Probability Matching and Strategic Decision Making*

Duk Gyoo Kim† Hee Chun Kim ‡

August 12, 2020

Abstract

This paper examines a link between an individual's (possibly limited) strategic thinking in the 11–20 money request game and (possibly non-rational) decision-making patterns in the matching pennies games. Experimental evidence shows that subjects' strategic behavior, which used to be understood as a result of cognitive iterations, is closely related to probability-matching patterns. Ignoring some individuals' choice randomization overestimates the variance of levels in cognitive iterations. Probability matchers do not seem to have less ability of cognitive iteration in strategic decision making. The relationship requires attention because the assumption that individuals are rational in the decision-theoretic sense may create a sizable misinterpretation of strategic behavior.

Keywords: Level-k reasoning, Probability matching, Cognitive bound, Preference for randomization

JEL: C72, C91, C92, D81

1 Introduction

A growing number of studies in social science consider bounded rationality both in a non-strategic setting, where a single player decides in an uncertain state, and in a

*The previous version of this paper was titled "Mixing Propensity and Strategic Decision Making." We thank Colin Camerer and Brian Rogers for their encouragement, PJ Healy, Youngdeok Hwang, Sang-Hyun Kim, Kirby Nielsen, Sangwon Park, Henrik Orzen, Jörg Oechssler, and participants in several conferences and seminars for their comments, and Kiril Lavrov for his assistance. 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*

strategic setting, where she responds to the other agents' unknown intentions and actions. When deciding on the non-strategic setting, individuals are often cognitively limited: They may not recognize or understand all the aspects that affect their payoffs or lack the cognitive ability to draw an ideal decision as much as they need. Observations from strategic environments 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.

This paper's primary goal is to examine how individuals' non-strategic—and possibly non-rational—decision-making patterns over probabilistic events are related to their strategic—perhaps cognitively limited—ones. This examination is important when inferring from observations in strategic situations.

To analyze (cognitively-limited) observations in strategic situations, 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*. That is, if there is a stochastically dominant action for a player given her belief about the other player's limited reasoning, then the literature has assumed that she must have chosen the stochastically dominant one all the time. Meanwhile, experimental work shows that when subjects are asked to make repetitive decisions under uncertainty, a significant fraction (varying from 20% to 50% by study) of subjects do not make decisions that maximize their expected payoff. Instead, they tend to *match* their decision frequencies to the probability of events, called *probability matching* (Rubinstein, 2002; Neimark and Shuford, 1959). For example, suppose that people are asked to play ten rounds of Matching Pennies (MP) games wherein each game, a coin, with a 70% probability of landing on heads, will be tossed independently. Some of them choose heads for seven out of the ten rounds and tails for the other three rounds, to match their relative choice frequencies with the probability of events, while they should have chosen heads for all the rounds to maximize the expected payoff. Although investigating why some people have such a preference for randomizing their choices is worthwhile,[1] we want to clarify that the primary purpose of this study is not to rationalize the probability-matching behavior. Rather, we take

[1]Many studies provide models of preferences for randomization and empirical support of them. Dwenger et al. (2018) provide a theory of responsibility aversion, which implies a demand for randomization. Levitt (2020) 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. (2020) consider convex preferences to account for the affinity towards randomization among equally preferred options. Miao and Zhong (2018) find experimental evidence of choice randomization in social settings. Richter and Rubinstein (2019) provide an axiomatic model generalizing the standard Euclidean definition of convex preferences.

their choice patterns from the non-strategic environment as given and investigate further whether and to what extent the existence of the probability-matching, or broadly speaking, choice-randomizing players affects the analysis of the cognitive bounds inferred from the observations in strategic situations.

We claim that when some individuals' choice-randomization behaviors are ignored, it is challenging to correctly map the individual's strategic actions to her underlying belief. Two leading theories formalizing bounded rationality in strategic decision making, 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), share an assumption that individuals use only finite (=k) steps of iterated dominance, and such k varies by individual. To analyze experimental observations, previous studies relying on one of those two models 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 subjects can have a limited capacity of reasoning, but within their capacity, no subject exhibits any choice randomization behavior, including probability matching. It may create a sizable misinterpretation *if* the choice randomization behavior in a non-strategic situation is maintained in a strategic situation. An individual who tends to exhibit probability matching may show heterogeneous choice patterns even when she has a homogeneous belief, and she may be merely regarded as a less sophisticated player due to "noisier" choices. Naturally followed questions include whether choice randomization behavior in a non-strategic setting is indeed related to the behavior in a strategic setting, and how the additional observations on choice randomization help elicit the belief distribution in the level of reasoning.

To address our questions, we conducted two sets of within-subjects laboratory experiments: a set of two repeated matching pennies (MP) games, and a set of repeated 11–20 (ET) money request games (Arad and Rubinstein, 2012). In a nutshell, from the MP experiment, we distinguish individuals randomizing choices from those who consistently make stochastically dominant choices. Then we observe how the observations from the ET experiment vary by choice pattern categorized in the MP experiment. We also incentivized subjects to reveal their beliefs about choice distributions at the end of those two experiments.

Our observations are summarized as follows. First, in the MP games, about 75% of the subjects were classified as a rational optimizer (RO), each of whom makes consistent choices of maximizing the expected payoff, and a quarter of them were classified as a probability matcher (PM), each of whom exhibits a probability matching tendency. Second, in the ET game, the choice variance of the PMs is larger than that of the ROs, suggesting that the probability matcher's tendency of choice randomization is also maintained in the strategic situations. Third, the average number of cognitive iterations of the PMs was not statistically different from that of the ROs, which suggests that the choice randomization in the ET games is not due to the less cognitive capacity of the PMs. Fourth, there is a strong correlation between a subject's choice randomization and his/her belief in the choice distributions.

Altogether, we find a subject's strategic behavior observed in the ET games is closely related to the decision-making patterns in the MP games. When we ignore the probability matchers, the distribution of the estimated cognitive iterations in the strategic situations would be with an overestimated variance.

The rest of this paper is organized as follows. In the following subsection, we review related studies. Section 2 describes the details of the experimental design and procedure. Section 3 illustrates, with two testable hypotheses, how the information on probability matchers can crucially affect the elicitation of the cognitive ability from the observations. 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. 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)—share two assumptions: (1) individuals are rational in the decision-theoretic sense as 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 the 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 the adoption of the Level-k model. However, some subjects explicitly mix two or more different strategies, each representing 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 result 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 subjects assume that their opponent plays either the L1 strategy or the L0 strategy; the latter is a uniform random strategy, although the appropriateness of such an assumption about the L0 behavior is often arguable. Depending on their 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 and 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 (the best response to the heterogeneous belief as a whole) can be strictly better than mixing several choices, each of which respectively corresponds to the best response to a part of the heterogeneous belief.

In the sense that we try to better understand higher-order rationality, our goal is consistent with that of Kneeland (2015), who proposes a more explicit design of experiments to identify higher-order rationality. Rather than adopting Kneeland's ring games of many (more than three) players, we stick to the two-person guessing-style 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 money request game introduced by Arad and Rubinstein (2012) is an excellent tool for eliciting higher-order rationality.[2] We conduct the 11–20 game as a part of our experiment because it is simple and less arguable on the assumption about the L0 behavior. The only, but crucial difference from Arad and Rubinstein (2012) is that we ask subjects repeatedly make decisions without feedback. This repetition allows us to examine further the cognitive ability of the subject and potential misrepresentation.

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

[2] Alaoui et al. (2020) also used a variation of the 11–20 money request game. One of their goals is to discern whether the level-k behavior is induced by the cognitive capacity or by the belief about the matched player's cognitive capacity. For example, L3 behavior can be observed either because the player is capable of performing cognitive iterations up to three times or because the (more capable) player best-responds to the matched player who is believed to behave as L2. They find that the cognitive bound is binding for a large fraction of subjects.

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 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 vast majority of experiment participants exhibit stochastic choice when asked to answer the same questions several times in a row. Dwenger et al. (2018) reported that university applications in Germany exhibit a choice pattern consistent with a preference for randomness. If a 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 (e.g., 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 the process regarding individual optimization patterns. In a similar vein, Agranov et al. (2020) examine whether a subject's type regarding randomization behaviors is stable across different choice environments, and they find that randomization preferences are

[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 for all 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.

highly correlated across domains. Although Agranov et al. (2020) also compare the choice patterns in individual choice questions with those in games as we do, their main goal is to examine whether the individual heterogeneity is indicative of a stable distribution of types, while ours is to find the implications of the link between the choice randomization and strategic behaviors.

2 Experimental Design and Procedure

We used a within-subject design. Each subject participated in two different experiments and a follow-up belief elicitation: In the Matching Pennies (MP) games, subjects made a streak of decisions in which payoffs depend on realized (but unknown) events. In the 11–20 Token Request (ET) games, subjects made a streak of decisions in which payoffs depend on the randomly matched subject's decisions. After that, they were asked to guess the distributions of entire choices made in the session, and the ones with closest guesses earned additional payoffs.

In the MP games, the subjects make eight choices in total to earn points from two games. The payoff matrix in Table 1a describes the first game. A subject's options, U and D, are on the left column. A probability distribution, (H, 3/4; T, 1/4), is on the top: Event H is realized with a 3/4 of chance, and event T is with a 1/4 of chance.[4] When a subject chooses U and event H is realized, the subject earns 1 point. Each point that the subject had won in the MP games was converted into 80 cents. A new event is independently realized before making each decision. The subjects make decisions without knowing the realized events. After making four independent decisions without feedback, the subjects play the second game, described in Table 1b. In this game, the subject's options are U, M, and D, and the event will be L with a probability 1/2, C with a probability 1/4, and R with a probability 1/4. Based on the subjects' choice patterns from the two different games, we categorize them into two types: the rational optimizer (RO) who chooses U (the stochastically dominant option) consistently, and the probability matcher (PM) who randomizes choices close to the probability distribution.

Table 2 shows how a player with a particular type would choose actions. When a player is expected to choose an action A∈ {U,M,D} for n times, it is denoted by An. An RO will play U, the choice that gives the largest expected payoff, all of the time. We denote the RO's play by U4. A PM will match the frequency of her choices with the probability of events. Thus, in Game 1, she will mix three Us and one D, and in Game 2, she will mix two Us, one M, and one D, up to permutation. Similarly, such a play from the PM is

[4] We instructed what we mean by a probability distribution and how an event is independently drawn from the probability distribution in plain words.

	H, 3/4	T, 1/4
U	1	0
D	0	1

(a) MP: Game 1

	L, 1/2	C, 1/4	R, 1/4
U	1	0	0
M	0	1	0
D	0	0	1

(b) MP: Game 2

Table 1: MP games

denoted by U3D1 in Game 1 and U2M1D1 in Game 2 respectively.

MP	Game 1	Game 2
Rational Optimizer	U4	U4
Probability Matcher	U3D1*	U2M1D1*

*: up to permutation

Table 2: Predicted Behaviors of the Two Types in MP Games

In the ET games (Arad and Rubinstein, 2012), the subjects make eight decisions in total, with knowing that they are randomly matched with an anonymous participant for the first four decisions and another match for the last four decisions. In each decision round, each subject chooses one of the integers $r \in \{11, \ldots, 20\}$. The subject's payoff is $r+20$ if the choice of the match in that round is $r+1$, and r otherwise. That is, if the subject believes that her match would choose, for example, 19, then the best response is to choose 18 so that the payoff can be 38 (=18+20).[5] One of the eight rounds was randomly selected for payment.[6] The tokens earned in the selected round were converted into euros at the rate of 1 token to 30 cents. For counter-balancing, the ET games were conducted before the MP games for 24 subjects in two sessions.

The only, but crucial difference from Arad and Rubinstein (2012) is that we ask the subjects to make repetitive decisions without feedback. As we will illustrate in the next section, unless a subject believes that the match would play the mixed strategy of the unique equilibrium, an RO has little reason for randomizing the eight decisions. However, a PM may randomize the decisions if they have an heterogeneous belief about the match's strategy. If the ET game were to be conducted once, the only observation might

[5]The unique mixed-strategy Nash equilibrium is to mix 15, 16, 17, 18, 19, and 20 with probabilities 0.25, 0.25, 0.20, 0.15, 0.10, and 0.05, respectively.

[6]Azrieli et al. (forthcoming) show that the random payment mechanism is incentive compatible in a very general environment, but they also show that the pay-all mechanism is too incentive compatible in a slightly more restricted environment. One advantage of paying for all rounds is to minimize unnecessary effects of risk preferences and to induce risk-neutral behavior (Walker et al., 1990), so we pay all rounds in the MP games.

not help us to recover the subject's cognitive ability. By observing repetitive decisions, we intend to elicit differences in responses according to subjects' type.

After making 16 decisions (4 for the first MP, 4 for the second MP, and 8 for the ET) in total, the subjects were incentivized to correctly guess the aggregate choice distributions. For example, with 15 participants in a session, there are 60 choices from the first MP games in aggregate. The subject whose guess (about how many of 60 choices were U and how many were D) is closest to the actual choice distribution won 4 extra euros.[7] Similarly, they were asked to guess the choice distributions for the second MP games and the ET games.

Seven sessions of laboratory experiments were conducted with a total of 106 participants at the Mannheim Laboratory for Experimental Economics (mLab) in Fall 2019.[8] The participants were drawn from the mLab subject pool. Python and its application Pygame were used to computerize the games and to establish a server–client platform. After the subjects were randomly assigned to separate desks equipped with a computer interface, they were asked to carefully read the instructions before they took a quiz to prove their understanding of the experiment. Except for mentioning that there are three different tasks, the instructions cover the upcoming task only. Those who failed the quiz were asked to re-read the instructions and to retake the quiz until they passed. An instructor answered all questions until every participant thoroughly understood the experiment. In the ET games, although new pairs were formed at the end of the fourth round, there was no physical reallocation of the subjects, and they only knew that they were randomly shuffled. They were not allowed to communicate with other participants during the experiment, nor allowed to look around the room. It was also emphasized to participants that their allocation decisions would be anonymous. Payments (10.63 euros on average) were made in private, and subjects were asked not to share payment information. Each session ran less than an hour.

3 An Illustration

Before reporting the experimental findings, we illustrate how the assumption about decision-theoretic rationality could yield a sizable difference about the inference of the experimental evidence on strategic behavior, when a substantial fraction of the popula-

[7] The subjects were told that the prize would be shared in case of multiple winners, but it did not happen.

[8] Pilot sessions with 62 subjects were conducted in the Missouri Social Science Experimental Lab at the Washington University in St. Louis in 2015, and the data are available upon request. Although the observations from the pilot experiments deliver messages qualitatively similar to what we report in this paper, we did not combine the two datasets because each sample set represents a different population at a distinct time window. The experimental design and the pay scale were also different: Rather than the 11–20 games, we conducted two-person guessing games.

tion are randomizing their choices as if they do in the decision-theoretic situations.

Only for this section, we maintain two assumptions, which will later serve as testable hypotheses. First, all players believe that their matched players are cognitively limited. Second, a probability matcher, if one exists, matches the choice frequencies with the underlying belief about other's cognitive levels.

Consider the following hypothetical situation. Suppose that there are several decision makers (DMs) whose level of cognitive iteration is known to be less than three. Each DM makes a streak of ten decisions against one anonymous match to maximize her payoffs, and the optimal decisions simply depend on the matched player's level of cognitive iteration. Along with the ET games, suppose that the DM's optimal strategy to Level-k player is to perform $a_k \in \mathbb{N}$, for $k = 0, 1, 2$. That is, if the DM believes that the match is a $L1$ type for sure, then she must choose a_2. Table 3 illustrates exemplary choices from several DMs, which are categorized into one of the six equally-populated patterns. (Some of the patterns are observationally identical, but we split them to make the population share of each pattern equal.)

Patterns	1	2	3	4	5	6	7	8	9	10	Type
Pattern1	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	RO
Pattern2	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	a_1	RO
Pattern3	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	RO
Pattern4	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	a_2	PM
Pattern5	a_1	a_1	a_1	a_2	a_1	a_1	a_2	a_1	a_1	a_2	PM
Pattern6	a_2	a_2	a_1	a_2	a_2	a_2	a_1	a_2	a_1	a_2	PM

Table 3: Illustration: Representative Choice Patterns

Suppose that an econometrician, who observes each DM's choices but does not know her type, tries to estimate her level of cognitive ability and the belief distribution, with assuming decision-theoretic rationality. DMs, who exhibit Pattern1, Pattern3, and Pattern4, made consistent a_2 decisions, so they will be coded as Level-2 players. DMs with Pattern2 will be coded as Level-1 for sure. DMs with Pattern5 will likely be coded as Level-1, and those with Pattern6 coded as Level-2, although there are three "noisy" decisions for each of them.

Now we suppose that the econometrician knows the type of the DMs: Then substantially different conclusions can be drawn. If the econometrician knows that DMs with Pattern5 are PMs who exactly match their choice frequencies to their underlying belief about their partner's cognitive level, then their choices can be interpreted as the following: DMs with Pattern5 are Level-2 players who believe that 70% of the population is Level-0, and 30% of it is Level-1. Similarly, DMs with Pattern6 are also Level-2 players but believe that the population distribution consists of 30% of Level-0 and 70% of Level-1. After learning this, we are no longer sure if DMs with Pattern2 are Level-1: We can interpret that they might have the same belief of DMs with Pattern5 but make stochastically dominant choices to maximize her expected payoff by best-responding to Level-0 players' action, the more likely one. If we assume further that the belief distributions of the ROs are similar to those of the PMs, then we could interpret that choices of DMs with Pattern1 and Pattern3 come from the same belief of those with Pattern4 and Pattern6, respectively. Thus, DMs with Pattern2 might be sharing the same belief with those with Pattern5, and hence, be coded as Level-2 players. Table 4 summarizes our arguments.

DMs w/	When PM types are ignored	When types are informed
Pattern1	L2 who believe 100% of L1	L2 who believe 100% of L1
Pattern2	L1 who believe 100% of L0	L2 who believe 30% of L1 and 70% of L0
Pattern3	L2 who believe 100% of L1	L2 who believe 70% of L1 and 30% of L0
Pattern4	L2 who believe 100% of L1	L2 who believe 100% of L1
Pattern5	L1 who believe 100% of L0	L2 who believe 30% of L1 and 70% of L0
Pattern6	L2 who believe 100% of L1	L2 who believe 70% of L1 and 30% of L0

Table 4: Illustration: Estimated Levels of Cognitive Ability

From this illustration, we want to emphasize how recognizing probability matchers is vital for two reasons. First, the assumption of decision-theoretic rationality can lead to a substantial overestimation of the variance of the cognitive levels. When the DM's choice-randomization types are considered, the entire population is coded as L2, while without taking the types into account, a third of the population are regarded as L1. Second, repeated choices from the PMs can help better understand the DM's true underlying belief about the cognitive ability of the population. Of course, we drew the second claim with the assumption that the PMs match the choice frequencies with their underlying belief, so it must be examined with the experimental data.

4 Experimental Findings

To examine possible effects from the order of two games, we compared the observations from the 24 subjects who first played the ET games with others (81 subjects) who played the MP games first. We ensure that the order of the two games did not have a meaningful impact. Two-sample Kolmogorov-Smirnov (KS) test on each of the four batches of observations—the first MP, the second MP, the first half of the ET, and the second half of the ET—does not reject the null hypothesis that two data samples come from the same distribution (combined KS p-values are 1.000, 1.000, 0.840, and 0.647, respectively). From now on, we combine two data sets for summarizing our four findings.

Observation 1. *A quarter of the subjects were classified as Probability Matchers.*

First, in the MP games, about 75% of the subjects (80 out of 106) were classified as ROs, each of whom consistently chose the stochastically dominant choice, U, in all decision rounds. A quarter of the subjects were classified as PMs, each of whom exhibits a probability-matching tendency. However, the PM's decisions cannot be deemed as evidence of pure random choices. No subjects chose stochastically dominated choices more: For example, in the second MP games, the mode choice of every PM was U. Table 5 summarizes the aggregated relative choice frequencies.

	U	D
All	0.96	0.04
RO (75.5%)	1.00	0.00
PM (24.5%)	0.86	0.14

(a) MP: Game 1

	U	M	D
All	0.93	0.06	0.01
RO (75.5%)	1.00	0.00	0.00
PM (24.5%)	0.64	0.22	0.13

(b) MP: Game 2

Table 5: Relative Choice Frequencies

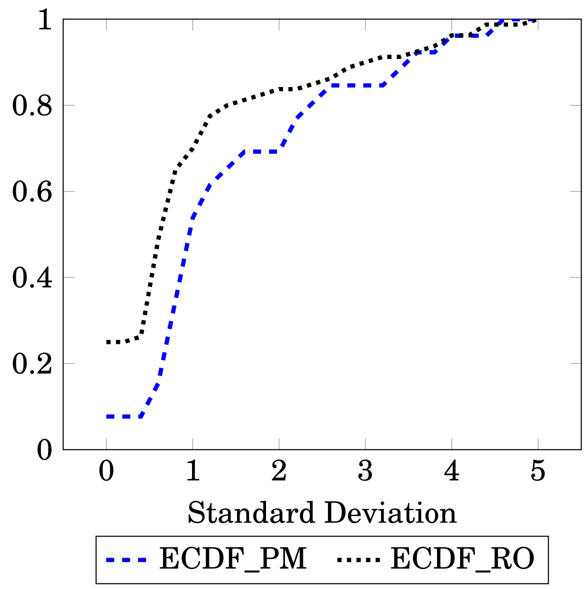
Observation 2. *The choices of the PMs in the ET games varied more.*

Second, in the ET games, the choice variance of the PMs was larger than that of the ROs. Figure 1 shows the empirical CDFs of the standard deviations of individual choices in the ET games, grouped by type. The PMs diversify their choices more than the ROs (KS test=1.4769, p-value=0.0255). A quarter of the ROs made the eight decisions constant. These patterns are unchanged when we draw similar empirical CDFs for the choices with the first match or the second match only. Relating to the choices in the MP games, this finding suggests that the tendency of choice randomization is also maintained in strategic situations.

Observation 3. *Both ROs and PMs did not play the Nash equilibrium.*

Third, neither the behavior of the ROs nor that of the PMs can be explained by Nash equilibrium. Our finding confirms the reports of Arad and Rubinstein (2012). Table 6 shows the unique symmetric Nash equilibrium distribution (assuming that players maximize the expected monetary payoff) and the actual choice distributions. The choice distribution is significantly different from the Nash equilibrium (χ^2 goodness of fit test, p-value<0.0001). Only 7 percent of the whole actions were 15 and 16, which is significantly

Figure 1: CDFs of the Standard Deviations of Individual Choices in the ET Games



smaller than 50 percent, the mixing probability in equilibrium. The vast majority of the actions were 17, 18, and 19, and those are corresponding to Level-3, -2, and -1 cognitive iterations, respectively.

Action	11	12	13	14	15	16	17	18	19	20
Equilibrium (%)					25	25	20	15	10	5
All (%)	7	1	1	2	3	4	15	27	26	13
RO (%)	8	1	1	1	3	1	13	30	26	13
PM (%)	6	2	2	4	4	7	20	19	25	11

Table 6: Equilibrium and Actual Distributions of the ET Games

The average level of cognitive iterations of the PMs (2.8125) was not statistically different from that (2.4922) of the ROs (t=1.6684, p-value=0.0956). Rather, the PMs' average level is slightly higher than that of the ROs. This finding suggests that the choice randomization observed in the ET games may not be due to the less cognitive ability of the PMs.

Observation 4. *Choices of both ROs and PMs in the ET reflected their beliefs.*

Fourth, there is a strong correlation between a subject's choice randomization behavior and his/her belief of the actual choice distributions. That is, a PM believes the choice distributions are dispersed more than what an RO does. This finding adds more confirmation to our main hypothesis that a subject's type of choice randomization in a non-strategic situation is likely to maintain in a strategic situation.

	U	D
Actual	0.96	0.04
RO	0.935	0.065
PM	0.802	0.198

(a) MP: Game 1

	U	M	D
Actual	0.93	0.06	0.01
RO	0.891	0.057	0.052
PM	0.644	0.199	0.168

(b) MP: Game 2

	11	12	13	14	15	16	17	18	19	20
Actual (%)	7	1	1	2	3	4	15	27	26	13
RO (%)	10	2	2	2	3	4	9	16	25	27
PM (%)	13	3	2	5	7	6	12	18	19	16

(c) ET Games

Table 7: Relative Choice Frequencies and Guesses by Type

Table 7 shows how subjects guessed the actual choice distributions. In the MP games, the ROs' guess was distinctively closer to the actual distributions, although they consistently chose U. This observation suggests that the ROs have a good sense that a small fraction of the whole subjects would exhibit a sort of choice-randomization tendency. Meanwhile, the PMs seem to believe that the entire subjects would behave like themselves. Their guesses are close to their actual choice distributions in Table 1. For example, in the second MP games, 64%, 22%, and 13% of the PMs' choices were U, M, and D, respectively, and they guessed the actual distributions would be around 64%, 20%, and 17%.

In the ET games, the ROs' guess was closer to the actual distribution, although they overly weighted the L0 behavior. The average level of cognitive iterations, according to their guess is 2.47, which is exactly corresponding to their actual average level of cognitive iterations, 2.49. 69% of the ROs' choices were 17, 18, and 19, and 68% of their guesses were 18, 19, and 20. Therefore, ROs were indeed best-responding to their beliefs. The PMs, who randomized their decisions more than the ROs, guessed that the actual choice distribution would be more dispersed than what they chose. Similar to what we observe for the ROs, PMs also diversified their choices in the ET games to best respond to what they believe. 71% of the PMs' choices were between 16 and 19, and 65% of their guesses were between 17 and 20.

To provide further evidence on how much each type reflects their guesses to choices, we develop a measure capturing the degree of discrepancies between individual's choices and guesses. The discrepancy measure is the sum of the squared differences between the belief distribution induced by the individual's choices and the self-reported belief distribution. To be more illustrative, we took subject 1 in our first session as an example. In the ET games, her choices in rounds 1 to 8 were (18, 18, 18, 18, 17, 17, 17, 17). When asked to guess the choice frequencies over 11 to 20 in the whole session of 20 subjects, her guesses were (0, 0, 0, 0, 0, 0, 40, 90, 20, 10). Her discrepancy measure is calculated in the following way. First, we standardize the ratio of the guesses induced from an individual's choices. Due to the structure of the ET game, her choices correspond to her guesses. For example, the choice of 18 corresponds to the guess of 19 since 18 is the best response to the match's action of 19. Thus, her standardized belief distribution induced by her choices is (0,0,0,0,0,0,0,0.5,0.5,0). Second, we standardize the ratio of the self-reported guesses. Her standardized self-reported distribution of guesses was (0,0,0,0,0,0,0.25,0.5625,0.125,0.0625). The measure of the individual discrepancy is the sum of the squared differences between the two distributions, which is $6(0-0)^2+(0-0.25)^2+(0.5-0.5625)^2+(0.5-0.125)^2+(0-0.0625)^2 = 0.2109$. We can calculate the discrepancy measure for every subject in the same way. Figure 2 shows the box and whisker plots of the individual discrepancies by type.

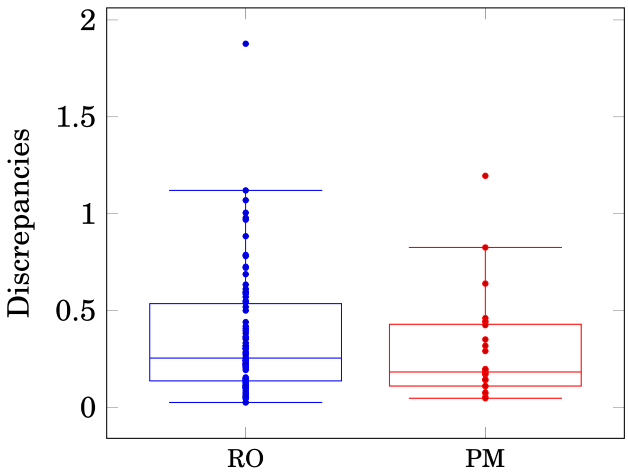


Figure 2: Box Plots of Individual Discrepancies by Type

Figure 2 shows that the two distributions by type seem to have no significant differences, except for a few outliers. The mean discrepancy of the ROs is 0.3647, and that of the PMs is 0.2885. We cannot reject the null hypothesis that two distributions have the same mean in discrepancy (p-value of two-sample t-test=0.276) nor reject the null hypothesis that two distributions are from the same population distribution (combined KS p-value=0.097.) Along with the points that the PMs' choices and guesses are more dispersed than those of the ROs, this result confirms that both the ROs and the PMs are reflecting their beliefs to a similar degree.

5 Concluding remarks

In this paper, 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 the subjects into two different types: the rational optimizer (RO), and the probability matcher (PM) by the observed choices in the MP games. We found more than a quarter of the subjects show choice patterns other than rational optimization. Our main observation is that when asked to make strategic decisions in the ET games, each type's choice patterns were different. In particular, PMs diversified their actions to multiple levels of cognitive iterations in the ET games in a similar way of diversifying their decisions in the ET games. The relationship between the decision-making patterns in the MP games and the ET games suggests that the literature, by assuming the decision-theoretic rationality, may have underestimated the level of cognitive iterations and overestimated the variance of belief distributions. If every subject were the PM-type, the belief structure estimated by the Level-k theory must be downward biased. On the other hand, if every subject were the RO, the belief distribution estimated by the Cognitive Hierarchy model would underestimate the variance of the distribution and conclude that such ROs were to have homogeneous underlying belief.

The main implication of our findings is twofold. First, an individual's decisions under probabilistic events can deviate from the rational-theoretic predictions, but such deviations can be predictable. Second, we can improve recovering the underlying belief of decision makers once we confirm that such prediction for deviation remains valid in strategic situations. We observed a statistical correlation between non-strategic and strategic decision making by exploiting our within-subject design that bridges DM's type of choice randomization and strategic decision making. Such an observation provides us somewhat limited but hopeful evidence that we can recover dominated parts of underlying belief, which were behind the dominant part of belief distribution, by making use of the PMs' choices.

References

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

_ , **Paul J. Healy, and Kirby Nielsen**, "Stable Randomization," 2020.

Alaoui, Larbi, Katharina A. Janezic, and Antonio Penta, "Reasoning about Others' Reasoning," *Journal of Economic Theory*, 2020, p. 105091.

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

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, forthcoming, *0* (0), 000–000.

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," *American Economic Review*, 2020, *107* (7), 2425–2445.

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: Evidence from university applications," *Journal of Public Economics*, 2018, *167*, 240–250.

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.

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," *The Review of Economic Studies*, 05 2020.

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

Miao, Bin and Songfa Zhong, "Probabilistic social preference: how Machina's Mom randomizes her choice," *Economic Theory*, 2018, pp. 1–24.

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.

Richter, Michael and Ariel Rubinstein, "Convex preferences: A new definition," *Theoretical Economics*, 2019, *14* (4), 1169–1183.

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.

Walker, James M., Vernon L. Smith, and James C. Cox, "Inducing risk-neutral preferences: An examination in a controlled market environment," *Journal of Risk and Uncertainty*, 1990, *3*, 5–24.

Appendix. Experiment Instructions

Welcome.

You will perform three different tasks. In each task, you will input your decisions to the computer interface. Your payment will be based on the decisions you make, the decisions that other participants make, and some luck. Your payment will be informed at the end of all three tasks.

Everyone starts the same task at the same time. If you finish some part of the experiment earlier than others, you will be asked to wait.

In each task, you will read the instructions first and, if necessary, will take a quiz to check your understanding of the instructions. Please read the instructions carefully.

Task 1: Decision-Making under Uncertainty

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 1/2 each. Then, the probability distribution of coin tossing can be described as (H, 1/2; T, 1/2).

When we say "an event (here, the face of a coin) is randomly drawn from (H, 1/2; T, 1/2)," it means that we toss a coin, and either H or T is realized with an equal probability, but we will not disclose what the actual realization is. Also, note that when each event is independently drawn, that event has nothing to do with the previous events whatsoever.

Another example: "an event is randomly drawn from (L, 0.2; C, 0.5; R, 0.3)" means the following three. (1) An event L (, C or R) will be drawn with a 20% (, 50% or 30%) of chance. (2) One among L, C, and R is realized according to their chances. (3) We will not disclose what the realization is.

During this task, 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.

For further explanation, please ask the experimenter at any time.

Your Task:

Your task is to make eight choices in total, to earn points from two games. The following payoff matrix describes the first game.

	H, 3/4	T, 1/4
U	1	0
D	0	1

Your options—U and D—are shown on the left. A probability distribution is on the top; (H, 3/4; T, 1/4): H happens in a 3/4 of chance, and T happens in a 1/4 of chance. The matrix shows your payoff. For example, if you choose U, and an event H is randomly drawn from the probability distribution, you will earn 1 point. If you choose D, when an event T is drawn, you will also earn 1 point.

You make four choices. For each choice, an event is randomly and independently drawn from the probability distribution.

The following payoff matrix describes the second game.

	L, 1/2	C, 1/4	R, 1/4
U	1	0	0
M	0	1	0
D	0	0	1

In this game, your options are U, M, and D, and the event will be L with a probability 1/2, C with a probability 1/4, and R with a probability 1/4. For example, if you choose D when the event L is drawn, you earn 0 points. If you choose M when the event C is drawn, you earn 1 point.

For each choice, an event is randomly and independently drawn from the probability distribution.

Payment:

All the points that you have earned in Task 1 will be converted into euros at the rate of 1 point = 80 cents.

Quiz:

Before you make your choices, you will answer two multiple-choice questions to check your understanding of this instruction. You can proceed only with all correct answers. You may ask an experimenter for help.

Q1 Suppose that the following matrix describes a certain game. Which of the following is true?

	H, 0.7	T, 0.3
U	1	0
D	0	1

(1) You can choose either H or T. (2) If you choose D, when an event T is drawn, you earn 0 points. (3) At the time you choose one option, you will not know your payoff. (4) The event will be H with a 50% of chance.

Q2 Suppose a probability distribution over events is (L,0.3; C,0.5; R,0.2). Which of the following is NOT TRUE? (1) When the realized event was L in the previous round, it is more likely to have an event R. (2) It is possible to face an event C both in the previous round and a current round. (3) In each round, a new event is randomly drawn from the probability distribution. (4) It is possible to face an event L in the previous round and face an event R now.

Task 2: 11–20 Token Request

Your Task:

Your task is to make four decisions with a randomly matched participant and make another four decisions with another randomly matched participant. In total, you make eight decisions. You will not know who your matches are, and they will not know you either.

In each decision round, choose one of the integers between 11 and 20, including 11 and 20. You will get the tokens (the currency in this task) corresponding to the integer you chose. Also, if your choice is one token less than your match's choice in that round, you will earn 20 extra tokens.

Payment:

One of the eight rounds will be randomly selected, and the earnings of that round will be paid. The tokens you have earned in that round will be converted into euros at the rate of 1 token=30 cents. Each round is equally possible to be selected, so it is of your best interest to consider every round equally seriously.

Quiz:

Before you make your choices, you will answer two multiple-choice questions to check your understanding of this instruction. You can proceed only with all correct answers. You may ask an experimenter for help.

Q1 Which of the followings is NOT TRUE? (1) You choose an integer between 11 and 20 in each of eight rounds. (2) In each round, you are randomly matched with a new participant. (3) In a particular round, if you choose 18, and your match chooses 17, you earn 18 tokens. (4) In a particular round, if you choose 19, and your match chooses 20, you earn 39 tokens.

Q2 Which of the followings is TRUE? (1) Your match will always choose the same integer for all four rounds. (2) Three randomly selected rounds out of the eight rounds will be paid. (3) You will know who your matches are. (4) At the end of the fourth rounds, you will be randomly matched with a new participant.

Task 3: Bonus Prizes

In this session are N participants, including you. You will win bonus prizes if you correctly guess how N participants answered in the previous two tasks in aggregate.

Bonus 1: In the first game of Task 1, each made four choices of either U or D. The probability distribution was (L: 3/4, R: 1/4). In total, there are $4 * N$ choices. Guess the total choice frequencies of U and D. If your guess is closest to the actual choice frequencies, you will win 4 extra euros. (When multiple winners, the prize will be split.)

Bonus 2: In the second game of Task 1, each made four choices among U, M, and D. The probability distribution was (L: 1/2, C: 1/4, R: 1/4). In total, there are $4 * N$ choices. Guess the total choice frequencies of U, M, and D. If your guess is closest to the actual choice frequencies, you will win 4 extra euros. (When multiple winners, the prize will be split.)

Bonus 3: In Task 2, every participant submitted eight integers between 11 and 20. In total, there are $8 * N$ choices. Guess the total choice frequencies of each integer. If your guess is closest to the actual choice frequencies, you will win 4 extra euros. (When multiple winners, the prize will be split.)