from several games. One point is equivalent to 25 cents. The specific form for each game will vary, but it will similarly be described as the following payoff matrix.

	H, 0.3	T, 0.7
U	1	0
M	0.15	0.35
D	0	1

Your options will be shown on the left (U, M, and D, in this example). A probability distribution will be on the top (H,0.3; T,0.7). The matrix shows your payoff. For example, if you choose M and event H is randomly drawn from the probability distribution, you will earn 0.15 points. If you choose D when event T is drawn, you will earn 1 point.

Here is another example. Suppose another game is described as the following matrix.

	L, 0.5	C, 0.3	R, 0.2
U	1	0	0
M	0	1	0
D	0	0	1
B	0.25	0.15	0.1

In this example, your options are U, M, D and B, and the event will be L with a probability 0.5, C with a probability 0.3, or R with a probability 0.2. For example, if you choose B when event L is drawn, you will earn 0.25 points. If you choose U when event C is drawn, you will earn 0 points.

Structure of Experiment:

The experiment consists of 4 sessions. In each session you will play a game several times. Your available actions, a probability distribution over events, and payoffs will be informed.

Each session consists of 4 sets. In the beginning of each set, an event is randomly drawn from the probability distribution. Given this event, you will make choices for four rounds of the set. Note that the realized event will be unknown during those rounds. In the beginning of the next set, another event is newly drawn from the same probability distribution. You will make another four choices, and so on.

Since a session consists of four sets and you have four rounds per set, one session runs 16 rounds. In a new session you will repeat the process with a new game that consists of new available actions, a new probability distribution, and new payoffs. It is important for you to understand how a session and a set are defined because in each session all 16 rounds look the same. For the first 4 rounds (i.e., the first set), the realized event is the same. For the next 4 rounds (i.e., the second set), an event is newly drawn and the realized event is the same.

Screening Test:

After you read the instructions, you will answer three multiple choice questions to check your understanding of the instructions. You may want to go back to review or ask an experimenter for help. You can participate in the experiment only if ALL of your answers are correct. The main purpose of this screening test is to help you understand the instructions, not to cause you any stress. It is okay for you to ask an experimenter to help you if you are in doubt.

Q1. The experiment will consist of (A) sessions. In each session, you will have (B) sets. In each set, you will play (C) rounds of game. What are the appropriate numbers in (A), (B) and (C)?

Q2. A game is described as the following matrix. [A matrix is displayed] Which of the following is true?

Q3. Suppose a probability distribution over events is (L,0.3; C,0.5; R,0.2). Which of the following is NOT true?

B.2. The SDM Experiment

General description:

This experiment consists of 8 sets, and each set consists of 5 rounds of decision making.

In each set, you are randomly matched to one experiment participant in this lab. You will play with your match for the set (5 rounds). In the next set, you will be matched to another participant. You will not know who your matches are, and your matches will not know you either.

In each round, you choose an integer in a certain range (e.g., between 100 and 300). You earn more when your number is closer to a certain target number (e.g., 0.7) times your match's actual choice. Your match will do the same task, but his/her range and target may be different.

Example:

Suppose that in the first set you choose a number between 0 and 300, your match chooses a number between 100 and 500, your target is 0.5, and your match's target is 1.1. If your match chooses, say, 200, then your payoff is maximized when you choose 100 because your target (0.5) times the match's choice (200) is 0.5*200 = 100. On the other hand, if your match believes that you will choose 100, then she would choose 110 because her target (1.1) times your choice (100) is 1.1*100 = 110.

Payoff:

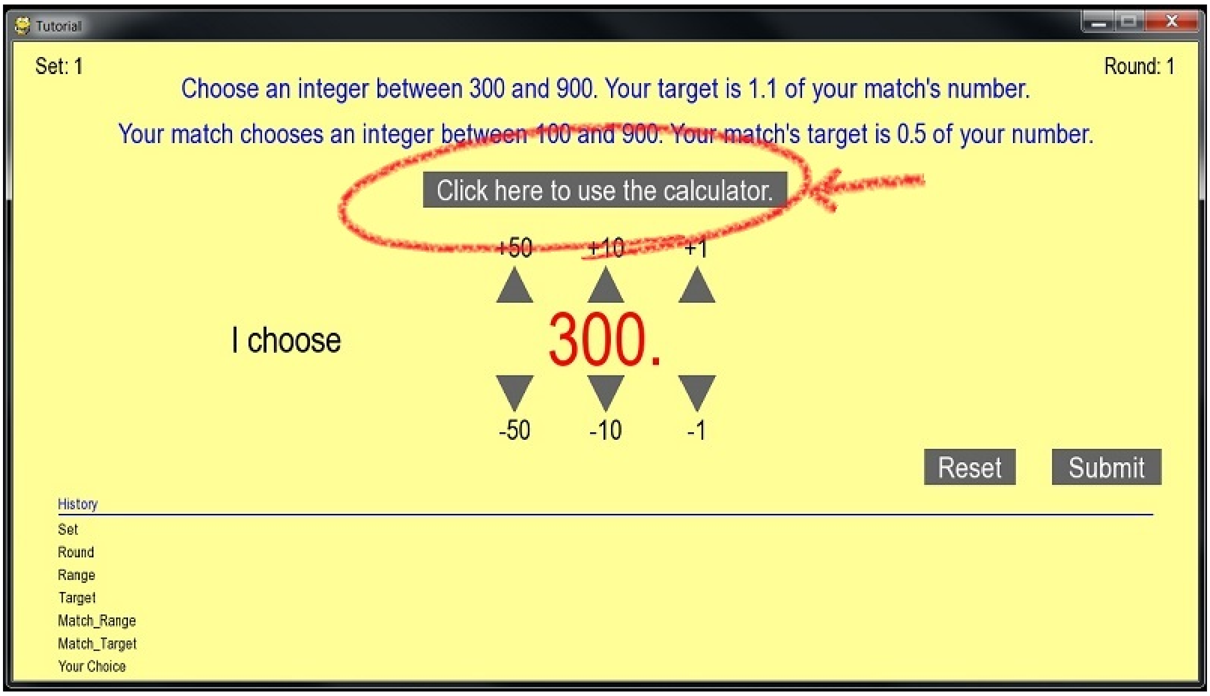
Showing you the formula calculating the payoff is overkill: Because the components of the game (i.e., your range and target, and your match's range and target) change in every set, the formula contains some messy mathematical normalizations and adjustments. Instead, we provide a calculator for you. Details will follow.

For now, please remember that your payoff is larger the more accurately you guess your match's choice. The calculation panel will do the remainder. If your choice is exactly the same as the 'ideal choice' (= your target * the match's choice), you will earn 25 cents for the round. The greater the difference between your choice and the ideal choice, the smaller you earn. The worst choice may give you 0 cents for the round.

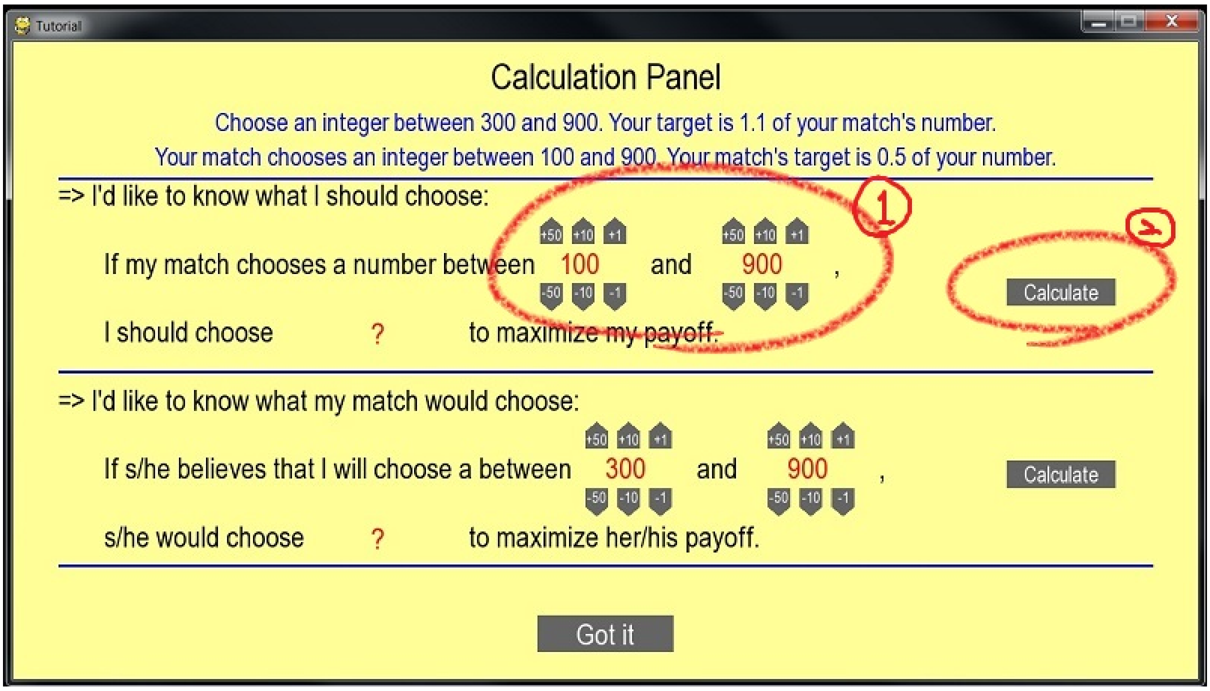
Calculation Panel:

We want you to guess your match's choice, but we do not want you to feel mathematically burdened. We provide a calculation panel. We highly recommend that you rely on the panel. If you input your guess about your match's choice, then the calculation panel outputs your ideal choice. Also, if you input your choice, then the calculation panel outputs your match's ideal choice.

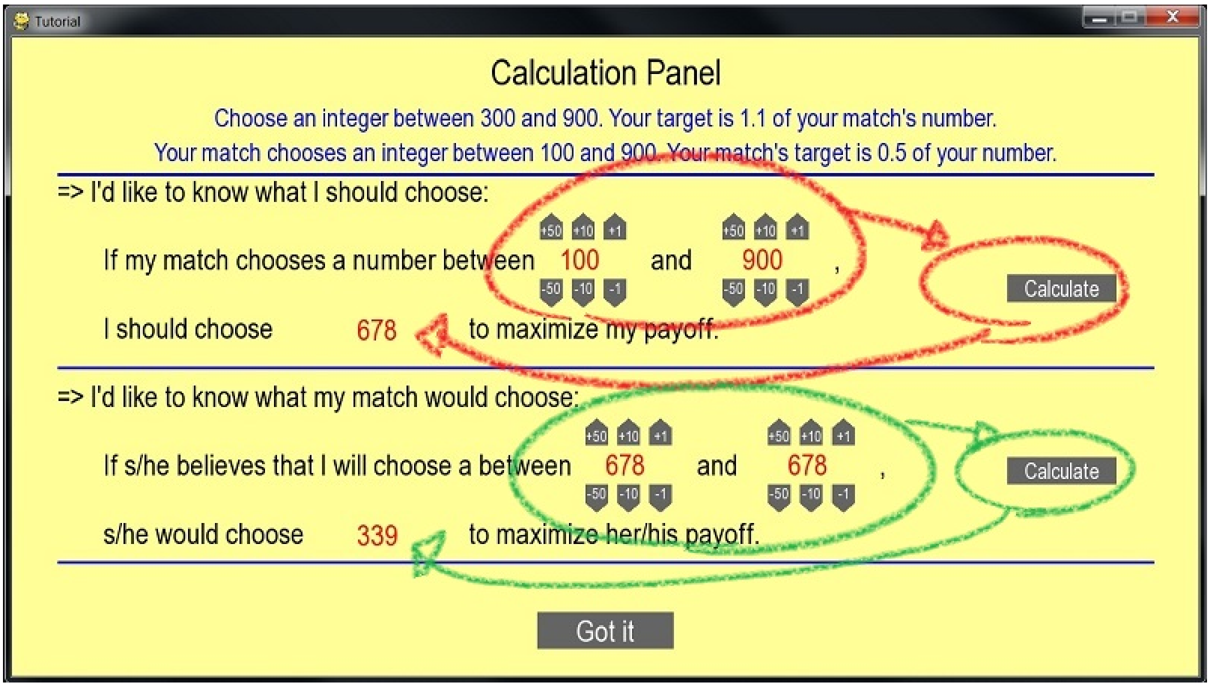
You can use the calculation panel as many times as you want. Check the following screen-shots. You may also use paper and pencil to write down the past results of your calculation. The paper will be wasted anonymously. Your usage of the calculation panel and paper will not affect your payoff.



Screenshot 1 [Set 1, Round 1]: To use the calculation panel, click the button indicated.



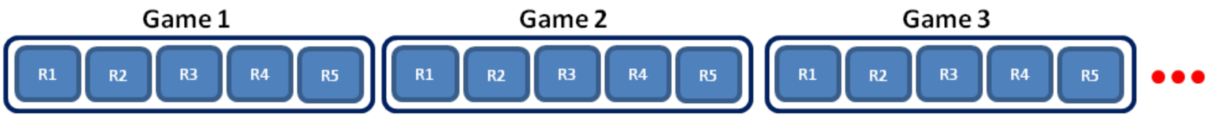
Screenshot 2 [Calculation Panel (1)]: If you want to know what you should choose, input your guess, and click 'Calculate' button. Similarly you can calculate what your match would choose.



Screenshot 3 [Calculation Panel (2)]: You can use this calculation panel as many times as you want. Note that it is okay to input the same number in the form of the range. (See the green circle above.)

Structure of the Experiment:

Again, this experiment consists of 8 sets. In the first set, you will play a game with a match for 5 rounds. In the next set, you will play another game with another match for five rounds, and so on. Even though the game will be the same for the five rounds within a set, we do not know whether your match will make the same choice for all five rounds.



* You play a game with a match for five rounds. After that, you will be matched with another participant.

Screening Test:

You will answer four multiple choice questions to check your understanding of this instruction. You may want to go back to review, or ask an experimenter for help. You can participate in the experiment only if ALL of your answers are correct. The main purpose of this screening test is to help you understand the instructions, not to cause you any stress. It is okay for you to ask an experimenter to help you if you are in doubt.

Q1. The experiment will consist of (A) sets. In each set, you will make (B) rounds of decision-making. What are appropriate numbers in (A) and (B)?

Q2. Suppose that your range is [100,500], your target is 0.5, your match's range is [100,900], and the match's target is 1.1. Which of the following is NOT true?

Q3. You play a game with a match for five rounds. Which of the following is true?

Q4. Suppose that you somehow guess that your match will pick any number between 200 and 500. Use the part of the calculation panel below and choose what you should choose. [The calculation panel is provided.]