國立臺灣大學社會科學院經濟學系
碩士論文
Department of Economics
College of Social Sciences
National Taiwan University
Master Thesis

用經濟學實驗方法探索人們的策略性思考能力

Uncover Components of Individual's Strategic IQ: An Experimental Study

劉書彧
Shu-Yu Liu

指導教授: 王道一 博士
Advisor: Joseph Tao-yi Wang, Ph.D.

中華民國 103 年 8 月
August, 2014

摘要

本文透過「優勢策略可解賽局」、「同時選擇」與「先後選擇」的「空間版選美結果預測賽局」來探索人們的策略性思考能力。利用主成分分析法, 我們從實驗資料中歸納出五個策略性思考能力, 用以解釋人們在這些賽局中對戰所有人之期望報酬 56% 的變異: (一)、人們的逆推思考能力 (最能解釋人們在這些賽局中表現的主成分); (二)、人們的在高維空間下的逆推思考能力; (三)、人們的風險偏好; (四)、人們對於他人社會偏好的認知信念; (五)、人們對於他人認知雙方行爲的高層次信念。

關鍵詞: 因素分析、優勢可解賽局、選美預測賽局、策略性思考能力、多層次思考模型

Uncover Components of Individual's Strategic IQ: An Experimental Study

Shu-Yu Liu Joseph Tao-yi Wang*

August, 2014

Abstract

We employ principal component analysis to identify components of subject's strategic IQ in the following three classes of games: The two-stage dominance-solvable game, Chen, Huang and Wang (2013)'s simultaneous spatial beauty contest game, and the first-mover spatial beauty contest game. Parallel analysis retains the first five principal components (PCs), which account for 56% of the total variance of subject's normalized expected payoffs for each of the 33 games. We interpret these PCs as five strategic IQs: The first SIQ indicates subjects' abilities to perform backward induction and it is also the common g-factor that can predict subjects' performances in most games. The second SIQ could be interpreted as subjects' abilities to perform high dimensional backward induction. The third SIQ controls for subjects' attitudes toward risk. The fourth SIQ reflects subjects' beliefs about social preferences. The fifth SIQ measures subjects' accuracy of higher order beliefs about others.

Keywords: *factor analysis; dominance-solvable game; two-person guessing game; strategic IQ; level-k thinking*

JEL classification: *C91*

*Department of Economics, National Taiwan University, 21 Hsu-Chow Road, Taipei 100, Taiwan. Shu-Yu Liu: r99323001@ntu.edu.tw; Joseph Tao-yi Wang: josephw@ntu.edu.tw.

Contents

List of Figures

List of Tables

1 Introduction

Since Stahl and Wilson (1995) and Nagel (1995), researchers have explored human limits of strategic thinking and the existence of heterogeneous levels of beliefs about such cognitive limitations. In the "level-k" model pioneered by these authors, subjects anchor their beliefs in a strategically naïve initial assessment of others' likely responses to the game called "level-0" (L0), and then adjust them via "thought-experiments" with iterated best responses: level-1 (L1) best responds to L0, level-2 (L2) to L1, and so on. Players' levels (types) of strategic thinking are heterogenous, but each player's level (type) is usually assumed to be drawn from a common distribution. Camerer, Ho and Chong (2004) developed a closely related model, known as the "cognitive hierarchy" (CH) model, that assumes Lk types best respond to a mixture of lower types, which distribution is a Poisson distribution, but "truncated" at L(k-1). Recently, such level-k models have been widely developed to explain strategic behavior in various classes of games, including two-player guessing games, initial responses in hide-and-seek games, auctions, coordination games, cheap talk games, field settings such as Swedish LUPI lotteries, movie reviews, and even lookup patterns captured by various techniques of eyetracking (See Crawford, Costa-Gomes and Iriberri, 2013, for a review.).

Strategic IQ, first proposed by Camerer and Ho (2004),[1] measures "the degree in individual's ability to think strategically by analyzing and anticipating what others might know or do, and subsequently choosing rational responses that will outwit the opponents." For example, Bhatt and Camerer (2005) defined strategic IQ as the normalized expected payoffs one earns in eight 2-player matrix games from making decisions and predicting accurately other's choices (and predictions). They found that strategic IQ is negatively correlated with activity in the insula, suggesting that low strategic IQ subjects are too self-focused. In contrast, strategic IQ is positively correlated with caudate activity, suggesting that high strategic IQ subjects spend more mental energy predicting the opponent's behavior. Interestingly, they find no correlation between the "theory of mind" regions and strategic IQ, indicating that a simple average of normalized expected payoffs alone cannot account for one's strategic abilities.

In this study, we conduct a battery of games that induces heterogeneous responses, including two-stage dominance-solvable games, Chen, Huang and Wang (2013)'s simultaneous spatial beauty contest (SBC) games, and first-mover spatial beauty contest (1st-mover SBC) games. First, the two-stage dominance-solvable

[1]The strategic IQ site: http://128.32.75.8/siq/default2.asp

game is a simple extensive form game which involves two players acting sequentially. The first player (Player 1) chooses between action *left*, which enforces an "outside option" payoffs on the two players, and action *right*. If *right* is chosen, the responder (Player 2) determines the allocation of payoffs by choosing between *up* and *down*. Although the structure of this game is simple, it is sufficient to reproduce the main deviations from rational choice considered by previous studies. We adopt games from Beard and Beil (1994), Goeree and Holt (2001), and Ert, Erev and Roth (2011), which show heterogeneity in subjects' decisions in their studies. Secondly, Chen, Huang and Wang (2013)'s simultaneous SBC game is a spatial variant of Costa-Gomes and Crawford (2006)'s asymmetric two-person guessing game. In this game, two players are asked to choose locations simultaneously on a given two-dimensional grid map with different targets. One's target is defined as a relative location to the opponent's choice of location and is common knowledge for both players. The closer a player's choice is away from his target, the higher payoffs he earns. We adopt 6 games from Chen, Huang and Wang (2013) to identify subjects' levels of reasoning. Lastly, the 1st-mover SBC game is a sequential variant of the simultaneous SBC games, in which subjects choose first, playing against a computerized profit-maximizing player. Unlike the first two classes of games, solving the 1st-mover SBC game does not involve subject's belief about what others might know or do. Hence, it can be considered as a working memory task which reflects subject's ability to play best response and perform backward induction.

We define five ad hoc indicators on subjects' performance representing various strategic abilities in each class of games.In particular, CS-DSG and EV-DSG summarize subject's performance in the two-stage dominance-solvable games. CS-DSG counts the times subjects violate comparative static predictions and reflects subject's inability to respond to changes in game payoffs. EV-DSG represents subject's ability to perform backward induction and the accuracy of his belief about Player 2 subjects. In addition, EV-1st1D and EV-1st2D summarize subject's performance in the 1st-mover SBC games, reflecting the ability to perform backward induction against the preprogrammed second mover. Lastly, EV-SBC summarizes subject's performance in the simultaneous SBC games, reflecting subject's ability to perform backward induction and the accuracy of his belief (and higher order belief) about the opponents' choices of locations. Note that except for CS-DSG, the remaining four indicators are all defined by subjects' average expected payoffs across certain games that belong to the same predefined class. The results of these indicators show the heterogeneity in subject's strategic abilities.

Since the above indicators are ad hoc and the classification of games could be rather arbitrary, we employ principal component analysis to form several linear combinations of the normalized expected payoffs of the 33 games used in the experiment. The first five principal components are selected based on Horn (1965)'s parallel analysis and can be interpreted as the following strategic IQs, which represent different strategic abilities: SIQ_1 reflects the ability to perform backward induction. SIQ_2 indicates the ability to perform multi-dimensional backward induction. SIQ_3 could be interpreted as (and controls for) subjects' attitudes toward risk. SIQ_4 measures subjects' beliefs about others' social preferences. SIQ_5 captures subjects' accuracy of higher order beliefs about the opponents in the simultaneous SBC games. These strategic IQs are correlated with some of our ad hoc indicators, meaning that these indicators are not as arbitrary as one may think.

The rest of the paper is organized as follow. The next section describes the game structure and the theoretical predictions of each game. Section 3 describes the design of our experiment. Section 4 reports the aggregate results of the experiments. Section 5 explores subjects' strategic abilities in the experiment by establishing various indicators that reflect subjects' performance and the underlying strategic abilities in the experiment. Strategic IQs, which are formed by principal component analysis, are provided to summarize subjects' performance in all games used in the experiment. Section 6 concludes and sketches future research.

2 Game Structure and Theoretical Predictions

2.1 Two-Stage Dominance-Solvable Games

The two-stage dominance-solvable game is a simple extensive form game which involves two players acting sequentially. The game is presented in Figure 1. The first player (Player 1) decides to choose either "left" (L) to obtain an assured payoff $\pi_1(L)$, giving the second player (Player 2) $\pi_2(L)$, or "right" (R) to put Player 2 on the move. Under the latter, if Player 2 chooses "down" (D), the two players would earn $\pi_1(R,D)$ and $\pi_2(R,D)$, respectively; if Player 2 chooses "up" (U), they would earn $\pi_1(R,U)$ and $\pi_2(R,U)$, instead. To make this game interesting, we assume $\pi_1(R,D) > \pi_1(L) > \pi_1(R,U)$.

Assuming that Player 1 is self-interest and believes that Player 2 is also self-interest, subgame perfect equilibrium (SPE) makes specific predictions in this game. When $\pi_2(R,D) > \pi_2(R,U)$, SPE predicts that Player 2 would choose D (giving Player 1 $\pi_1(R,D)$), and Player 1 hence chooses R (since $\pi_1(R,D) > \pi_1(L)$). In contrast, when $\pi_2(R,U) > \pi_2(R,D)$, Player 2 would respond to Player 1's R choice by choosing U (giving Player 1 $\pi_1(R,U)$), and Player 1 hence chooses L (since

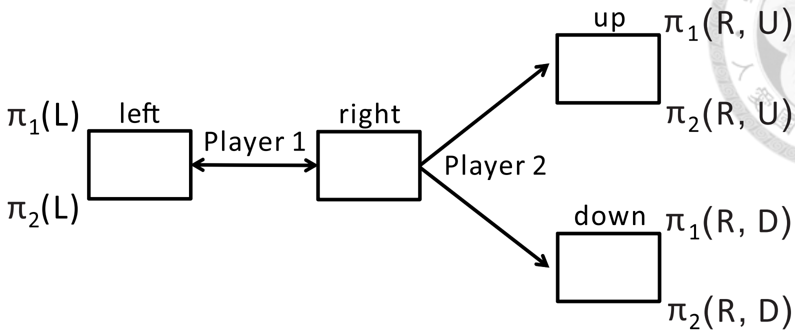


Figure 1: Two-Stage Dominance-Solvable Game

$\pi_1(R,U) < \pi_1(L)$).

When Player 1 does not think that all Player 2 subjects are self-interest and obey dominance, his belief about his opponent's rationality would affect his decision. In particular, a risk neutral Player 1 first forms the belief of the probability (or frequency) that a randomly selected Player 2 would choose D following R, $p(D|R)$, and uses it to calculate the expected payoff of choosing R, $E[\pi_1(R)] = p(D|R) \cdot \pi_1(R,D) + (1 - p(D|R)) \cdot \pi_1(R,U)$. Then, he compares this expected payoff with the assured payoff, $\pi_1(L)$, and chooses R if $E[\pi_1(R)] > \pi_1(L)$. Similarly, a risk averse Player 1 compares the assured payoff with the expected *utility* of choosing R, $u(\pi_1(R)) = p(D|R) \cdot u(\pi_1(R,D)) + (1 - p(D|R)) \cdot u(\pi_1(R,U))$, and demand a risk premium to compensate for the risk of choosing R. The threshold probability, $\hat{p}(D|R)$, represents Player 1's belief about Player 2's rationality required to justify choosing R. For a risk neutral (risk averse) Player 1, this threshold is the belief of the frequency of D choices that makes the expected payoff (utility) of choosing R equal to the assured payoff (utility) by choosing L:

$$\hat{p}(D|R) = \frac{\pi_1(L) - \pi_1(R,U)}{\pi_1(R,D) - \pi_1(R,U)}$$
$$\left(\hat{p}(D|R) = \frac{u(\pi_1(L))_{-}u(\pi_1(R,U))}{u(\pi_1(R,D)) - u(\pi_1(R,U))}\right)$$

Table 1 presents the payoffs and the SPE prediction of each game used in the experiment.[2] The payoffs selected for these games are motivated by a desire to induce various influences on Player 1 subjects' decisions. Game D1, D2, D3 and their variants have different threshold probabilities but the same SPE prediction, (R, D). Game RP and RP-VLR are rational punishment games in which Player

[2] Game D1 and its variants are adopted from Beard and Beil (1994). Game D2, D3 and their variants are similar to Goeree and Holt (2001). The remaining games (rational punishment games and trust games) are inspired by Ert, Erev and Roth (2011).

Table 1: Two-Stage Dominance-Solvable Games and Their Theoretical Predictions

Game	Payoffs: (Player 1, Player 2)			SPE	Risk Neutral Threshold
	L	(R,U)	(R,D)		$\hat{p}(D\|R)^{\dagger}$
Beard and Beil					
D1 (baseline 1)	$(9.75, 3)$	$(3, 4.75)$	$(10, 5)$	(R, D)	0.96
D1-LR (less risk)	$(7, \cdot)$	$(\cdot, \cdot)$	$(\cdot, \cdot)$	(R, D)	0.57
D1-MRs (more resentment)	$(\cdot, 6)$	$(\cdot, \cdot)$	$(\cdot, \cdot)$	(R, D)	0.96
D1-MRc (more reciprocity)	$(\cdot, 5)$	$(5, 9.75)$	$(\cdot, 10)$	(R, D)	0.95
D1-MA (more assurance)	$(\cdot, \cdot)$	$(\cdot, 3)$	$(\cdot, \cdot)$	(R, D)	0.96
Goeree and Holt					
D2 (baseline 2)	$(7, 6)$	$(6, 1)$	$(9, 5)$	(R, D)	0.33
D2-LA (lower assurance)	$(\cdot, \cdot)$	$(\cdot, 4.75)$	$(\cdot, \cdot)$	(R, D)	0.33
D3 (baseline 3)	$(8, 5)$	$(2, 1)$	$(9, 7)$	(R, D)	0.86
D3-LA (lower assurance)	$(\cdot, \cdot)$	$(\cdot, 6.75)$	$(\cdot, \cdot)$	(R, D)	0.86
D3-VLA (very low assurance)	$(40, 25)$	$(10, 34.75)$	$(45, 35)$	(R, D)	0.86
Ert, Erev and Roth					
RP (rational punishment)	$(6, 4)$	$(0, 3)$	$(14, 0)$	(L, U)	0.43
RP-VLR (very low risk)	$(1, 13)$	$(0, 4)$	$(14, 0)$	(L, U)	0.07
TG (trust game)	$(4, 1)$	$(0, 10)$	$(9, 9)$	(L, U)	0.44
TG-LRc (less reciprocity)	$(2, 0)$	$(0, 3)$	$(9, 2)$	(L, U)	0.22
TG-CR (costly repay)	$(3, 0)$	$(0, 10)$	$(8, 1)$	(L, U)	0.38

Note: $(\cdot, \cdot)$ indicates the payoffs are the same as those in the baseline game.
†: Individual threshold probability depends on subject's attitude toward risk. Here, we provide the threshold probability for a risk neutral Player 1 as a benchmark.

2's U choice not only maximizes his own payoff but also "punishes" Player 1's R choice that makes him earn less. Game TG, TG-LRc, and TR-CR are trust games designed to incorporate Player 1's belief about Player 2's social preference. In these games, Player 1 could choose the SPE prediction L to obtain the assured payoff, or choose R to increase Player 2's potential payoffs and expect a reciprocal, but dominated choice from Player 2.

The predictions for various influences on the probability (or frequency) that a randomly selected Player 1 would choose the secure option L ($p(L)$) of 15 games are as follows:

The first baseline game, Game D1, has a high threshold probability.In particular, the difference between $\pi_1(L)$ and $\pi_1(R, D)$ is only \$0.25 and the the difference between them and $\pi_1(R, U)$ are around \$7. Therefore, the risk neutral Player 1's threshold probability, $\hat{p}(D|R)$, of this game is high (0.96). Hence, some Player 1 subjects might choose L to earn \$9.75 for sure, violating the SPE prediction.

Game D1-LR, D1-MRc, D1-MRs, and D1-MA vary the payoffs of Game D1 to induce a change in Player 1's behavior. In particular, Game D1-LR lowers

Player 1's L payoff from \$9.75 to \$7. This lowers $\hat{p}(D|R)$ (the risk neutral $\hat{p}(D|R)$ decreases to 0.57) and makes it "less risky" to choose R. As a result, Player 1 is less likely to select L. In addition, Game D1-MRs raises $\pi_2(L)$ from \$3 to \$6, so that $\pi_2(L)$ becomes greater than $\pi_2(R,U)$ (\$4.75) and $\pi_2(R,D)$ (\$5). This induces "resentment" in Player 2 and likely makes him "retaliate" by choosing U. Hence, Player 1 is more likely to select L. Thirdly, Game D1-MRc raises Player 2's potential payoffs from \$5 ($\pi_2(L)$) to around \$10 ($\pi_2(R,U)$ and $\pi_2(R,D)$), making it more likely that Player 2 would "reciprocate" by choosing R. This added motivation would let Player 1 be less likely to select L. Finally, Game D1-MA lowers $\pi_2(R,U)$ from \$4.75 to \$3, which increases the cost of Player 2 mistakenly choose U instead of D. This increases Player 1's "assurance" that Player 2 would choose D, so he is less inclined to choose the secure option L. To sum up, we have:

Hypothesis 1. *Compared with Game D1, Player 1 is*

a. less likely to select L in Game D1-LR since choosing R now involves "less risk" for himself.

b. more likely to select L in Game D1-MRs since choosing R now induces "more resentment" for Player 2.

c. less likely to select L in Game D1-MRc since choosing R now creates "more reciprocity" for Player 2.

d. less likely to select L in Game D1-MA since he now has "more assurance" that Player 2 would obey dominance.

Goeree and Holt (1999) introduce Game D2, D3 and their variants to test similar hypotheses regarding assurance. In particular, Game D2 is the second baseline game with low threshold probability, in which most Player 1 subjects would choose R.[3] Compared with Game D2, Game D2-LA raises $\pi_2(R,U)$ from \$1 to \$4.75, lowering the assurance that Player 2 would choose D. Hence, we predict that:

Hypothesis 2. *Player 1 is more likely to select L in Game D2-LA than in Game D2 since he now has "lower assurance" that Player 2 would obey dominance.*

Similarly, Game D3 is the third baseline game with intermediate threshold probability, so the fraction of L choices by Player 1 subjects is expected to be lower than that in Game D1 but higher than that in Game D2. Starting from Game D3, Game D3-LA lowers the assurance that Player 2 would choose D by raising $\pi_2(R,D)$ from

[3] In addition, Game D2 and D2-LA also induce resentment for Player 2 since $\pi_2(L)$ is greater than $\pi_2(R,D)$ and $\pi_2(R,U)$ in both games.

\$1 to \$6.75. Game D3-VLA further lowers this assurance by multiplying all payoffs of Game D3-LA by approximately 5, making the difference between $\pi_2(R, D)$ and $\pi_2(R, U)$ only 0.7%, though still \$0.25 in absolute terms. As a result, we have:

Hypothesis 3. *Since the assurance that Player 2 would obey dominance is decreasing, the likelihood that Player 1 selects L increases across Game D3, D3-LA, and D3-VLA.*

Game RP is a rational punishment game, in which Player 2 has little incentive to violate dominance. In this game, if Player 1 chooses L, Player 2 can earn \$4. In contrast, if Player 1 chooses R, Player 2 can only earn \$3 by choosing U (giving Player 1 \$0) and \$0 by choosing D (giving Player 1 \$14). Thus, the choice U by Player 2 is not only a rational response but also a punishment for Player 1's R choice. As a result, Player 2 has little incentive to deviate from the SPE prediction, U, and most Player 1 subjects might respond it by choosing L to earn \$6 for sure.

Game RP-VLR involves very low risk of choosing R, so some Player 1 subjects would choose R. Compared with Game RP, Game RP-VLR considerably decreases the risk of choosing R by lowering $\pi_1(L)$ from \$6 to \$1. Actually, the threshold probability for a risk neutral Player 1 is only 0.08. Therefore, some Player 1 subjects would choose R, hoping to meet the irrational choice D by Player 2.

Player 1's choice in Game TG reflects his belief about the reciprocal behavior by Player 2 subjects. In this game, SPE predicts the outcome (L, U), letting Player 1 and 2 earn \$4 and \$1, respectively. However, Player 1 can express his trust on Player 2 by choosing R, which increases Player 2's potential payoffs ($\pi_2(R, U) = \$10$ and $\pi_2(R, D) = \$9$), expecting to receive the reciprocal choice D by Player 2. Since the payoff augmentation from Player 1's R choice is high (increases from \$1 to at least \$9) and the costs of choosing the reciprocal choice D is low (the difference between $\pi_2(R, D)$ and $\pi_2(R, U)$ is only \$1), some Player 1 subjects would believe that Player 2 would reciprocate his trust, and hence choose R.

Game TG-LRc lowers both Player 2's potential payoffs and Player 1's threshold probability, making it unclear which direction would Player 1's choice move. On the one hand, Player 2's potential payoffs decrease from \$9-10 to \$2-3. This would deter Player 2's willingness to reciprocate Player 1. On the other hand, the threshold probability for a risk neutral Player 1 is only 0.22, so some Player 1 subjects might still select the risky option R.

Game TG-CR substantially increases the cost of repayment for Player 2, so most Player 1 subjects would choose L. Compared with Game TG and TG-LRc, the costs of choosing D by Player 2 extensively increases from \$1 to \$9. This astronomical cost decreases the likelihood of reciprocal behavior from Player 2. Consequently, we expect most Player 1 subjects would follow the SPE prediction by choosing L.

2.2 Simultaneous Spatial Beauty Contest Games

Chen, Huang and Wang (2013)'s simultaneous SBC game is a spatial variant of Costa-Gomes and Crawford (2006)'s asymmetric two-person guessing game. In the original asymmetric two-person guessing game, one player would like to choose a number which equals to α times his opponent's choice and his opponent would like to choose a number which equals to β times his choice. In the simultaneous SBC game, two players are asked to choose locations instead of numbers simultaneously on a two-dimensional grid map to hit their target locations. One's target location is defined as a relative location to the opponent's choice of location by a pair of coordinates (a, b) in the standard Euclidean coordinate. For instance, $(0, 2)$ means a player's target location is "two squares above the opponent's choice of location," and $(-4, 0)$ means a player's target location is "four squares to the left of the opponent's choice of location." Targets of both players are common knowledge.

Payoffs are determined by how "far" a player's choice of location is away from his target location. Specifically, suppose player i chooses (x_i, y_i) with the target (a_i, b_i), and his opponent $-i$ chooses (x_{-i}, y_{-i}). The payoff to player i is determined by the following equation:

$$p_i(x_i, y_i; x_{-i}, y_{-i}; a_i, b_i) = \bar{s} - \lambda(|x_i - (x_{-i} + a_i)| + |y_i - (y_{-i} + b_i)|)$$

where $\bar{s}$ and λ are constants,[4] and $(x_{-i} + a_i, y_{-i} + b_i)$ is the target location for player i. Note that the target location may not be available. For example, consider a player who is assigned to choose a location on a 7×7 grid map with the target $(4, 0)$. For the purpose of illustration, suppose the player's opponent has chosen the center location $((0, 0))$. Then, to hit his target, the ideal choice/response is $(4, 0)$. However, location $(4, 0)$ is not available since it is outside the map. Among all 49 feasible choices of locations on the map, location $(3, 0)$ is the optimal choice of location since it is the only feasible location that is one square from the ideal response (target location) $(4, 0)$.

Table 2 lists the 6 simultaneous SBC games used in the experiment. In these games, both players have one-dimensional targets, one horizontal, one vertical. To report Player 2 subjects' behavior, we also define the sister game, Game SBC-mR, to be the same as Game SBC-m (where $m = 1, 2, \ldots, 6$) but with reversed roles for

[4] In our experiment, $\bar{s}$ is 10 and λ is 0.5.

Table 2: Simultaneous Spatial Beauty Contest Games and Their Theoretical Predictions

Game	Map Size	Player 1 Target	Player 2 Target	Player 1 Choice of $L0$	$L1$	$L2$	$L3$	NE	*Soph*	$\bar{k}$
SBC-1	9×9	-2, 0	0,-4	0, 0	-2, 0	-2,-4	$\underline{-4,-4}$	$\underline{-4,-4}$	-4,-3	3
SBC-2	7×7	2, 0	0,-2	0, 0	2, 0	2,-2	3,-2	3,-3	3,-2	4
SBC-3	11×5	2, 0	0, 2	0, 0	2, 0	2, 2	4, 2	5, 2	4, 2	5
SBC-4	9×7	-2, 0	0,-2	0, 0	-2, 0	-2,-2	-4,-2	-4,-3	-3,-3	4
SBC-5	7×9	-4, 0	0, 2	0, 0	-3, 0	$\underline{-3, 2}$	$\underline{-3, 2}$	-3, 4	-3, 2	4
SBC-6	7×9	2, 0	0, 2	0, 0	2, 0	2, 2	3, 2	3, 4	3, 3	5
SBC-1R[a]	9×9	0,-4	-2, 0	0, 0	0,-4	$\underline{-2,-4}$	$\underline{-2,-4}$	$\underline{-4,-4}$	$\underline{-4,-4}$	4
SBC-2R	7×7	0,-2	2, 0	0, 0	0,-2	2,-2	2,-3	3,-3	2,-3	4
SBC-3R	11×5	0, 2	2, 0	0, 0	0, 2	$\underline{2, 2}$	$\underline{2, 2}$	5, 2	4, 2	6
SBC-4R[b]	9×7	0,-2	-2, 0	0, 0	0,-2	-2,-2	-2,-3	$\underline{-4,-3}$	$\underline{-4,-3}$	4
SBC-5R	7×9	0, 2	-4, 0	0, 0	0, 2	-3, 2	$\underline{-3, 4}$	$\underline{-3, 4}$	$\underline{-3, 4}$	3
SBC-6R	7×9	0, 2	2, 0	0, 0	0, 2	2, 2	2, 4	3, 4	2, 4	5

[*] Non-separating types are underlined.
[a] In Game SBC-1R, $L2$ and $L3$ make identical predictions, and so does NE and *Soph*.
[b] Besides $(-4,-3)$, $(-3,-3)$ is also a *Soph* prediction in Game SBC-4R.

the two players.[5] For example, Game SBC-1R is identical to Game SBC-1, Game SBC-2R is identical to Game SBC-2, and so on.

Chen, Huang and Wang (2013) adopt the level-k model to explain the results of simultaneous SBC games. In particular, they assume that a $L0$ player would randomly choose any location on the map, which is on average the center $(0, 0)$. To best respond to a $L0$ player, a $L1$ player with the target (a, b) would choose the location (a, b), or the nearest feasible location if (a, b) is outside the map. Similarly, a $L2$ player with the target (c, d) plays best response to a $L1$ player who chooses (a, b), by choosing (closest to) $(a + c, b + d)$.[6] A $L3$ player best responds to a $L2$ player, and so on. Chen, Huang and Wang (2013) show that there exists a smallest positive integer $\overline{k}$ such that for all $k \geq \overline{k}$, the level-k predictions are all the same, making them mutual best responses, or the Nash equilibrium (NE). For example, Figure 2 shows the various level-k predictions of Game SBC-2. Specifically, the predictions for Player 1 with target $(2, 0)$ are $L1_1$, $L2_1$, $L3_1$, and E_1; the predictions for Player 2 with target $(0, -2)$ are $L1_2$, $L2_2$, $L3_2$, and E_2. O represent the prediction of $L0$ for both players. Notice that Lk_1 (Lk_2) are the best responses to $L(k-1)_2$ ($L(k-1)_1$), and so on. For example, $L2_1$'s choice $(2, -2)$ is the best response to $L1_2$, since $(0, -2) + (2, 0) = (2, -2)$. For $k \geq 4$, the level-k predictions of both players coincide with the NE predictions.

[5] These games are adopted from Game 1 to 12 of Chen, Huang and Wang (2013).
[6] To ensure uniqueness, in all our games, we have $a + c \neq 0$ and $b + d \neq 0$.

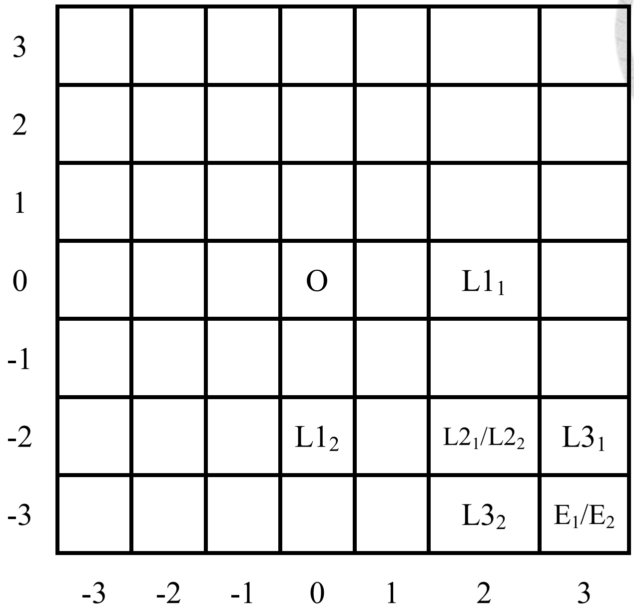


Figure 2: Level-k and NE Predictions of a 7×7 Simultaneous SBC Game with Targets $(2, 0)$ (Player 1) and $(0, -2)$ (Player 2) (Game SBC-2).

In addition to the Lk and NE types, we also define the *Sophisticated* (*Soph*) type to capture the possibility that some subjects have a prior understanding of others' decisions. A *Soph* player has a precise belief about others' decisions, and best responds to the empirical distribution of the opponents' decisions. The *Soph* prediction of each game is presented in the next-to-last column of Table 2. Note that the *Soph* prediction coincides with NE when (most) players play NE.[7]

2.3 First-Mover Spatial Beauty Contest Games

The 1st-mover SBC game is a sequential variant of the simultaneous SBC game. In the simultaneous SBC game, two subjects play against each other and choose simultaneously. Notwithstanding, in the 1st-mover SBC game, each subject chooses individually, then a computerized player who is preprogrammed to maximize its own profit reacts and plays best response. This design controls for subjects' beliefs about the opponent's level of reasoning, and their decisions hence only reflect the ability to play best response and perform backward induction.

Given the same targets and map size, the equilibrium prediction of a 1st-mover SBC game may differ from the simultaneous one. For instance, consider a SBC

[7]In our study, only *Soph* predictions of Game SBC-1R, SBC-4R, and SBC-5R are identical to the NE predictions. However, the *Soph* predictions of the remaining games are also close, being at most two squares away.

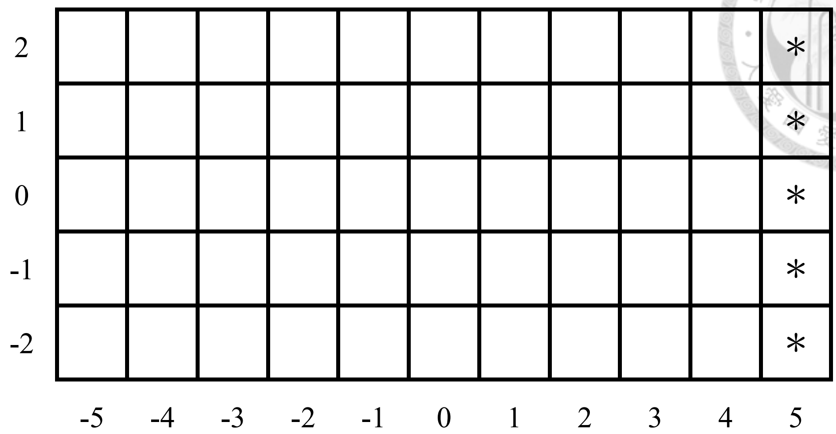


Figure 3: Optimal Choices of a 11×5 1st-Mover SBC Games with Targets $(0, 2)$ (First Mover) and $(2, 0)$ (Second Mover) (Game 1st-3R).

game with targets $(0, 2)$ and $(2, 0)$ for both players on a 11×5 grid map. If both players choose simultaneously (Game SBC-3/3R), there is an unique NE, $((5, 2), (5, 2))$. However, as shown in Figure 3, the sequential variant of this game (Game 1st-3R) with targets $(0, 2)$ for the first mover (subject) and $(2, 0)$ for the second mover (computer) has 4 other SPE (all labeled with $*$). In fact, if the first mover chooses (l, m), the computerized second mover would play best response by choosing $(\min(5, l + 2), m)$, which is $(l + 2, m)$ provided that it is on the map and $(5, m)$ otherwise. Hence, the first mover's ideal choice would be $(\min(5, l+2), m+2)$. By backward induction, the first mover would choose $(5, m)$ to minimize the distance between his/her choice (l, m) and ideal choice $(\min(5, l + 2), m + 2)$. In other words, among all feasible 55 choices of locations, locations $(5, -2)$, $(5, -1)$, $(5, 0)$, $(5, 1)$, and $(5, 2)$ are optimal for the first mover.

We derive the SPE predictions for the general case as follows. Consider a 1st-mover SBC game with target (a_1, b_1) for the first mover and (a_2, b_2) for the second mover. Suppose the first mover chooses location (x_1, y_1) on a map $G \equiv \{-X, -X + 1, \ldots, 0, \ldots, X\} \times \{-Y, -Y + 1, \ldots, 0, \ldots, Y\}$, where X and Y are positive integers and $(0, 0)$ is the center of the map.[8] Then, the choice (x_2, y_2) of the computerized profit-maximizing second mover can be characterized by the following "boundary-adjusted" best-response function:

$$\begin{aligned}(x_2, y_2) &= BR(X, Y; x_1, y_1; a_2, b_2)\\ &= (\min\{X, \max\{-X, x_1 + a_2\}\}, \min\{Y, \max\{-Y, y_1 + b_2\}\})\end{aligned}$$

[8]For example, $(x_1, y_1) = (X, Y)$ means the first mover chooses the Top-Right corner of the map.

Like simultaneous SBC games, there is no interaction between the choices of x_i and y_i in 1st-mover SBC games. Hence, first mover's maximization can be obtained by choosing x_i and y_i separately. We thus focus on the case for x_i. The case for y_i is analogous. Without loss of generality, we assume that $a_2 \geq 0$. If $a_1 > -a_2$, the first mover can maximize his payoff by inducing the second mover to choose the upper bound, X. Hence, the SPE (x_1^e, x_2^e) is $x_1^e = X + \min(a_1, 0)$ and $x_2^e = X$. Note that when $a_1 \geq 0$, $(x_1^e, x_2^e) = (X, X)$.[9] In contrast, if $a_1 \leq -a_2$, the first mover can only lower the distance between his choice and the second mover's choice to a_2 instead of $|a_1|$. Hence, $(-X, -X + a_2), (-X + 1, -X + 1 + a_2), \ldots,$ and $(X - a_2, X)$ are all SPE. Note that if $a_2 = 0$, the second mover chooses the same location as the first mover, making $(-X, -X), (-X + 1, -X + 1), \ldots, (X, X))$ all SPE. To sum up, we obtain:

Proposition 1. Consider a 1st-mover spatial beauty contest game with target (a_1, b_1) for the first mover and (a_2, b_2) for the second mover. Without loss of generality, we assume $a_2, b_2 \geq 0$. Suppose the first mover and the second mover choose locations (x_1, y_1) and (x_2, y_2) on the map $G \equiv \{-X, -X + 1, \ldots, 0, \ldots, X\} \times \{-Y, -Y + 1, \ldots, 0, \ldots, Y\}$, $-2X \leq a_1, a_2 \leq 2X$ and $-2Y \leq b_1, b_2 \leq 2Y$. The SPE $\big((x_1^e, y_1^e), (x_2^e, y_2^e)\big)$ of this game can be characterized by:

$$(x_1^e, x_2^e) \in \begin{cases} \{(X + \min(a_1, 0), X)\} & \text{(unique) if} \quad a_1 > -a_2 \\ \{(-X, -X + a_2), \ldots, (X - a_2, X)\} & \text{if} \quad a_1 \leq -a_2 \end{cases}$$

$$(y_1^e, y_2^e) \in \begin{cases} \{(Y + \min(b_1, 0), Y)\} & \text{(unique) if} \quad b_1 > -b_2 \\ \{(-Y, -Y + b_2), \ldots, (Y - b_2, Y)\} & \text{if} \quad b_1 \leq -b_2 \end{cases}$$

Table 3 presents the 12 1st-mover SBC games used in the experiment and the first mover's optimal choices for these games. Game 1st-3R to 1st-6R are 6 games with one-dimensional targets (1st-1D SBC games). Game 1st-7 to 1st-12 are 6 games with two-dimensional targets (1st-2D SBC games). The 6 1st-1D games are sequential variants of the original simultaneous SBC games in Table 2. Game 1st-3R is the sequential variant of Game SBC-3R, Game 1st-4 is the sequential variant of Game SBC-4, and so on. Game 1st-7 to 1st-11 are sequential variants of Game 13, 16, 19, 22, and 24 in Chen, Huang and Wang (2013). Game 1st-12 is a spacial game in which the uniqueness condition of Proposition 1 are satisfied on *both* dimensions $(a_1 > -a_2, b_1 > -b_2)$, so the number of optimal choices reduces to one.

[9] If $-a_2 < a_1 \leq 0$, the first mover can exactly hit his target by choosing the ideal location $X + a_1$. However, if $a_1 > 0$, the first mover can only choose the upper bound, X, which is a_1 squares from his ideal location, $X + a_1$.

Table 3: First-Mover Spatial Beauty Contest Games and Their Theoretical Predictions

Game	Map Size	1st Mover Target	2nd Mover Target	1st Mover Optimal Choice(s)
1D Targets				
1st-3R	11×5	0, 2	2, 0	$(5, m)$, $m = -2, -1, \ldots, 2$
1st-4	9×7	-2, 0	0,-2	$(l, -3)$, $l = -4, -3, \ldots, 4$
1st-5	7×9	-4, 0	0, 2	$(l, 4)$, $l = -3, -2, \ldots, 3$
1st-5R	7×9	0, 2	-4, 0	$(-3, m)$, $m = -4, -3, \ldots, 4$
1st-6	7×9	2, 0	0, 2	$(l, 4)$, $l = -3, -2, \ldots, 3$
1st-6R	7×9	0, 2	2, 0	$(3, m)$, $m = -4, -3, \ldots, 4$
2D Targets				
1st-7	9×9	-2,-6	4, 4	$(2, m)$, $m = -4, -3, \ldots, 0$
1st-8	7×7	4,-2	-2, 4	$(l, 1)$, $l = -1, 0, \ldots, 3$
1st-9	9×7	-6,-2	4, 4	$(l, 1)$, $l = -4, -3, \ldots, 0$
1st-10	7×9	4, 2	-2,-4	$(l, -2)$, $l = -1, 0, \ldots, 3$
1st-11	7×9	4,-4	-2, 6	$(l, 0)$, $l = -1, 0, \ldots, 3$
1st-12	11×5	-2, 4	6, 2	$(3, 2)$

3 Experimental Design

The experiments were conducted with graphic user interfaces using version 3.3.11 of Zurich Toolbox for Readymade Economic Experiments (z-Tree, Fischbacher, 2007) at the California Social Science Experimental Laboratory (CASSEL) in University of California, Los Angeles (UCLA). Students were recruited via CASSEL's online recruiting website. A total of 6 sessions were run between April 17, 2012 and April 19, 2012, in which 144 UCLA undergraduate students participated.

Each session consisted of four classes of games. Upon arrival at the laboratory, subjects were instructed to sit at separate computer terminals. Subjects were not given any paper instructions. All instructions were projected on the screen and read aloud by the experimenter. Graphical user interfaces and practice-rounds were provided to ensure that all subjects had understood the rules of each class of games. Subjects played (in order) 15 dominance-solvable games (with one practice-round as Player 1 against a computerized Player 2 who chooses D), 10 third-party punishment games,[10] 6 simultaneous SBC games (either playing the same role twice or switching to play both roles once)[11] with 10 second-mover (2nd-mover) SBC

[10] In this paper, we do not discuss the results of third-party punishment games.

[11] 54 of the 72 Player 1 subjects (in dominance-solvable games) played Game SBC-1 to SBC-6 twice as Player 1, and 54 of the 72 Player 2 subjects played each game twice as Player 2 (or Game SBC-1R to SBC-6R as Player 1). For these subjects, we adopt their first-time choices. The remaining 36 subjects, 18 Player 1s and 18 Player 2s, switched and played both roles in the 6 SBC

games as practice,[12] and 12 1st-mover SBC games. Subjects formed groups of three in the third party punishment games, and groups of two in the dominance-solvable games and the simultaneous SBC games. They played individually in the 1st-mover SBC games. Subjects remained the same role in the third party punishment games, and dominance-solvable games.[13] To avoid possible order and learning effects, games (within each class) were presented randomly to each subject and no feedback was provided.[14]

At the end of the session, one game of each class of games was randomly selected and played out against randomly matched opponents to determine subjects' earnings. When announcing the results, we first show subjects' own decisions in the selected game. Then, subjects were informed about the other player's choice and consequently their payoffs. Subjects' total earnings were the sum of payoffs in one randomly selected game in each of the four classes plus a \$5 show-up fee. The average subject earned US\$33.4, ranging from US\$20 to US\$72.5.

4 Basic Results

4.1 Results of Dominance-Solvable Games

The experimental results for all 15 dominance-solvable games are summarized in Table 4. We first note that 43.3% of the Player 1 subjects violate the SPE prediction by choosing L (R) in the first 10 (last 5) games. The frequency of SPE violation varies from 15% (Game RP) to 78% (Game D1-MRs). On the other hand, the average frequency of choices violating dominance by Player 2 subjects is 15.8%, varying from 1% (Game D2 and D3) to 46% (Game TG). These results show that the SPE predictions do not fare particularly well for Player 1 (though Player 2 subjects obey dominance most of the time), and subject decisions indeed vary a lot across games.

Next, we turn to test the predictions discussed in Section 2.1. In our study, we have *within-subject* results for all games. Accordingly, we can compare changes in

games.

[12]Since the rules of simultaneous SBC games are complicated, we employed 10 2nd-mover SBC games as practice rounds, in which subjects chose after seeing a "pre-programmed" computer agent's decision. The computer agent was programmed to always choose the Top-Left corner on the map. The target location (which may be outside the map) and the optimal location were shown in the end of each practice round. Table 6 presents the game structure, the optimal choice of location, and the result of each 2nd-mover SBC game. Subjects played these games in the same order.

[13]Instructions for the simultaneous SBC games were symmetric with labeling either subject as Player 1 or 2. In fact, players were simply referred to as "You" and "Other."

[14]To make sure subjects understood the rules of the game, results of the practice rounds were shown after the decision, and they were all presented in the same order to the subjects.

Table 4: Frequency of Subjects' Choices in the Dominance-Solvable Games

	Payoffs: (Player 1, Player 2)			Frequency (%) of [a]						$\hat{p}(D\|R)$[b]	Em.[c]	McNemar	McNemar	Z
Game	L	(R,U)	(R,D)	L	D	LL	LR	RL	RR	(%)	BR	Stat.	Exact p	Stat.
D1	$(9.75,3)$	$(3,4.75)$	$(10,5)$	58	79					96	L			
vs. D1-LR	$(7,\cdot)$	$(\cdot,\cdot)$	$(\cdot,\cdot)$	38	76	33	25	4	38	57	R	10.71	0.0015**	−2.50*
vs. D1-MRs	$(\cdot,6)$	$(\cdot,\cdot)$	$(\cdot,\cdot)$	78	75	53	6	25	17	96	L	8.91	0.0043**	2.50*
vs. D1-MRc	$(\cdot,5)$	$(5,9.75)$	$(\cdot,10)$	50	92	42	17	8	33	95	L	2.00	0.2379	−1.00
vs. D1-MA	$(\cdot,\cdot)$	$(\cdot,3)$	$(\cdot,\cdot)$	56	85	44	14	11	31	96	L	0.22	0.8145	−0.34
D2	$(7,6)$	$(6,1)$	$(9,5)$	25	99					33	R			
vs. D2-LA	$(\cdot,\cdot)$	$(\cdot,4.75)$	$(\cdot,\cdot)$	28	89	14	11	14	61	33	R	0.22	0.8145	0.38
D3	$(8,5)$	$(2,1)$	$(9,7)$	33	99					86	R			
vs. D3-LA	$(\cdot,\cdot)$	$(\cdot,6.75)$	$(\cdot,\cdot)$	47	85	24	10	24	43	86	L	4.17	0.0639	1.70
vs. D3-VLA	$(40,25)$	$(10,34.75)$	$(45,35)$	57	79	24	10	33	33	86	L	9.32	0.0033**	2.85**
D3-LA	$(8,5)$	$(2,6.75)$	$(9,7)$	47	85					86	L			
vs. D3-VLA	$(40,25)$	$(10,34.75)$	$(45,35)$	57	79	35	13	22	31	86	L	1.96	0.2295	1.17
RP	$(6,4)$	$(0,3)$	$(14,0)$	85	6					43	L			
vs. RP-VLR	$(1,13)$	$(0,4)$	$(14,0)$	56	8	50	35	6	10	7	R	15.21	0.0001**	−3.82**
TG	$(6,4)$	$(0,3)$	$(14,0)$	56	46					44	R			
vs. TG-LRc	$(2,0)$	$(0,3)$	$(9,2)$	43	28	35	21	8	36	22	R	3.86	0.0784	−1.50
vs. TG-CR	$(3,0)$	$(0,10)$	$(8,1)$	81	7	51	4	29	15	38	L	13.50	0.0003**	3.22**
TG-LRc	$(2,0)$	$(0,3)$	$(9,2)$	43	28					22	R			
vs. TG-CR	$(3,0)$	$(0,10)$	$(8,1)$	81	7	40	3	40	17	38	L	23.52	0.0000**	4.63**

* two-sided $p < 0.05$, ** two-sided $p < 0.01$.

[a] LL (RR) means Player 1 chooses L (R) in the first game and chooses L (R) in the other game, and so on. The frequency of subjects violating comparative statics are underlined.

[b] The risk neutral threshold probability.

[c] The empirical best response represents a risk neutral Player 1's best response to the empirical choice distribution of Player 2 subjects. The empirical best responses which do not coincide with the SPE predictions are underlined.

subjects' behavior across the 12 pairs of games presented in Table 4. We employ the *exact* McNemar's test to see if Player 1 subjects' decisions are significantly different for any pair of games.[15] The two-sided McNemar's exact p-values for 12 pairs of games are reported in the next-to-last column of Table 4.

All 4 hypotheses regarding Game D1 and its variants are confirmed, though only two of them are statistically significant. First, Game D1-LR lowers the risk of choosing R, so the frequency of L choices by Player 1 subjects significantly decreases from 58% to 38% (two-sided McNemar's exact $p = 0.0015$), confirming Hypothesis 1a. Second, Game D1-MRs creates more resentment for Player 2, inducing 78% of Player 1 subjects to choose L (significantly higher than 58%, two-sided McNemar's exact $p = 0.0043$), even though the frequency of D choices by Player 2 subjects only slightly decreases from 79% to 75%. This confirms Hypothesis 1b. Thirdly, the frequency of the reciprocal choice D by Player 2 subjects increases to 92% in Game D1-MRc, but the frequency of L choices by Player 1 subjects insignificantly decreases from 58% to 50% (two-sided McNemar's exact $p = 0.2379$). Lastly, the frequency of D choices by Player 2 subjects increases to 85% in Game D1-MA. However, the frequency of L choices by Player 1 subjects (56%) is not significantly lower than the 58% in Game D1 (two-sided McNemar's exact $p = 0.8145$). Consequently, we find weak evidence to support Hypothesis 1c and 1d.

Consistent with Goeree and Holt (2001), Game D2, D3, and their variants provide more evidence to support the hypotheses regarding assurance. To begin with, Game D2-LA lowers the assurance that Player 2 would obey dominance, increasing the frequency of L choices by Player 1 subjects from 25% (Game D2) to 28%. This difference is not statistically significant (two-sided McNemar's exact $p = 0.8145$), but the direction is right. In fact, 14% of Player 1 subjects are sensitive to the change of payoffs, they choose R in Game D2, but move to L in Game D2-LA. Similarly, the frequency of L choices by Player 1 subjects increase across Game D3 (33%), D3-LA (47%), and D3-VLA (57%), confirming Hypothesis 3, although only the difference between Game D3 and D3-VLA is statistically significant (two-sided McNemar's exact $p = 0.0033$). Thus, we conclude that in general, Player 1 subjects do respond to the change of assurance that Player 2 would select D.

[15] McNemar's test is like a paired χ^2 test for differences between two correlated proportions. Its test statistics follow a χ^2 distribution with $df = 1$ asymptotically. However, since the number of LR/RL observations in our study is small, the McNemar's statistics may not be well-approximated by the chi-squared distribution. In this case, the exact version of McNemar's test (using a binomial distribution) is employed instead. Notwithstanding, we still report the McNemar's statistics in the third-last column of Table 4. Note that unlike our study, Beard and Beil (1994) conducted their experiment using a *between-subject* design, so they employed the proportion Z test instead. As shown in the last column of Table 4, the proportion Z test yields similar results to that of the exact McNemar's test in our data, but has less power.

Most Player 1 subjects follow the SPE prediction by choosing L in Game RP, while some of them alter their choices from L to R in Game RP-VLR. In Game RP, only 6% of Player 2 subjects choose D, which is much lower than the threshold probability justifying a risk neutral Player 1 to choose R (43%). Hence, 85% of Player 1 subjects respond by choosing L, which is also the SPE prediction. Compared with Game RP, Game RP-VLR considerably lowers the risk of choosing R, inducing 35% of Player 1 subjects to alter their choices from L to R. In fact, the frequency of D choices by Player 2 subjects (8%) is slightly higher than the risk neutral Player 1's threshold probability (7%). This makes choosing R also the empirical best response for a risk neutral Player 1.

Player 1 subjects' frequencies of the entrusting choice R in Game TG, TG-LRc, and TG-CR change according to our predictions. In particular, in Game TG, 44% of Player 1 subjects choose R, and 46% of Player 2 subjects choose the reciprocal choice D. In addition, Game TG-LRc lowers Player 2's potential payoffs when receiving the entrusting choice R, so the frequency of reciprocal behavior D decreases from 46% to 28%. Notwithstanding, since this frequency is still higher than the risk neutral Player 1's threshold probability (22%), the frequency of R by Player 1 subjects increases to 57%, though insignificantly (two-sided McNemar's exact $p = 0.784$). Lastly, in Game TG-CR, since the costs of reciprocation is high ($9), only 7% of Player 2 subjects choose the reciprocal response D. The frequency of R choices by Player 1 subjects drops to 19%, significantly lower than the 44% (57%) in Game TG (TG-LRc) (two-sided McNemar's exact $p < 0.001$).

4.2 Results of Simultaneous/1st-Mover Spatial Beauty Contest Games

Table 5 presents the frequency of Player 1 subjects' choices in the simultaneous SBC games used in our experiment. We use the difference measure (Selten, 1991), which is the choice frequency minus the fraction of choices predicted, to account for the size of the prediction.[16] We have 90 observations in each game, since we have 36 subjects who played both roles and 54 subject who played Player 1 twice (and we only adapt their first-time choices),[17] The average frequency of all Lk choices (column 3) is 39.5%, ranging from 31.1% (Game SBC-3) to 56.7% (Game SBC-5). All of them are statistically significant under a binomial test. This may seem disappointing economically, but if we consider the locations within one location of

[16] For example, the level-k model predicts several cells, while NE predicts only one.

[17] Individual second-time choices are fairly consistent with their first-time choices, though they do not exactly coincide. In fact, 36.7% of them are exactly the same as the first-time choice, and 52.9% (70.5%) of them are one (two) step(s) away. The average difference between the two choices is 1.855 steps.

Table 5: Player 1 Subjects' Choices in the Simultaneous SBC Games

Game	Obs.	Frequency of					Difference Measure				
		Lk	$Lk \pm 1$	NE	$Lk \setminus NE$	*Soph*	Lk	$Lk \pm 1$	NE	$Lk \setminus NE$	*Soph*
SBC-1	90	41.1	58.9	30.0	11.1	6.7	36.2**	40.4**	28.8**	7.4**	5.4**
SBC-2	90	45.6	66.7	31.1	14.4	2.2	35.4**	36.1**	29.1**	6.3**	0.2
SBC-3	90	31.1	46.7	20.0	11.1	2.2	22.0**	17.6**	18.2**	3.8	0.4
SBC-4	90	38.9	62.2	23.3	15.6	1.1	31.0**	35.2**	21.7**	9.2**	−0.5
SBC-5	90	56.7	75.6	38.9	17.8	11.1	50.3**	53.3**	37.3**	13.0**	9.5**
SBC-6	90	34.4	48.9	21.1	13.3	1.1	26.5**	21.9**	19.5**	7.0**	−0.5
SBC-1R†	90	40.0	55.6	25.6	14.4	25.6	35.1**	38.3**	24.3**	10.7**	24.3**
SBC-2R	90	42.2	64.4	18.9	23.3	3.3	32.0**	33.8**	16.8**	15.2**	1.3
SBC-3R	90	36.7	62.2	17.8	18.9	2.2	27.6**	34.9**	16.0**	11.6**	0.4
SBC-4R†	90	35.6	58.9	15.6	20.0	20.0	27.6**	31.9**	14.0**	13.7**	18.4**
SBC-5R†	90	38.9	66.7	22.2	16.7	22.2	32.5**	42.9**	20.6**	11.9**	19.0**
SBC-6R	90	33.3	65.6	12.2	21.1	6.7	25.4**	38.6**	10.6**	14.8**	5.1**
Mean		39.5	61.0	23.1	16.5	8.7	31.8**	35.4**	21.4**	10.4**	6.9**

Note: All results are presented in percentage (%).
* two-sided $p < 0.05$, ** two-sided $p < 0.01$.
† Games in which *Soph* coincide with *NE*.

the Lk predictions, the frequency of "Lk with noises" ($Lk \pm 1$) choices is on average 61.0%, varying from 46.7% (Game SBC-3) to 75.6% (Game SBC-5). Again, all are statistically significant, as shown in the ninth column of Table 5. Note that there is an unusual concentration of NE choices, accounting for 58.5% of the Lk choices (23.1%/39.5%). In fact, all NE choices occur significantly above chance (two-sided binomial test $p < 0.01$). This is very different from most previous studies on the beauty contest game (aka guessing game), and is likely due to the simplicity of the graphic interface and the training through practice rounds. Nonetheless, binomial test results still show that the remaining Lk choices are chosen significantly above random ($p < 0.03$) for all but Game SBC-3, which has $p = 0.124$. In fact, as shown in the sixth column of Table 6, the frequency of best-responses by subjects increases from 32.6% (Game 2nd-I) to 95.8% (Game 2nd-X), indicating that most subjects had understood the rules and learn to play best reponse after 10 rounds of practice. In contrast, as shown in the seventh column of Table 5, the frequency of *Soph* choices is on average 8.7%, with only 3 of 12 games having frequencies above 12%, all of which the *Soph* predictions coincide with NE. In fact, in the remaining 9 games in which *Soph* predictions differ from the NE predictions, only three of them have difference measures significantly greater than zero (one of them above 7%). Hence, we conclude that even though the frequency of NE choices is around 25%, few subjects play best response against the empirical distributions of the opponent choices.

Table 7 shows the frequency of subjects' optimal choices in the 1st-mover SBC

Table 6: Results of 2nd-Mover SBC Games (Practice Rounds)

Game	Map Size	1st Mover Choice	2nd Mover Target	2nd Mover BR	Frequency of BR (%)	Diff. Measure
2nd-I	3×3	-1, 1	0, 1	-1, 1	32.6	21.5**
2nd-II	7×7	-3, 3	-1, 2	-3, 3	40.3	38.2**
2nd-III	7×9	-3, 4	-1,-4	-3, 0	83.3	81.7**
2nd-IV	9×7	-4, 3	4, 2	0, 3	88.9	87.3**
2nd-V	7×9	-3, 4	-2, 1	-3, 4	61.1	59.5**
2nd-VI	7×7	-3, 3	0,-1	-3, 2	94.4	92.4**
2nd-VII	11×5	-5, 2	3, 0	-2, 2	95.1	93.3**
2nd-VIII	9×9	-4, 4	-1, 0	-4, 4	65.3	64.0**
2nd-IX	11×5	-5, 2	4,-2	-1, 0	84.7	82.9**
2nd-X	9×9	-4, 4	2, 1	-2, 4	95.8	94.6**
Mean					74.2	71.6**

Note: Number of observations is 144.
* two-sided $p < 0.05$, ** two-sided $p < 0.01$.

games, with games with one-dimensional targets (1D games) on the left panel and games with two-dimensional targets (2D games) on the right. We have 144 observations for all games since these are individual decisions made against a payoff-maximizing computer. As shown in the left panel of Table 7, 79.2% of subjects' choices are optimal in the 6 1D games, ranging from 74.3% (Game 1st-6R) to 84% (Game 1st-6). However, when targets become two-dimensional in the 6 2D games, the average frequency of subjects' optimal choices decreases to 41.1% , ranging from 36.8% (Game 1st-9) to 46.5% (Game 1st-8) (right panel of Table 7). These results show that most subjects could solve 1D 1st-mover SBC games, but only some subjects could also solve the 2D games.

Table 7: Subjects' Choices in the 1st-Mover SBC Games

1D Game	*Optimal* (SPE)(%)	Diff. Measure	Difference in Deviations	2D Game	*Optimal* (SPE)(%)	Diff. Measure
1st-3R	75.7	66.6**	-2.23	1st-7	43.8	37.6**
1st-4	78.5	64.2**	-1.22	1st-8	46.5	36.3**
1st-5	80.6	69.5**	-1.21	1st-9	36.8	28.9**
1st-5R	81.9	67.6**	-0.64	1st-10	40.3	32.4**
1st-6	84.0	72.9**	-1.40	1st-11	39.6	31.7**
1st-6R	74.3	60.0**	-1.33	1st-12	39.6	37.8**
Mean	79.2	66.8**	-1.34	Mean	41.1	34.1**

Note: Number of observations is 144.
* two-sided $p < 0.05$, ** two-sided $p < 0.01$; the binomial test.

We now compare the subjects' choices in the 6 1D 1st-mover SBC games to that of the simultaneous SBC games which have the same map sizes and targets. We are interested in deviations from the *EQ* prediction in each class of games. Since all horizontal choices are optimal in Game 1st-4, 1st-5, and 1st-6, we consider the vertical distance between subjects' choices and *EQ* predictions in the corresponding simultaneous SBC games (Game SBC-4, SBC-5, and SBC-6). Similarly, we consider only the horizontal distance between subjects' choices and *EQ* predictions in the simultaneous games which two players' roles are reversed (Game SBC-3R, SBC-5R, and SBC-6R) since all vertical choices are optimal. As shown in the last column of the left panel of Table 7, the average difference in deviations between 1st-mover and simultaneous SBC games is -1.34. This indicates that subjects' choices are on average 1.34 squares closer to *EQ* predictions in the 1st-mover SBC games than in the simultaneous ones. In fact, 45% of the subjects do not play *EQ* in the simultaneous SBC games, but choose optimally in the 1st-mover SBC games. This indicates that subjects choose closer to equilibrium when their beliefs about the opponent are controlled.

5 Subjects' Strategic IQ

Given the basic results reported in section 4 are mostly consistent with the literature, we now attempt to identify individual's strategic abilities using their choice sequences. Section 5.1 describes several subject performance indicators, and investigates the correlations between them. Section 5.2 employs principal component analysis to identify components of strategic IQs (SIQs) which explain the variation across subjects' standardized expected payoffs for each game, and interprets them as various strategic abilities.

5.1 Subject Performance Indicators

We define six different performance indicators that reflect the following strategic abilities: the ability to play best response, perform backward induction, form beliefs about others, and perform complicated backward induction on multi-dimensional action space. Table 8 reports the basic statistics of each indicator, and compare them to various benchmarks: The expected scores of $L0$, $L1$, *EQ*, and *Soph* subjects.[18] To make *within-subject* comparisons, we report results only from 72 subjects who were Player 1 in dominance-solvable games. Table 9 list the corresponding

[18] A $L0$ subject chooses randomly; a $L1$ subject best responds to the $L0$ opponent who chooses randomly; an *EQ* subject plays according to the equilibrium; a *Soph* subject knows the exact choice distribution of the opponents in each game and best responds to that distribution.

Table 8: Statistics and Predicted Scores for Each Performance Indicator

Measure	Obs.	Mean	Std.	Min	Max	*L*0 (Rand.)	*L*1	*EQ*	*Soph* (Opti.)
CS-DSG	72	0.78	1.08	0	4	2.00	0	0	0
EV-DSG	72	8.86	0.25	8.21	9.23	8.56	8.58	8.78	9.24
EV-2ndSBC	72	9.14	0.32	7.85	9.40	6.67	–	–	9.40
EV-1st1D	72	8.63	0.36	7.67	8.83	7.93	–	8.83	8.83
EV-1st2D	72	7.92	0.67	6.50	8.83	7.57	–	8.83	8.83
EV-SBC†	72	7.59	0.59	5.95	8.36	6.41	7.65	8.24	8.37

* Non-separating types are underlined.

† A $L0$ subject who randomly chooses in the maps would obtain 6.41 scores; however, a $L0$ subject who always chooses the the center of the maps would obtain 6.83 scores. In this case, two definitions lead to different predictions.

strategic abilities each indicator represents. Figure 4 to 9 show the distribution of each indicator. We discuss them one by one:

Table 9: Corresponding Strategic Abilities Represented by Each Indicator

	Strategic Abilities				
	BR	BI	Belief	Higher Order Belief	2D-BI†
CS-DSG	√				
EV-DSG	√	√	√		
EV-2ndSBC	√				
EV-1st1D	√	√			
EV-1st2D	√	√			√
EV-SBC	√	√	√	√	

† The ability to perform complicated backward induction in 1st-mover SBC games with high dimensional targets.

CS-DSG is each subject's total number of choice pairs which violate the comparative statics predictions of the dominance-solvable games discussed in section 2.1. According to Hypothesis 1, 2, and 3, we have 8 comparative statics predictions and failure to follow these predictions indicates inability to respond to changes in game payoffs. We underline the frequency of Player 1 subjects' choices in each of the 8 choice pairs violating these predictions in Table 4. For instance, Hypothesis 1d predicts that Player 1 is less likely to select L in Game D1-MA than in Game D1. However, 11% of Player 1 subjects choose R in Game D1 but choose L in Game 1D-MA, violating the comparative statics prediction in Hypothesis 1d. In this case, we deem that these subjects do not correctly respond to the change in payoffs, and this choice pair would count toward their CS-DSG scores. Hence, CS-DSG is a

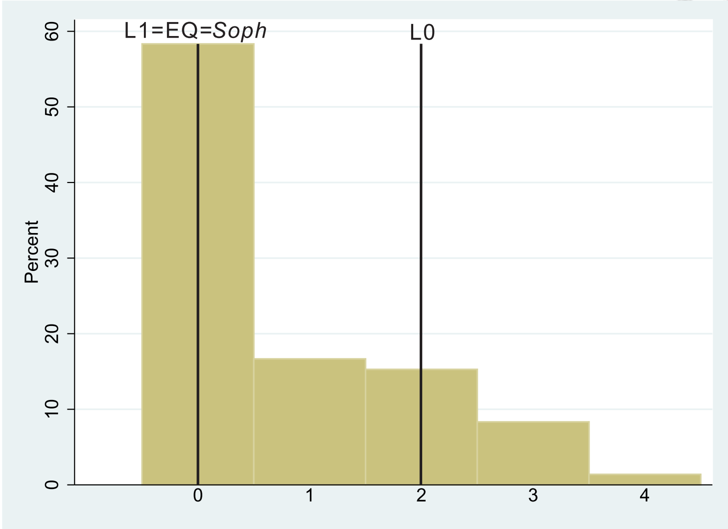


Figure 4: Histogram of CS-DSG (Sample Size = 72)

"counter-indicator."

As shown in the first row of Table 8, the average of CS-DSG is 0.78. This shows that on average less than 1 out of 8 comparative statics predictions are violated, indicating that most subjects are sensitive to the changes in game payoffs and respond to those changes rationally. The distribution of CS-DSG is skewed to right (Figure 4). In particular, more than 58% of subjects do not violate any comparative statics prediction, and none violate more than 4 comparative statics predictions. Only 8.3% (6 out of 72 subjects) violate 3 comparative statics predictions, and only one subject (out of 72) violates 4.[19] Note that since these eight comparative static predictions concern binary decisions that are not independent (especially those of Hypothesis 3), the maximum possible number of violations is 6.

The second indicator is Player 1 subjects' expected earnings averaged across 15 dominance-solvable games (EV-DSG) against the empirical distribution of Player 2 subjects. This measures subject's ability to perform backward induction by forming accurate beliefs about Player 2 subjects' choices and correctly reacting to them. For instance, in Game D1, only 79% of Player 2 subjects choose D. Hence, if Player 1 simply follows the SPE prediction by choosing R, his expected earnings would be \$8.54, lower than the assured payoff by choosing L (\$9.75). So, a risk neutral Player 1, who has the right belief about the frequency of Player 2 choices, would choose L.[20]

[19]It seems that subjects are less sensitive to changes in assurance that Player 2 would obey dominance. In the 5 comparative statics predictions regarding assurance, the frequency of subjects' choices violating the predictions are all greater than 10% in Table 4.

[20]Risk aversion does not play a role in this particular game because the assumed payoff yields

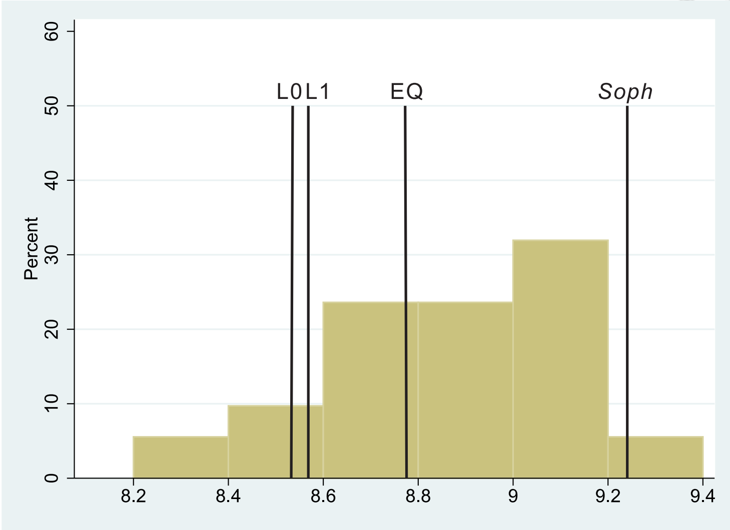


Figure 5: Histogram of EV-DSG (Sample Size = 72)

Figure 5 shows the distribution of EV-DSG. The average is 8.86, ranging from 8.21 to 9.23. Only 12.5% of subjects have EV-DSG scores lower than that of a $L1$ subject (8.58). In contrast, two thirds of the subjects have EV-DSG scores greater than that of an EQ subject (8.78). Moreover, 4 subjects have EV-DSG scores around 9.2–9.23, which is close to the maximum possible, or the expected score of a *Soph* subject (9.24). Thus, we conclude that subjects do not simply choose according to the SPE predictions. Instead, most subjects consider possible deviations of Player 2 subjects.

The remaining four performance indicators in Table 9 reflect various strategic abilities in the three types of SBC games. First, EV-2ndSBC is the average of subject's (hypothetical) earnings in 10 2nd-mover SBC games (practice rounds of simultaneous SBC games), which reflects subject's ability to best respond to a computerized player who always chooses the Top-Left corner on the map. In addition, EV-1st1D is subject's average earnings of 6 1st-mover SBC games with one-dimensional targets against a payoff-maximizing computerized player, reflecting subject's ability to perform backward induction. Thirdly, EV-1st2D represents subject's average earnings of 6 1st-mover SBC games with two-dimensional targets, reflecting subject's ability to perform high dimensional backward induction. Lastly, EV-SBC is subject's average expected earnings of 6 SBC games as Player 1 against the empirical distribution of Player 2 subjects, reflecting their level of reasoning and the accuracy of their belief about the opponent's level of reasoning.

The third row of Table 8 shows the basic statistics of EV-2ndSBC. The average

higher expected value. In other games, risk attitude may affect subject behavior.

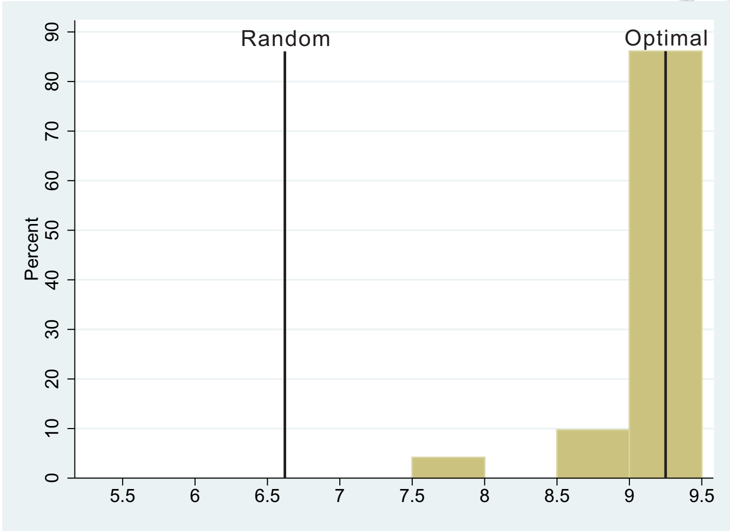


Figure 6: Histogram of EV-2ndSBC (Sample Size = 72)

is 9.14, which is close to the maximum possible (9.4). In fact, as shown in Figure 6, 86% of subjects have EV-2ndSBC scores greater than 9. Moreover, only 4.2% (3 out of 72) of subjects' EV-2ndSBC scores are lower than 8 (the minimum is 7.85), but still much higher than that of a $L0$ subject (6.67). These results indicate that most subjects understand the rules and play best response even without monetary incentives.

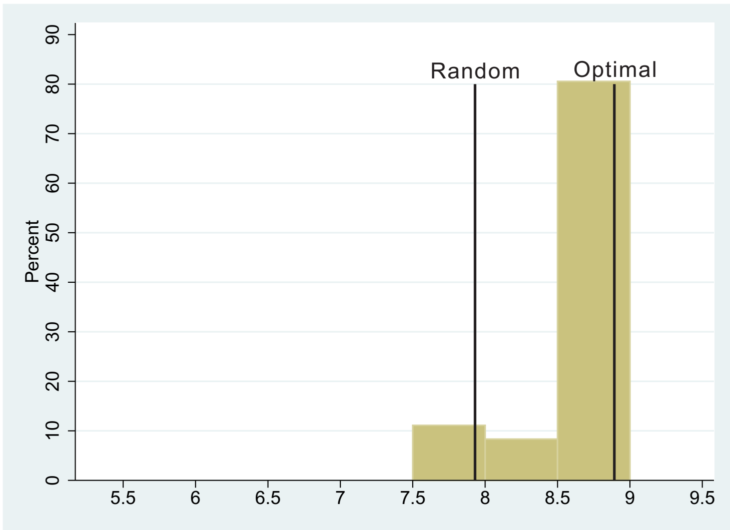


Figure 7: Histogram of EV-1st1D (Sample Size = 72)

Subjects' average EV-1st1D (8.63) is close to that of an optimal subject, indicating that most subjects can perform backward induction and earn the most payoffs. Like EV-2ndSBC, the distribution of EV-1st1D is skewed to left (Figure 7). In particular, 81% of subjects have EV-1st1D scores above or equal to 8.5, which is close to 8.83 (the maximum possible). However, the remaining subjects' average EV-1st1D (7.95) is close to that of a $L0$ subject (7.93), being as low as 7.67.[21]

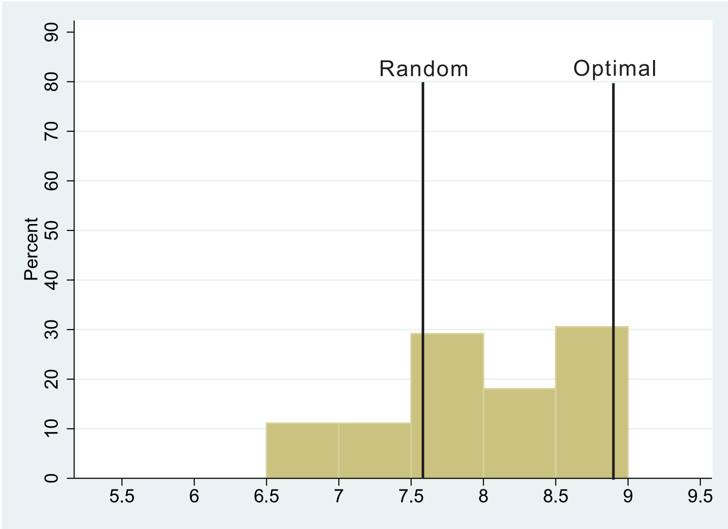


Figure 8: Histogram of EV-1st2D (Sample Size = 72)

The basic statistics of EV-1st2D show the diversity of subjects' ability to perform high dimensional backward induction. In particular, the average of EV-1st2D is 7.92, which is higher than that of a $L0$ subject (7.57) but much lower than that of an EQ subject (8.83). As shown in Figure 8, only 31% of subjects have EV-1st2D scores is close to that of an EQ subject (above or equal to 8.5). The remaining subjects' average EV-1st2D scores (7.58) is close to that of a $L0$ subject (7.57), being as low as 6.5.[22] Moreover, compared with EV-1st1D, EV-1st2D has lower average (7.92 vs. 8.63), higher range (1.16 vs. 2.33), and higher standard deviations (0.67 vs. 0.36). These results show that most subjects can perform backward induction on one-dimensional targets, but some of them fail to do it when there are two-dimensional targets. In particular, 50% of subjects have EV-1st1D scores ≥ 8.5 but EV-1st2D scores < 8.5. Therefore, the frequency of subjects' scores close to that of an optimal subject decreases from 81% (EV-1st1D) to 31% (EV-1st2D).

The average of EV-SBC is 7.59, which is close to $L1$ (7.65) and much higher than that of a $L0$ subject (6.41). Figure 9 shows that the distribution of EV-SBC is skewed to left. In particular, 12.5% of subjects have EV-SBC scores greater than an that of EQ subject (8.24),[23] and only 5.6% (4 out of 72 subjects) score lower

[21] Eight subjects have EV-1st1D scores even lower than that of a $L0$ subject.

[22] In particular, 12.5% of subjects' EV-1st2D scores are even lower than that of a $L0$ subject.

[23] The remaining subjects' average EV-SBC scores is 7.49, still much higher than that of a $L0$ subject (6.41).

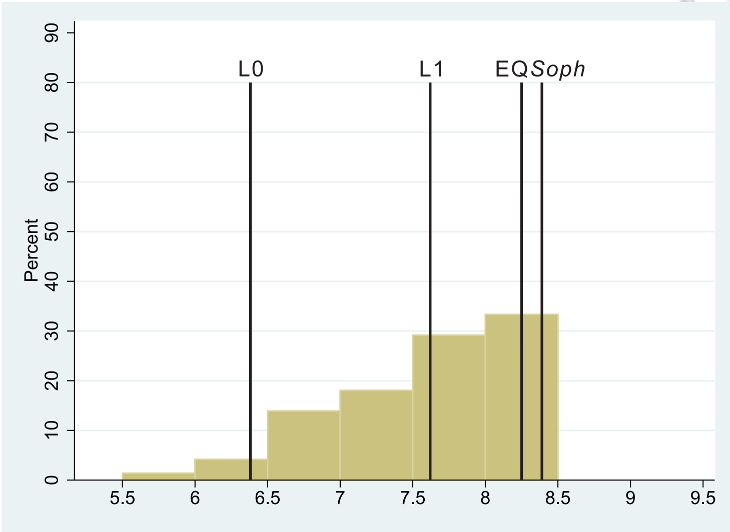


Figure 9: Histogram of EV-SBC (Sample Size = 72)

than that of a $L0$ subject (6.41). This indicates that most subjects do not choose randomly but attempt to earn more payoffs through some process of reasoning. In fact, EV-SBC has lower average and higher standard deviation than EV-1st1D. Specifically, the average of EV-SBC (7.59) is much lower than that of EV-1st1D (8.63), and the standard deviations and range of EV-SBC are 0.59 and 2.41, respectively, which is much higher than those of EV-1st1D (0.36 and 1.16, respectively). These results indicate that subjects' performance become better when we control for their beliefs about the opponents.

Table 10: Correlations Between Indicators

	CS-DSG	EV-DSG	EV-2ndSBC	EV-1st1D	EV-1st2D
CS-DSG	1				
EV-DSG	-0.463^{**}	1			
EV-2ndSBC	-0.054	0.003	1		
EV-1st1D	-0.008	-0.041	0.439^{*}	1	
EV-1st2D	-0.071	0.098	0.157	0.157	1
EV-SBC	0.046	0.040	0.189	0.402^{*}	0.186

* $p < 0.05$, ** $p < 0.01$

We now investigate the correlations between these performance indicators. Results from the Pearson correlation test with the Bonferroni correction (Table 10) show that most indicators are uncorrelated, indicating that the strategic abilities that affect subjects' performance differ across different classes of games. The only exceptions are as follows: First, we find that CS-DSG, a counter indicator of subject performance in dominance-solvable games, is negatively correlated with EV-DSG as predicted ($r = -0.463, p < 0.01$). In addition, EV-2ndSBC is positively correlated with EV-1st1D ($r = 0.439, p < 0.05$), indicating that to perform backward induction requires the ability to play best response. Lastly, EV-SBC is positively correlated with EV-1st1D ($r = 0.402, p < 0.01$), indicating that to perform higher levels of reasoning requires the ability to perform backward induction in 1D 1st-mover SBC games.

5.2 Principal Component Analysis

We employ principal component analysis to explain variation in the normalized expected payoffs of all 33 games in the experiment using a handful of linear combinations, also known as principal components (PC).[24] We normalize the data so that the mean of each variable is always 0 and variance equals to 1, because the relative size of the variances positively affects the weights in principal component analysis. We use 72 observations of Player 1 subjects in 15 dominance-solvable games, 6 simultaneous SBC games, 6 1D 1st-SBC games, and 6 2D 1st-SBC games.

Table 11 presents the entire set of PCs obtained and the corresponding percentage of the total variance of the data explained. The first PC (PC_1) accounts for 21.46% of the total variance of the data, the second PC (PC_2) accounts for 10.77%, and so on. Horn (1965)'s parallel analysis suggests that one should retain all PCs with corresponding variance explained significantly greater than 1 since this means they explain variation of more than one game.[25] This means retaining the first five PCs (PC_1 to PC_5), which account for 56% of the total variance in the data.[26]

The first five PCs could be identified as components of subjects' strategic IQs (SIQs) according to their loadings. The loading of a variable (normalized EV) on a PC is the correlation between this variable and the PC. The higher the loading, the more influential it is in forming the PC, and vice versa. Traditionally, researchers use a threshold of 0.5 to determine whether a given variable is influential in the formation of a PC. We present the loadings of the 5 SIQs in Table 12, and interpret the meanings of each SIQ as follows:

[24]Principal component analysis is a statistical technique of dimension reduction. As linear combinations of the original variables, the first PC accounts for the maximum variance in the data. The second PC accounts for the maximum *remaining* variance that has not been accounted for by the first PC, and so on. Hence, the PCs are uncorrelated among themselves. Ideally, only a few PCs would be needed to account for most of the variance in the data. The mathematic procedure of principal component analysis is provided in the Appendix.

[25]Since there are 33 PCs in total, some PCs would explain variance more than 1, the average variance of one game (out of 33). Hence, in parallel analysis, we simulate 33 iid uncorrelated random variables with mean equal to 0 and variance equal to 1 (each variable has 72 observations), and calculate the corresponding PCs. Using the distribution of these simulated PCs, we can determine whether each PC explains variance *significantly* above 1.

[26]The result of the parallel analysis is reported in the Appendix (Figure A.35).

Table 11: Weights of the Principle Components of Subjects' Normalized EV of the 33 Games

	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}	PC_{11}	PC_{12}	PC_{13}	PC_{14}	PC_{15}	PC_{16}	PC_{17}
D1	−0.15	−0.09	0.31	0.09	0.07	0.06	−0.04	−0.14	0.14	0.17	0.32	0.04	0.23	0.02	0.19	0.02	−0.38
D1-LR	0.14	0.17	−0.32	−0.02	0.01	0.09	0.11	−0.12	−0.03	0.05	−0.30	0.04	0.48	0.05	0.14	0.00	−0.11
D1-MRs	−0.14	0.06	0.27	−0.06	0.12	0.22	0.03	0.07	−0.10	−0.25	−0.03	0.26	0.51	0.15	−0.03	0.31	0.19
D1-MRc	−0.21	0.13	0.28	0.14	0.08	0.26	−0.04	0.06	0.02	−0.07	−0.01	−0.14	0.14	0.03	−0.18	0.05	−0.20
D1-MA	−0.19	−0.04	0.32	−0.05	0.12	−0.20	0.00	−0.06	0.30	0.01	−0.03	0.19	−0.11	0.19	0.13	−0.02	0.13
D2	0.06	0.04	0.06	0.34	−0.10	−0.17	0.02	0.58	0.13	0.12	−0.24	0.16	0.08	0.22	−0.06	−0.10	−0.06
D2-LA	−0.02	0.18	0.08	0.26	−0.18	−0.39	0.17	0.07	−0.10	0.01	−0.17	−0.40	0.14	0.13	0.31	0.12	−0.03
D3	0.17	0.03	−0.19	0.20	−0.23	−0.04	−0.06	0.00	0.34	0.04	0.32	−0.01	0.22	0.24	−0.22	−0.24	−0.12
D3-LA	−0.19	0.03	0.18	0.02	0.02	−0.43	0.04	−0.08	−0.23	−0.34	0.00	−0.08	0.14	−0.03	−0.26	−0.14	0.25
D3-VLA	−0.14	0.04	0.31	0.13	−0.02	0.12	−0.30	−0.13	0.06	0.06	−0.20	−0.34	−0.02	−0.01	−0.12	−0.35	−0.08
RP	0.06	−0.06	0.15	0.29	−0.18	−0.26	0.07	−0.32	0.18	0.26	−0.02	0.14	0.15	−0.47	−0.11	0.16	0.26
RP-VLR	−0.18	0.00	0.16	−0.21	0.20	−0.03	0.32	0.03	−0.15	0.36	−0.13	−0.09	0.00	−0.10	0.11	−0.16	−0.26
TG	−0.05	0.05	−0.09	−0.37	−0.06	−0.23	−0.36	0.16	0.26	0.04	0.03	−0.04	0.30	−0.30	−0.04	0.15	−0.12
TG-LRc	−0.09	0.18	0.01	−0.29	0.11	−0.38	−0.22	−0.01	0.13	−0.02	−0.26	0.17	−0.19	0.10	−0.20	0.05	−0.26
TG-CR	0.08	−0.08	0.13	0.33	−0.19	0.19	−0.20	−0.09	−0.06	−0.13	−0.39	0.30	−0.23	−0.19	0.00	0.15	−0.19
1st-3R	0.29	−0.07	0.19	−0.03	0.15	−0.07	0.14	−0.01	0.19	−0.01	0.03	−0.07	0.04	−0.12	0.19	−0.11	0.18
1st-4	0.32	−0.10	0.12	−0.05	0.20	−0.01	−0.10	0.03	0.08	−0.04	0.00	−0.08	0.00	−0.02	0.14	0.04	−0.06
1st-5	0.24	−0.10	0.11	0.03	0.17	−0.01	0.09	−0.06	0.00	0.37	0.11	−0.06	0.01	0.21	−0.50	0.26	−0.01
1st-5R	0.25	−0.10	−0.01	0.00	0.25	−0.12	−0.04	−0.18	−0.17	0.00	−0.20	0.21	0.17	0.05	−0.18	−0.25	−0.18
1st-6	0.26	−0.11	0.02	0.03	0.32	−0.12	−0.10	−0.02	−0.15	0.00	−0.11	−0.05	0.00	0.23	0.02	0.09	0.10
1st-6R	0.30	−0.07	0.12	−0.03	0.18	−0.04	−0.02	0.13	0.07	−0.16	0.01	−0.17	0.00	−0.23	0.20	0.07	−0.07
1st-7	0.09	0.43	0.02	0.05	0.03	−0.05	0.11	−0.05	−0.11	−0.02	0.14	0.10	−0.05	−0.16	0.04	−0.17	−0.18
1st-8	0.10	0.37	0.07	0.03	0.10	0.14	−0.08	−0.18	0.03	0.23	−0.06	0.05	−0.04	0.14	0.07	−0.26	0.35
1st-9	0.00	0.42	−0.06	0.07	0.01	0.00	−0.19	0.08	−0.17	0.04	0.18	−0.16	−0.08	−0.06	−0.13	0.22	0.08
1st-10	0.15	0.30	0.06	0.07	0.05	−0.04	−0.04	−0.23	0.30	−0.31	0.08	0.12	−0.10	0.22	0.22	0.10	−0.11
1st-11	0.01	0.41	0.01	−0.02	0.11	0.02	0.21	−0.01	0.02	0.16	−0.03	0.23	−0.02	−0.20	−0.07	0.08	−0.03
1st-12	0.22	0.11	0.18	0.15	0.17	0.01	0.04	0.36	−0.09	−0.13	0.20	−0.03	−0.10	−0.22	−0.16	0.04	−0.09
SBC-1	0.24	0.01	0.10	−0.08	−0.24	−0.04	−0.01	−0.33	−0.21	−0.21	0.00	−0.20	0.10	0.02	−0.14	0.02	−0.26
SBC-2	0.15	0.05	0.22	−0.29	−0.22	0.11	0.20	0.11	0.09	−0.23	0.05	0.08	0.02	−0.06	−0.16	−0.35	0.12
SBC-3	0.12	0.10	0.17	−0.27	−0.25	0.12	0.23	−0.06	0.22	0.06	−0.22	−0.25	−0.17	0.19	−0.10	0.34	0.04
SBC-4	0.18	−0.04	0.19	−0.18	−0.38	−0.01	0.18	0.16	−0.08	0.06	−0.08	0.19	0.05	−0.02	0.02	−0.06	−0.10
SBC-5	0.09	−0.02	0.20	−0.10	−0.27	−0.18	−0.20	−0.02	−0.42	0.20	0.31	0.26	−0.08	0.23	0.23	0.06	−0.03
SBC-6	0.17	0.10	0.14	−0.16	−0.10	0.15	−0.47	0.13	−0.08	0.23	−0.14	−0.08	0.14	−0.06	0.11	−0.08	0.23
Variance	7.08	3.55	3.14	2.54	2.13	1.40	1.18	1.14	1.05	0.96	0.90	0.84	0.74	0.70	0.63	0.59	0.52
%	21.46	10.77	9.53	7.69	6.46	4.25	3.57	3.44	3.17	2.91	2.73	2.54	2.25	2.13	1.92	1.79	1.56
Cum. %	21.46	32.23	41.75	49.44	55.90	60.15	63.72	67.16	70.33	73.24	75.96	78.51	80.75	82.88	84.80	86.59	88.16

Table 11: (*Continued*)

	PC_{18}	PC_{19}	PC_{20}	PC_{21}	PC_{22}	PC_{23}	PC_{24}	PC_{25}	PC_{26}	PC_{27}	PC_{28}	PC_{29}	PC_{30}	PC_{31}	PC_{32}	PC_{33}
D1	0.44	−0.16	−0.10	0.15	−0.09	−0.24	−0.15	0.08	0.08	0.10	−0.03	−0.12	0.13	−0.09	0.18	−0.10
D1-LR	−0.16	−0.05	0.02	−0.29	−0.21	−0.27	0.11	0.14	−0.02	−0.24	−0.03	−0.16	0.24	−0.01	0.21	0.06
D1-MRs	−0.14	0.12	−0.20	0.11	0.18	0.23	−0.21	0.01	−0.16	0.06	−0.04	0.04	0.07	0.02	−0.09	−0.05
D1-MRc	0.00	0.13	0.16	−0.26	−0.16	−0.06	0.37	0.15	0.20	0.01	0.05	0.40	−0.35	0.07	−0.02	0.14
D1-MA	−0.18	0.09	0.13	−0.13	−0.35	0.17	−0.02	−0.11	0.31	−0.15	−0.11	−0.34	0.10	0.24	−0.07	0.17
D2	0.10	−0.07	0.09	−0.15	−0.19	−0.06	−0.22	−0.14	−0.21	0.13	0.28	0.01	−0.06	−0.06	−0.08	−0.12
D2-LA	0.15	0.18	−0.19	0.21	0.22	0.07	0.17	0.06	0.14	−0.13	−0.08	−0.13	−0.10	0.04	−0.15	0.10
D3	−0.26	0.00	−0.04	0.09	0.20	0.07	−0.21	0.00	0.24	−0.13	−0.21	0.27	0.04	0.17	0.05	0.03
D3-LA	−0.05	−0.28	0.12	−0.01	0.01	−0.22	−0.09	−0.15	0.20	−0.07	0.01	0.11	0.04	−0.31	0.25	−0.03
D3-VLA	−0.38	−0.09	−0.05	0.24	−0.03	−0.02	0.08	0.09	−0.37	0.00	0.04	−0.14	0.21	−0.01	−0.07	−0.06
RP-VLR	−0.08	0.08	0.24	−0.14	0.34	0.11	−0.22	−0.27	−0.06	−0.17	−0.03	0.21	0.09	0.15	0.10	−0.06
RP	0.09	0.01	0.11	−0.18	0.02	0.06	0.01	0.16	−0.22	0.05	−0.04	0.18	0.12	0.10	−0.13	0.06
TG	−0.12	0.28	0.24	0.14	0.14	−0.19	0.07	−0.16	0.06	0.16	0.15	−0.17	−0.12	−0.05	−0.03	0.01
TG-LRc	0.13	0.01	−0.45	−0.16	0.07	0.06	0.01	0.26	−0.09	0.00	−0.08	0.20	0.08	0.07	0.13	−0.02
TG-CR	−0.06	0.20	0.01	0.12	0.20	−0.13	−0.14	−0.10	0.28	−0.21	0.05	−0.04	0.00	−0.02	0.22	−0.07
1st-3R	−0.17	0.11	−0.20	0.03	−0.09	0.02	−0.05	0.09	−0.15	−0.04	0.07	−0.02	−0.50	0.08	0.51	−0.19
1st-4	0.02	−0.05	−0.04	0.09	−0.02	0.06	−0.17	−0.04	−0.07	−0.18	0.26	0.22	0.11	−0.25	−0.03	0.71
1st-5	−0.02	0.08	−0.10	−0.03	0.08	0.10	0.26	−0.12	0.09	−0.22	0.20	−0.23	0.08	−0.26	0.04	−0.16
1st-5R	0.17	0.13	0.18	0.24	−0.16	0.02	−0.03	−0.04	−0.11	−0.04	−0.43	−0.04	−0.32	−0.10	−0.27	0.01
1st-6	0.02	−0.04	0.34	0.21	0.11	−0.09	0.04	0.32	0.16	0.27	0.20	0.10	0.22	0.42	0.10	−0.10
1st-6R	−0.13	0.04	−0.19	−0.08	−0.23	−0.07	0.03	−0.16	0.20	−0.05	−0.13	0.29	0.29	−0.08	−0.33	−0.43
1st-7	−0.15	0.14	0.16	−0.03	−0.05	0.37	−0.17	0.39	0.20	0.17	0.18	−0.13	0.05	−0.35	−0.03	−0.09
1st-8	0.12	0.32	−0.19	−0.09	0.10	−0.27	0.00	−0.29	0.13	0.36	−0.03	0.08	0.09	−0.10	0.01	0.09
1st-9	0.17	0.21	0.14	0.21	−0.38	0.09	−0.17	−0.15	−0.19	−0.27	−0.12	0.10	0.13	0.19	0.26	−0.03
1st-10	0.07	−0.19	0.32	−0.09	0.26	0.07	0.27	−0.23	−0.31	−0.09	0.01	0.03	−0.01	−0.09	0.01	−0.16
1st-11	−0.16	−0.46	−0.17	0.36	−0.05	−0.18	0.05	−0.10	0.14	−0.07	0.17	0.01	−0.19	0.24	−0.21	0.04
1st-12	−0.09	−0.07	−0.02	−0.25	0.25	−0.20	0.03	0.06	−0.09	0.12	−0.40	−0.34	0.08	0.13	0.06	0.23
SBC-1	0.05	−0.05	−0.13	−0.27	−0.12	0.12	−0.21	−0.30	0.01	0.25	0.25	−0.15	−0.10	0.33	−0.07	0.04
SBC-2	0.34	0.20	0.04	0.05	0.08	−0.17	0.06	0.19	−0.03	−0.36	0.22	−0.07	0.14	0.16	−0.15	−0.07
SBC-3	−0.06	−0.07	0.20	0.05	−0.06	−0.23	−0.34	0.18	0.01	0.12	−0.25	0.03	−0.09	−0.19	−0.06	−0.02
SBC-4	−0.04	−0.07	0.02	0.25	−0.15	0.24	0.41	−0.15	0.03	0.28	−0.17	0.14	0.22	−0.04	0.31	0.12
SBC-5	−0.30	0.08	0.01	−0.11	−0.02	−0.29	0.02	0.12	−0.14	−0.15	0.03	0.10	−0.11	−0.02	−0.14	−0.02
SBC-6	0.21	−0.40	0.06	−0.18	0.13	0.29	−0.04	0.05	0.22	−0.17	−0.13	−0.06	−0.14	−0.01	0.00	−0.10
Variance	0.49	0.48	0.41	0.38	0.35	0.32	0.24	0.23	0.20	0.17	0.17	0.15	0.11	0.08	0.06	0.06
%	1.49	1.45	1.24	1.14	1.07	0.97	0.74	0.71	0.61	0.53	0.51	0.44	0.33	0.25	0.19	0.17
Cum. %	89.65	91.10	92.34	93.48	94.55	95.52	96.26	96.97	97.58	98.11	98.61	99.06	99.39	99.64	99.83	100.00

Table 12: Weights and Loadings of the Five Strategic IQs

	SIQ_1		SIQ_2		SIQ_3		SIQ_4		SIQ_5	
EV	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
D1	−0.15	−0.41	−0.09	−0.16	0.31*	0.56*	0.09	0.15	−0.07	−0.10
D1-LR	0.14	0.37	0.17	0.33	−0.32*	−0.57*	−0.02	−0.03	−0.01	−0.01
D1-MRs	−0.14	−0.37	0.06	0.11	0.27	0.48	−0.06	−0.10	−0.12	−0.17
D1-MRc	−0.21*	−0.56*	0.13	0.24	0.28*	0.50*	0.14	0.22	−0.08	−0.12
D1-MA	−0.19*	−0.51*	−0.04	−0.08	0.32*	0.56*	−0.05	−0.09	−0.12	−0.17
D2	0.06	0.15	0.04	0.07	0.06	0.10	0.34*	0.54*	0.10	0.14
D2-LA	−0.02	−0.05	0.18	0.34	0.08	0.14	0.26	0.42	0.18	0.26
D3	0.17	0.45	0.03	0.05	−0.19	−0.33	0.20	0.31	0.23	0.34
D3-LA	−0.19*	−0.52*	0.03	0.05	0.18	0.31	0.02	0.02	−0.02	−0.04
D3-VLA	−0.14	−0.37	0.04	0.07	0.31*	0.55*	0.13	0.20	0.02	0.04
RP	0.06	0.16	−0.06	−0.12	0.15	0.27	0.29	0.46	0.18	0.26
RP-VLR	−0.18	−0.49	0.00	0.01	0.16	0.28	−0.21	−0.34	−0.20	−0.29
TG	−0.05	−0.13	0.05	0.10	−0.09	−0.16	−0.37*	−0.59*	0.06	0.08
TG-LRc	−0.09	−0.24	0.18	0.35	0.01	0.01	−0.29	−0.46	−0.11	−0.16
TG-CR	0.08	0.21	−0.08	−0.15	0.13	0.23	0.33*	0.53*	0.19	0.28
1st-3R	0.29*	0.77*	−0.07	−0.14	0.19	0.33	−0.03	−0.05	−0.15	−0.23
1st-4	0.32*	0.85*	−0.10	−0.18	0.12	0.21	−0.05	−0.08	−0.20	−0.29
1st-5	0.24*	0.64*	−0.10	−0.19	0.11	0.20	0.03	0.04	−0.17	−0.25
1st-5R	0.25*	0.67*	−0.10	−0.20	−0.01	−0.02	0.00	0.01	−0.25	−0.37
1st-6	0.26*	0.69*	−0.11	−0.20	0.02	0.04	0.03	0.05	−0.32	−0.46
1st-6R	0.30*	0.79*	−0.07	−0.14	0.12	0.20	−0.03	−0.04	−0.18	−0.26
1st-7	0.09	0.24	0.43*	0.82*	0.02	0.03	0.05	0.08	−0.03	−0.05
1st-8	0.10	0.26	0.37*	0.70*	0.07	0.13	0.03	0.05	−0.10	−0.15
1st-9	0.00	−0.01	0.42*	0.79*	−0.06	−0.10	0.07	0.11	−0.01	−0.02
1st-10	0.15	0.39	0.30*	0.57*	0.06	0.10	0.07	0.11	−0.05	−0.07
1st-11	0.01	0.04	0.41*	0.78*	0.01	0.01	−0.02	−0.03	−0.11	−0.17
1st-12	0.22*	0.58*	0.11	0.21	0.18	0.32	0.15	0.23	−0.17	−0.25
SBC-1	0.24*	0.64*	0.01	0.02	0.10	0.17	−0.08	−0.12	0.24	0.35
SBC-2	0.15	0.39	0.05	0.10	0.22	0.38	−0.29	−0.45	0.22	0.32
SBC-3	0.12	0.32	0.10	0.18	0.17	0.30	−0.27	−0.43	0.25	0.37
SBC-4	0.18	0.48	−0.04	−0.08	0.19	0.34	−0.18	−0.29	0.38*	0.56*
SBC-5	0.09	0.23	−0.02	−0.04	0.20	0.36	−0.10	−0.16	0.27	0.39
SBC-6	0.17	0.46	0.10	0.19	0.14	0.24	−0.16	−0.25	0.10	0.15
Variance(%)	21.46		10.77		9.53		7.69		6.46	

Note: Column (1) are the weights of each SIQ, and Column (2) are the loadings of each SIQ.
* Absolute value of loadings greater than the threshold of 0.5.

The first SIQ (SIQ_1) is the component that explains the maximum variance (possible by one single dimension). This would be the data-identified common "g-factor" that predicts subject performance, and we interpret it as subjects' abilities to perform backward induction. This SIQ has loadings of 1st-mover SBC games with 1D targets all greater than 0.5, and the corresponding weights of these games are all between 0.24 to 0.32. In fact, its correlation with performance indicator EV-1st1D is 0.88. Thus, this SIQ is the (weighted) average EV of 6 easy 1st-mover SBC games, which corresponds to subjects' ability to perform backward induction.[27] Moreover, the loadings of dominance-solvable games have signs corresponding to the consistency of SPE and empirical best response. In particular, SIQ_1 has positive loadings on games where the SPE and empirical best response coincide (Games D1-LR, D2, D3, RP and TG-CR), and has negative loadings on games where the SPE and empirical best response differ (Games D1, D1-MRs, D1-MRc, D1-MA, D3-LA, D3-VLA, RP-VLR, TG and TG-LRc). This implies that those who are capable of performing backward induction in the 1st-mover SBC games with 1D targets are also more likely to play SPE in the dominance-solvable games, which is bad for their expected earnings when the empirical best response does not coincide with SPE.[28] This effect is so strong Games D1-MRc, D1-MA, and D3-LA have loadings greater than 0.5. SIQ_1 is also closely related to performance in SBC games, though only Game SBC-1 has loading greater than 0.5. Interestingly, Game 1st-12 has loading equal to 0.58, likely because it is the only game where both players have the same vertical target (of being above the opponent), effectively reducing it to a single dimension game.

The second SIQ (SIQ_2) could be interpreted as subjects' abilities to perform high dimensional backward induction. For all but one 1st-mover SBC games with two-dimensional targets, this SIQ has loadings greater than 0.5. The corresponding weights of these games are mostly between 0.30 to 0.42, so we could interpret this SIQ as subjects' ability to perform high dimensional backward induction. This ability is also reflected in EV-1st2D, which has a correlation of 0.90 with SIQ_2. This shows that our ad hoc performance indicators in Section 5.1 may not be as arbitrary as one may think, although not all games in the same class (with the same format) reflect the same abilities.

The third SIQ (SIQ_3) controls for subjects' attitudes toward risk. This SIQ has high loadings for Games D1, D3 and their variants, which have high risk neutral

[27]Subjects also need to know how to play best response, but the results of 2nd-mover SBC games show that most subjects have the ability to play best response.

[28]The only exception is Game D2-LA, which has a loading of -0.05 (close to zero), but both SPE and empirical B.R. are R for Player 1.

Table 13: Percentiles (%) of each SIQ for the 72 Subjects

Subject ID	SIQ_1	SIQ_2	SIQ_3	SIQ_4	SIQ_5	Subject ID	SIQ_1	SIQ_2	SIQ_3	SIQ_4	SIQ_5
101	37.5	1.4	41.7	56.9	36.1	413	68.1	56.9	34.7	18.1	41.7
102	22.2	34.7	84.7	12.5	8.3	414	80.6	38.9	36.1	45.8	54.2
103	18.1	59.7	29.2	65.3	100.0	415	62.5	45.8	87.5	34.7	59.7
104	26.4	20.8	20.8	58.3	4.2	416	44.4	86.1	33.3	1.4	9.7
105	100.0	83.3	37.5	66.7	86.1	417	77.8	73.6	83.3	70.8	55.6
106	65.3	22.2	68.1	29.2	88.9	418	50.0	54.2	8.3	15.3	80.6
113	45.8	62.5	88.9	23.6	11.1	501	76.4	50.0	47.2	4.2	51.4
114	8.3	66.7	76.4	27.8	98.6	502	84.7	95.8	55.6	19.4	61.1
115	27.8	5.6	62.5	54.2	47.2	503	40.3	12.5	97.2	69.4	63.9
201	51.4	29.2	5.6	22.2	40.3	504	54.2	41.7	4.2	94.4	44.4
202	59.7	11.1	52.8	9.7	45.8	505	56.9	91.7	86.1	86.1	34.7
203	69.4	23.6	50.0	13.9	77.8	506	43.1	75.0	100.0	55.6	23.6
204	12.5	76.4	22.2	47.2	97.2	513	52.8	27.8	80.6	36.1	84.7
205	61.1	55.6	54.2	62.5	38.9	514	72.2	94.4	11.1	75.0	6.9
206	20.8	6.9	1.4	100.0	72.2	515	98.6	77.8	23.6	83.3	62.5
213	2.8	72.2	65.3	51.4	93.1	516	31.9	36.1	81.9	5.6	31.9
214	6.9	84.7	30.6	97.2	87.5	517	13.9	52.8	43.1	6.9	94.4
215	93.1	33.3	12.5	61.1	65.3	518	70.8	43.1	51.4	43.1	68.1
301	47.2	87.5	63.9	95.8	16.7	601	38.9	25.0	98.6	44.4	43.1
302	83.3	100.0	56.9	41.7	66.7	602	94.4	88.9	70.8	76.4	69.4
303	58.3	47.2	79.2	73.6	26.4	603	16.7	37.5	73.6	77.8	91.7
304	15.3	79.2	31.9	90.3	5.6	604	4.2	80.6	18.1	8.3	73.6
305	5.6	51.4	75.0	68.1	95.8	605	1.4	81.9	19.4	37.5	2.8
306	33.3	15.3	93.1	33.3	20.8	606	95.8	48.6	25.0	80.6	83.3
313	81.9	40.3	6.9	2.8	56.9	613	73.6	44.4	94.4	50.0	52.8
314	36.1	65.3	77.8	98.6	30.6	614	55.6	63.9	95.8	63.9	50.0
315	19.4	61.1	38.9	38.9	1.4	615	91.7	93.1	48.6	30.6	76.4
316	86.1	30.6	44.4	59.7	90.3	616	29.2	26.4	58.3	81.9	27.8
317	97.2	31.9	26.4	40.3	75.0	617	75.0	4.2	45.8	16.7	79.2
318	9.7	70.8	16.7	26.4	18.1	618	88.9	16.7	27.8	48.6	70.8
401	63.9	13.9	2.8	93.1	19.4	625	34.7	2.8	61.1	84.7	13.9
402	25.0	8.3	91.7	31.9	25.0	626	30.6	19.4	15.3	52.8	81.9
403	48.6	68.1	90.3	91.7	48.6	627	90.3	90.3	9.7	72.2	33.3
404	23.6	18.1	40.3	11.1	29.2	628	66.7	97.2	72.2	20.8	22.2
405	87.5	69.4	59.7	87.5	58.3	629	11.1	58.3	66.7	88.9	12.5
406	79.2	98.6	13.9	25.0	15.3	630	41.7	9.7	69.4	79.2	37.5

thresholds. It also has coefficients with a negative sign when choosing R lowers one's payoffs in these games, implying that Player 1 subjects who choose the assured choice L would obtain higher SIQ_3 scores. Assuming that subjects could perform backward induction in dominance-solvable games (controlled for by SIQ_1), subjects who choose L in these games due to their attitudes toward risk. Therefore, we could interpret this SIQ as a variable to explain subjects' risk aversion.[29]

The fourth SIQ (SIQ_4) reflects subjects' beliefs about others' social preferences. In particular, the loadings for Game TG and TG-CR for this SIQ are -0.59 and 0.53, respectively, and the loading of the remaining trust game, Game TG-LRc, is -0.46. Since the empirical best response of these three games are choosing R, R, and L, respectively, these loadings imply that subjects who obtain higher SIQ_4 scores are more likely to choose L in the trust games. Therefore, Player 1 subjects who underestimate their opponent's reciprocity (so they always choose L in trust games) would obtain higher SIQ_4. In fact, the loadings of Game D2 are D2-LA are also high (0.54 and 0.42), while the SPE and empirical best response are both R. This means that subjects who obtain a higher SIQ_4 are more likely to choose R in this game, ignoring the possible resentment (negative reciprocity) caused by this action. Thus, SIQ_4 indicates beliefs regarding the likelihood of others (not) reciprocating.

The fifth SIQ (SIQ_5) measures subjects' accuracy of the higher order beliefs about the opponents in the SBC games. This SIQ only has a loading (Game SBC-7) which is greater than 0.5. Nevertheless, the loadings of simultaneous SBC games are all positive. Given SIQ_1 and SIQ_2 already account for subjects' abilities to play best response and perform backward induction in these games, we interpret this SIQ as a measure on subjects' accuracy of higher order belief about their opponents.

The percentiles of each SIQ for the 72 subjects are listed in Table 13.

6 Conclusion

In this paper, we employ dominance-solvable games, simultaneous and 1st-mover spatial beauty contest games to uncover different strategic abilities. The basic response confirm most comparative statics in the literature. We define six indicators on subjects' performance and each represents various strategic abilities. The results of these indicators show the heterogeneity in subject's strategic abilities. First, in the dominance-solvable games, two-thirds of subjects' performance (EV-DSG) are

[29]Alternatively, SIQ_3 could be viewed as reflecting people's belief regarding the likelihood of their opponent's lack of rationality, which is what drives risk averse subjects to choose the assured payoff in DSG games. This interpretation is partially supported by the positive loadings of SBC games, but none of them cross the 0.5 threshold.

even better than that of an EQ subject but there are still subjects who perform even worse than a $L0$ subject. Second, the distributions of EV-2ndSBC and EV-1st1D show that more than 80% of subjects can play best response and perform backward induction. However, with multi-dimensional targets, the frequency of subjects who can perform backward induction reduces to 31% in the 2D 1st-mover SBC games. Moreover, the remaining subjects have an average EV-1st2D close to that of a $L0$ subject. Lastly, when higher-order beliefs are required, there are more variations among subjects' expected payoffs. In fact, the range and standard deviation of EV-SBC are all greater than those of EV-2ndSBC, EV-1st1D, and EV-1st2D.

Since our indicators are somewhat ad hoc and the classification of games is rather arbitrary, we employ principal component analysis to form several linear combinations of the standardized expected payoffs of the 33 games used in the experiment. We interpret the first five PCs as subject's strategic IQs: The first SIQ (SIQ_1) indicates subjects' abilities to perform backward induction. The second SIQ (SIQ_2) could be interpreted as subjects' abilities to perform multi-dimensional backward induction. The third SIQ (SIQ_3) controls for subjects' attitudes toward risk. The fourth SIQ (SIQ_4) reflects subjects' beliefs about others' social preferences. The fifth SIQ (SIQ_5) measures subjects' accuracy of the higher order beliefs about the opponents in the SBC games.

References

Beard, T. Randolph, and Richard O. Beil. 1994. "Do People Rely on the Self-interested Maximization of Others: An Experimental Test." *Management Science*, 40(2): 252–262.

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

Chen, Chun-Ting, Chen-Ying Huang, and Joseph Tao-yi Wang. 2013. "A Window of Cognition: Eyetracking the Reasoning Process in Spatial Beauty Contest Games." National Taiwan University Working Papaer.

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

Crawford, Vincent P., Miguel Costa-Gomes, and Nagore Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature*, 51(1): 5–62.

Ert, Eyal, Ido Erev, and Alvin E. Roth. 2011. "A Choice Prediction Competition for Social Preferences in Simple Extensive Form Games: An Introduction." *Games*, 2(3): 257–276.

Goeree, Jacob K., and Charles A. Holt. 2001. "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions." *American Economic Review*, 91(5): 1402–1422.

Horn, John L. 1965. "A Rationale and Test for the Number of Factors in Factor Analysis." *Psychometrika*, 30(2): 179–185.

Jolliffe, Ian T. 2002. *Principal Component Analysis, 2nd Edition.* Springer.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review*, 85(5): 1313–1326.

Selten, Reinhard. 1991. "Properties of a Measure of Predictive Success." *Mathematical Social Sciences*, 21(2): 153–167.

Sharma, Subhash. 1995. *Applied Multivariate Techniques.* John Wiley & Sons, Inc.

Stahl, Dale O. II, and Paul W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, 10(1): 218–254.

Appendix

A Procedure for Principal Component Analysis

This mathematical appendix summarizes Chapter 4 of Sharma (1995) and Chapter 2 of Jolliffe (2002), which describe the mathematical procedure of principal component analysis we adopt.

Let $\mathbf{X}$ be a 33-component vector which contains 72 subjects' normalized EV of the 33 games used in our experiment. The covariance matrix, $\mathbf{\Sigma}$, is given by $\mathrm{E}(\mathbf{X}\mathbf{X}')$. Let $\boldsymbol{\omega}' = (w_1\, w_2 \cdots w_{33})$ be a vector of weights such that the new variable, $\boldsymbol{\xi} = \boldsymbol{\omega}'\mathbf{X}$, is a linear combination of the subjects' original normalized EV of the 33 games. The variance of the new variable is given by the $\mathrm{E}(\boldsymbol{\xi}\boldsymbol{\xi}')$, which equals to $\boldsymbol{\omega}'\mathbf{\Sigma}\boldsymbol{\omega}$. The purpose of PCA is finding the weight vector, $\boldsymbol{\omega}$, such that the variance, $\boldsymbol{\omega}'\mathbf{\Sigma}\boldsymbol{\omega}$, of the new variable is maximum over the class of linear combinations that can be formed subject to the constraint $\boldsymbol{\omega}'\boldsymbol{\omega} = 1$.

The solution to the maximization problem can be obtained as follows:
Let

$$\mathrm{Z} = \boldsymbol{\omega}'\mathbf{\Sigma}\boldsymbol{\omega} - \lambda(\boldsymbol{\omega}'\boldsymbol{\omega} - 1), \tag{A.1}$$

where λ is the Lagrange multiplier. The 33-component vector of the partial derivative is given by

$$\frac{\partial \mathrm{Z}}{\partial \boldsymbol{\omega}} = 2\mathbf{\Sigma}\boldsymbol{\omega} - 2\lambda\boldsymbol{\omega}. \tag{A.2}$$

The first order condition of this problem is setting the above vector of partial derivatives to zero. That is,

$$(\mathbf{\Sigma} - \lambda\mathbf{I})\boldsymbol{\omega} = \mathbf{0}. \tag{A.3}$$

For the above system of homogeneous equations to have a nontrivial solution the determinant of $(\mathbf{\Sigma} - \lambda\mathbf{I})$ should be zero. That is,

$$|\mathbf{\Sigma} - \lambda\mathbf{I}| = \mathbf{0}. \tag{A.4}$$

Equation A.4 is a polynomial in λ of order 33, and therefore has 33 roots. Let $\lambda_1 \geq \lambda_2 \geq, \ldots, \lambda_{33}$ be the 33 roots. That is, Equation A.4 results in 33 values for λ, and each value is called the root or eigenvalue of the $\mathbf{\Sigma}$ matrix. Each value of λ results in a set of weights given by the 33-component vector $\boldsymbol{\omega}$ by solving the following equations:

$$
\begin{aligned}
(\boldsymbol{\Sigma} - \lambda \mathbf{I})\boldsymbol{\omega} &= \mathbf{0} && \text{(A.5)}\\
\boldsymbol{\omega}'\boldsymbol{\omega} &= 1. && \text{(A.6)}
\end{aligned}
$$

As a result, the first eigenvector, $\boldsymbol{\omega_1}$, corresponding to the first eigenvalue, λ_1, is obtained by solving equations

$$
\begin{aligned}
(\boldsymbol{\Sigma} - \lambda_1 \mathbf{I})\boldsymbol{\omega_1} &= \mathbf{0} && \text{(A.7)}\\
\boldsymbol{\omega_1}'\boldsymbol{\omega_1} &= 1. && \text{(A.8)}
\end{aligned}
$$

Premultiplying Equation A.7 by $\boldsymbol{\omega_1'}$ gives

$$
\begin{aligned}
\boldsymbol{\omega_1'}(\boldsymbol{\Sigma} - \lambda_1 \mathbf{I})\boldsymbol{\omega_1} &= \mathbf{0}\\
\boldsymbol{\omega_1'}\boldsymbol{\Sigma}\boldsymbol{\omega_1} &= \lambda_1 \boldsymbol{\omega_1'}\boldsymbol{\omega_1}\\
\boldsymbol{\omega_1'}\boldsymbol{\Sigma}\boldsymbol{\omega_1} &= \lambda_1 && \text{(A.9)}
\end{aligned}
$$

as $\boldsymbol{\omega_1'}\boldsymbol{\omega_1} = 1$. The left-hand side of Equation A.9 is the variance of the new variable, $\boldsymbol{\xi_1}$, and is equal to the eigenvalue, λ_1. The first PC is hence given by the eigenvector, $\boldsymbol{\omega_1}$, corresponding to the largest eigenvalue, λ_1.

Let $\boldsymbol{\omega_2}$ be the second 33-component vector of the weights to form the next linear combination. $\boldsymbol{\omega_2}$ can be found such that the variance of $\boldsymbol{\omega_2'}\mathbf{X}$ is the maximum subject to the constraints $\boldsymbol{\omega_1'}\boldsymbol{\omega_2} = 0$ and $\boldsymbol{\omega_2'}\boldsymbol{\omega_2} = 1$ (The first constraint ensures that $\boldsymbol{\xi_1}$ and $\boldsymbol{\xi_2}$ are orthogonal). It can be shown that $\boldsymbol{\omega_2'}$ is the eigenvector of λ_2, and the second largest eigenvalue of $\boldsymbol{\Sigma}$. Similarly, it can be shown that the remaining vectors of weights to form PCs, $\boldsymbol{\omega_3'}, \boldsymbol{\omega_4'}, \ldots, \boldsymbol{\omega_{33}'}$, are also the eigenvectors corresponding to the eigenvalues, $\lambda_3, \lambda_4, \ldots, \lambda_{33}$, of the covariance matrix, $\boldsymbol{\Sigma}$. Consequently, the problem of finding the weights reduces to finding the eigenstructure of the covariance matrix. The eigenvectors give the vectors of weights and the eigenvalues represent the variances of the PCs.

B Additional Figures

B.1 Data from Practice 2nd-Mover SBC Games

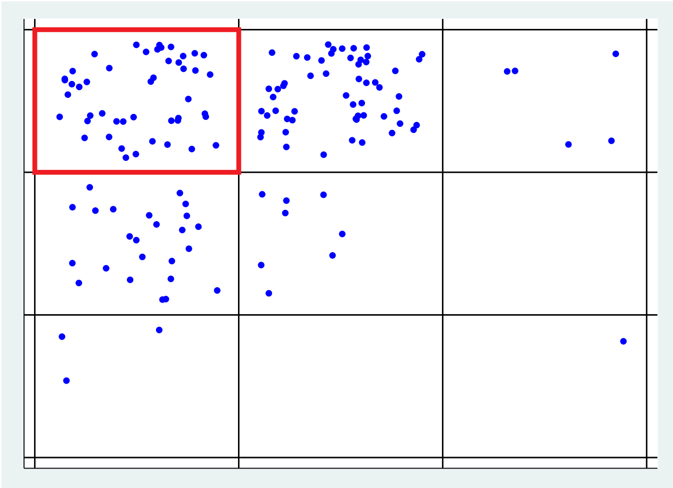


Figure A.1: Choice Distribution of Game 2nd-I with Targets $(0, 1)$ (own) and $(-1, 0)$ (computer) on a 3×3 map

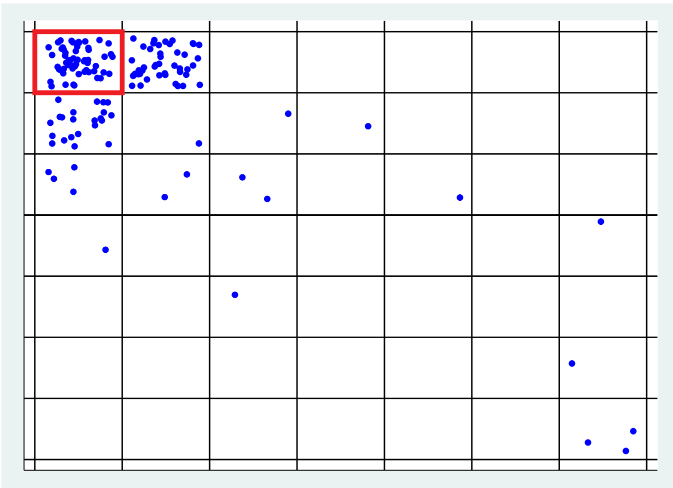


Figure A.2: Choice Distribution of Game 2nd-II with Targets $(-1, 2)$ (own) and $(4, 2)$ (computer) on a 7×7 map

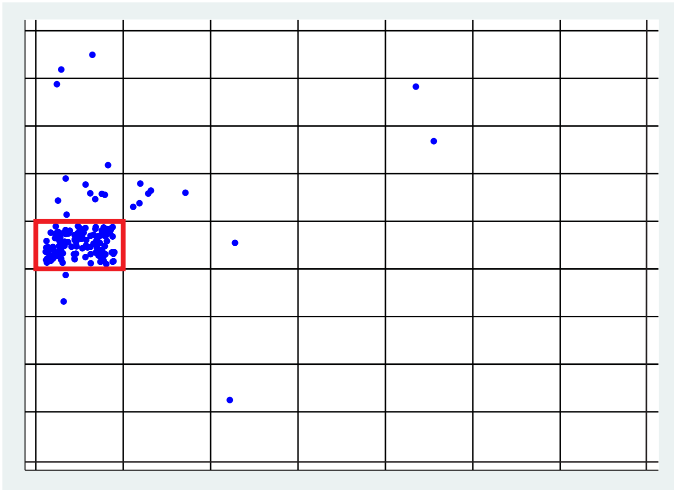


Figure A.3: Choice Distribution of Game 2nd-III with Targets $(-1,-4)$ (own) and $(4,2)$ (computer) on a 7×9 map

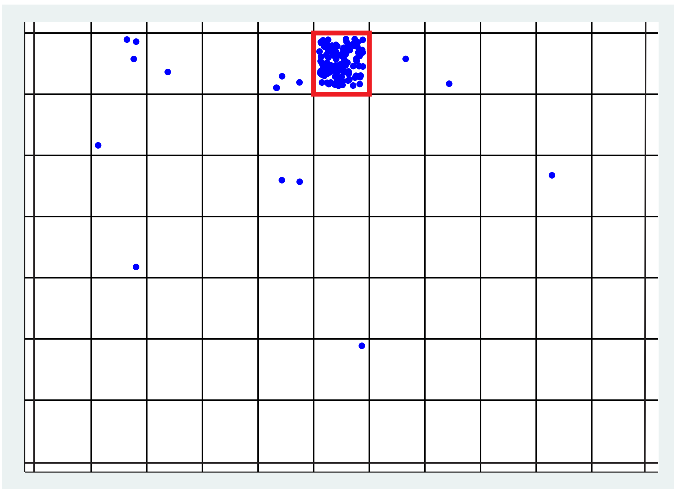


Figure A.4: Choice Distribution of Game 2nd-IV with Targets $(4,2)$ (own) and $(-6,-3)$ (computer) on a 9×7 map

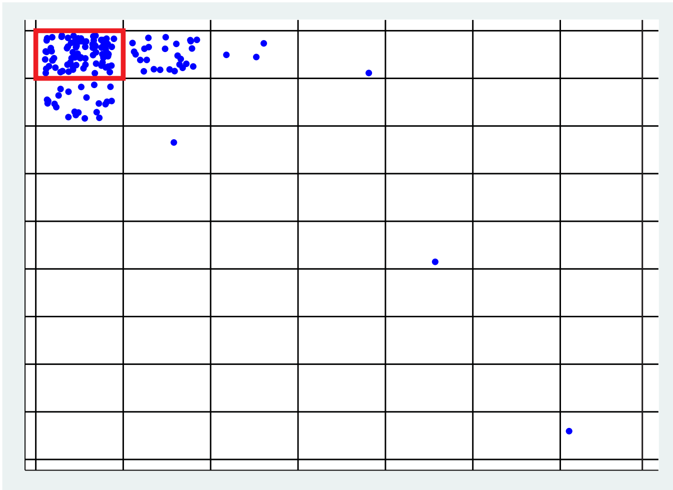


Figure A.5: Choice Distribution of Game 2nd-V with Targets $(-2, 1)$ (own) and $(4, -4)$ (computer) on a 7×9 map

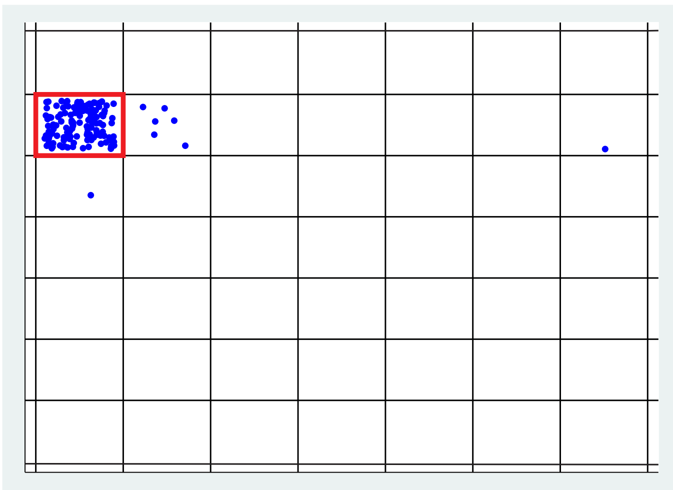


Figure A.6: Choice Distribution of Game 2nd-VI with Targets $(0, -1)$ (own) and $(1, 0)$ (computer) on a 7×7 map

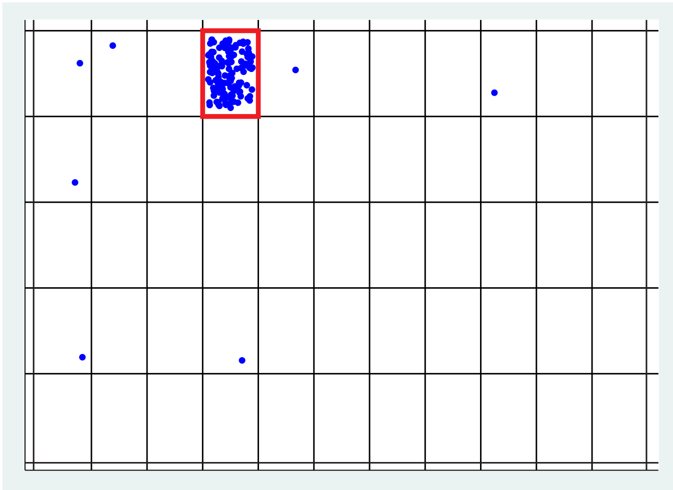


Figure A.7: Choice Distribution of Game 2nd-VII with Targets $(3, 0)$ (own) and $(0, 3)$ (computer) on a 11×5 map

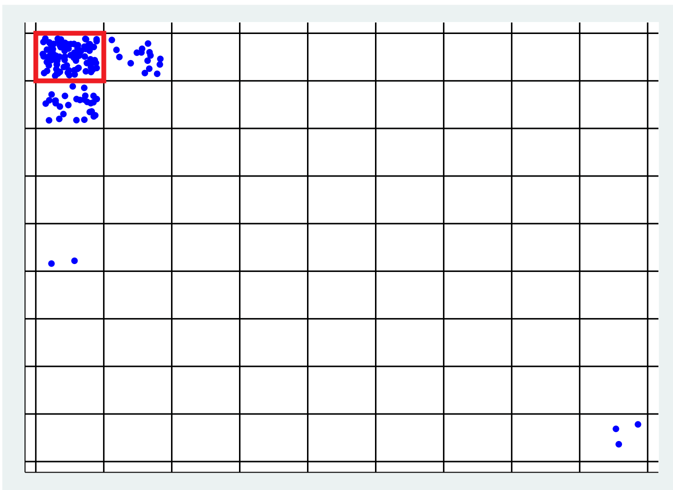


Figure A.8: Choice Distribution of Game 2nd-VIII with Targets $(-1, 0)$ (own) and $(0, -4)$ (computer) on a 9×9 map

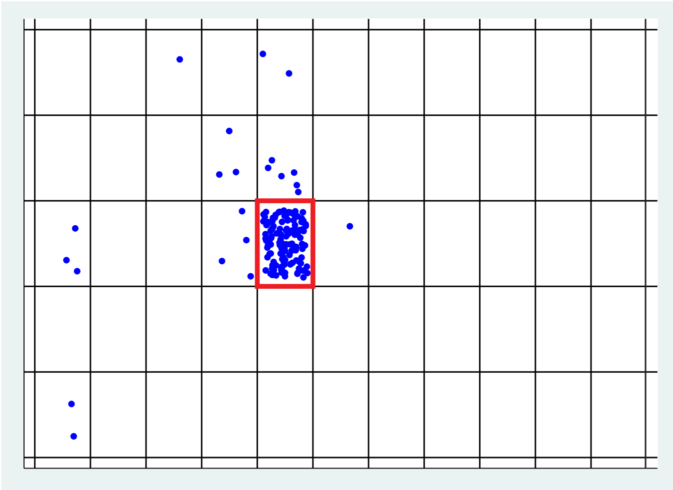


Figure A.9: Choice Distribution of Game 2nd-IX with Targets $(4, -2)$ (own) and $(-2, -4)$ (computer) on a 11×5 map

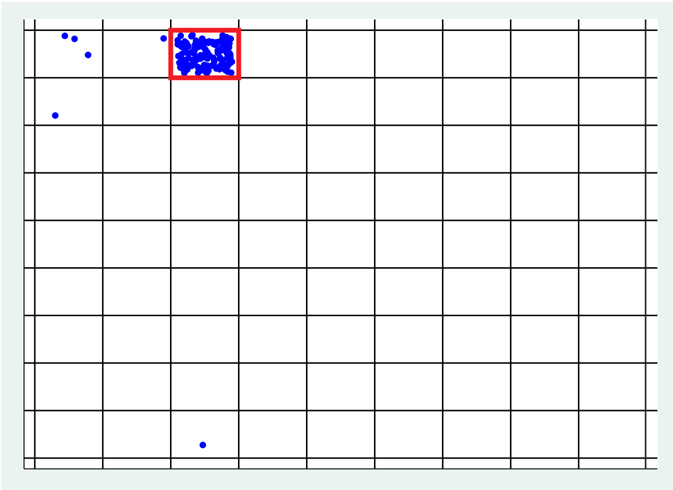


Figure A.10: Choice Distribution of Game 2nd-X with Targets $(2, 1)$ (own) and $(-2, -6)$ (computer) on a 9×9 map

B.2 Data from Simultaneously SBC Games

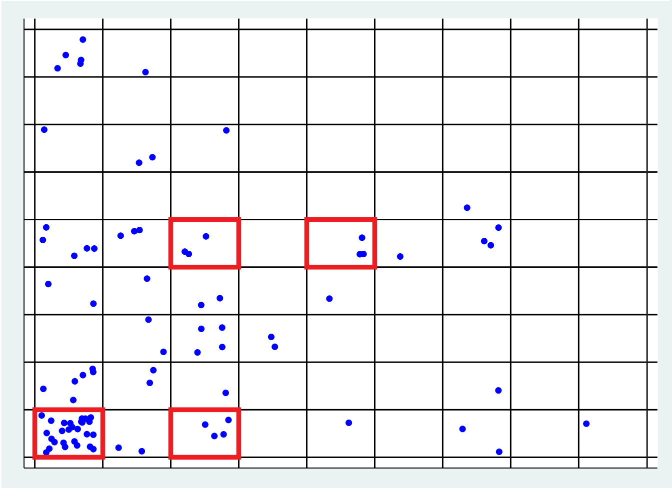


Figure A.11: Choice Distribution of Game SBC-1 with Targets $(-2, 0)$ (own) and $(0, -4)$ (opponent) on a 9×9 map

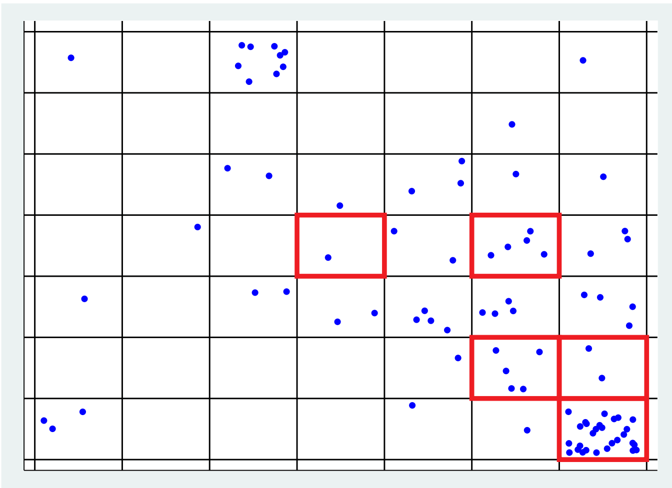


Figure A.12: Choice Distribution of Game SBC-2 with Targets $(2, 0)$ (own) and $(0, -2)$ (opponent) on a 7×7 map

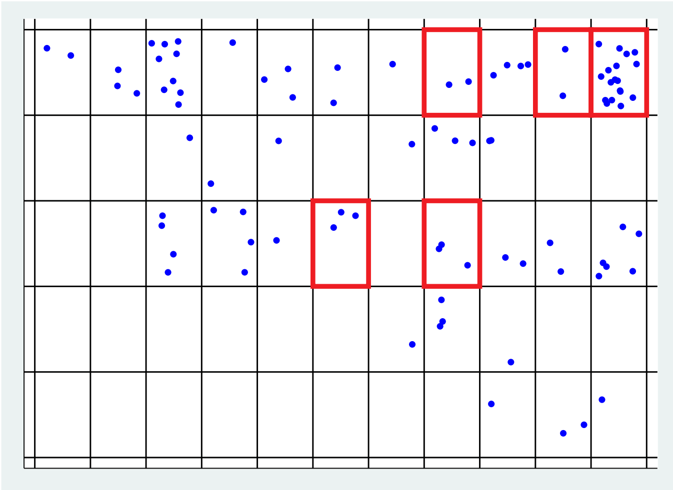


Figure A.13: Choice Distribution of Game SBC-3 with Targets $(2, 0)$ (own) and $(0, 2)$ (opponent) on a 11×5 map

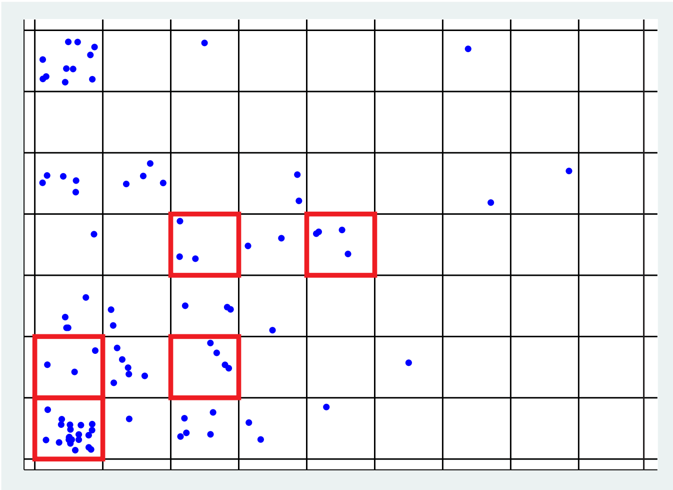


Figure A.14: Choice Distribution of Game SBC-4 with Targets $(-2, 0)$ (own) and $(0, -2)$ (opponent) on a 9×7 map

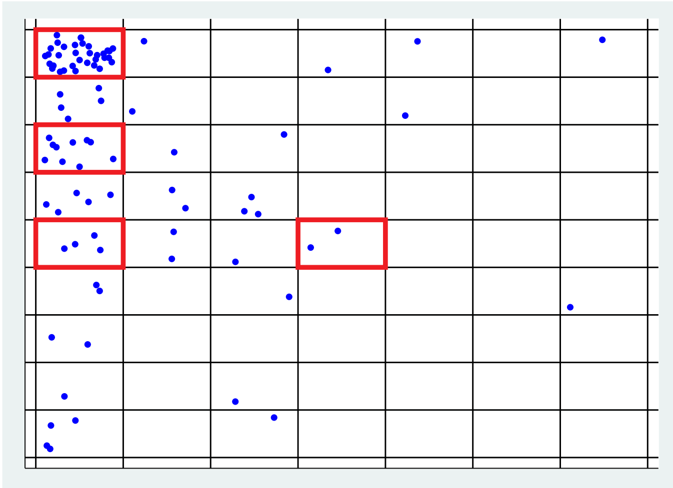


Figure A.15: Choice Distribution of Game SBC-5 with Targets $(-4, 0)$ (own) and $(0, 2)$ (opponent) on a 7×9 map

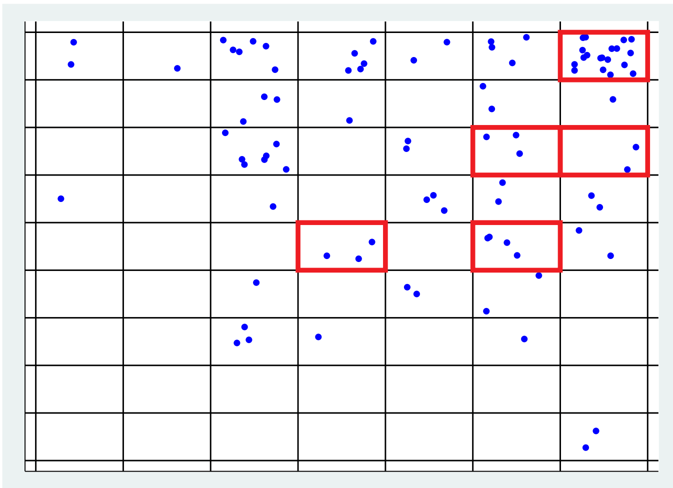


Figure A.16: Choice Distribution of Game SBC-6 with Targets $(2, 0)$ (own) and $(0, 2)$ (opponent) on a 7×9 map

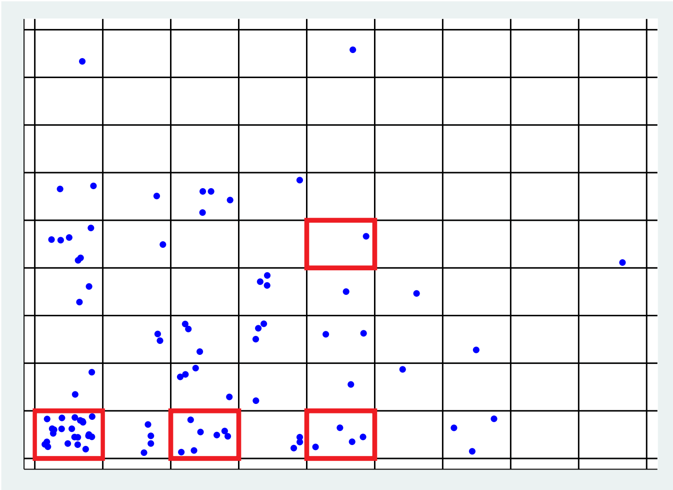


Figure A.17: Choice Distribution of Game SBC-1R with Targets $(0, -4)$ (own) and $(-2, 0)$ (opponent) on a 9×9 map

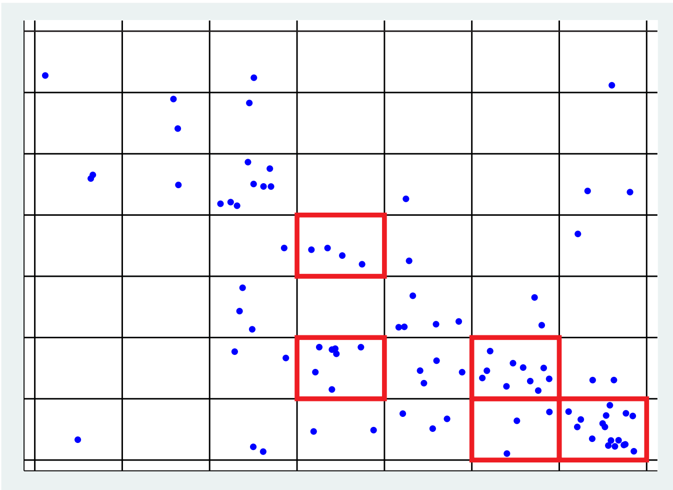


Figure A.18: Choice Distribution of Game SBC-2R with Targets $(0, -2)$ (own) and $(2, 0)$ (opponent) on a 7×7 map

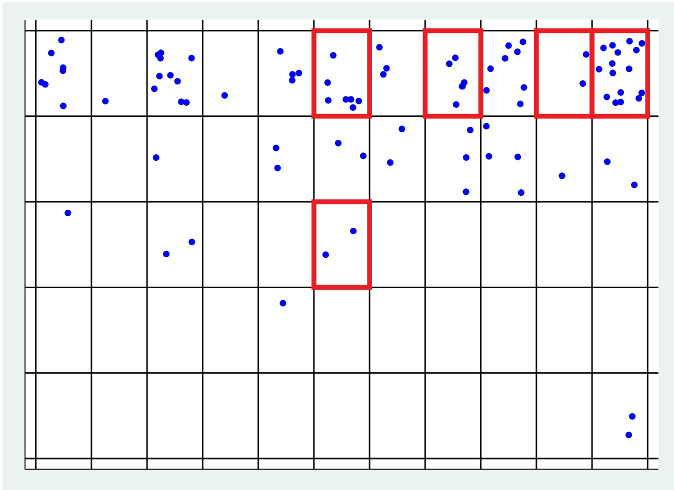


Figure A.19: Choice Distribution of Game SBC-3R with Targets $(0, 2)$ (own) and $(2, 0)$ (opponent) on a 11×5 map

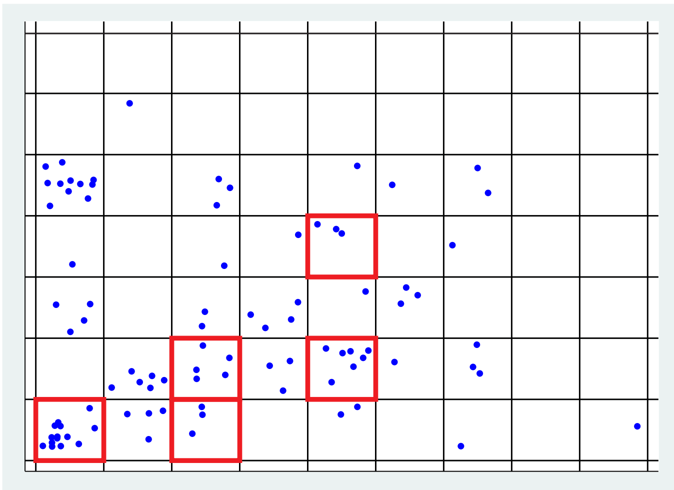


Figure A.20: Choice Distribution of Game SBC-4R with Targets $(0, -2)$ (own) and $(-2, 0)$ (opponent) on a 9×7 map

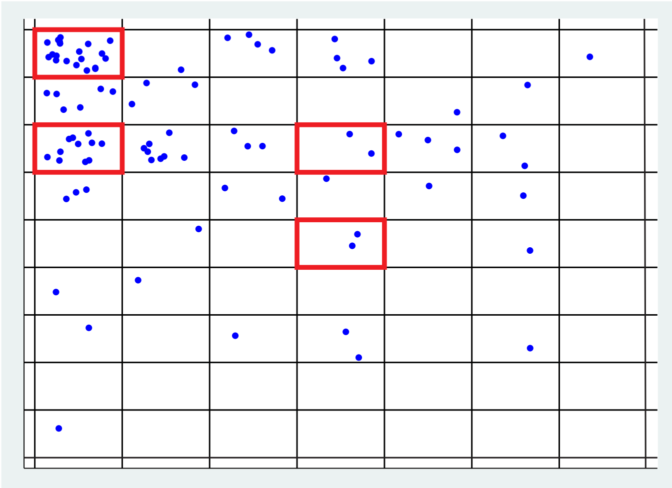


Figure A.21: Choice Distribution of Game SBC-5R with Targets $(0, 2)$ (own) and $(-4, 0)$ (opponent) on a 7×9 map

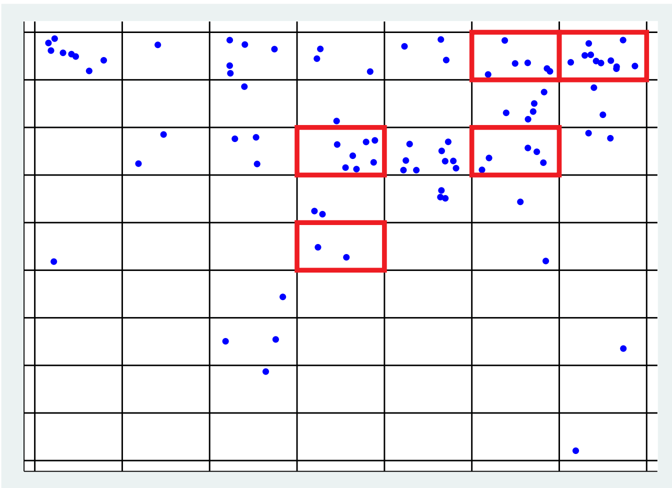


Figure A.22: Choice Distribution of Game SBC-6R with Targets $(0, 2)$ (own) and $(2, 0)$ (opponent) on a 7×9 map

B.3 Data from 1st-Mover SBC Games

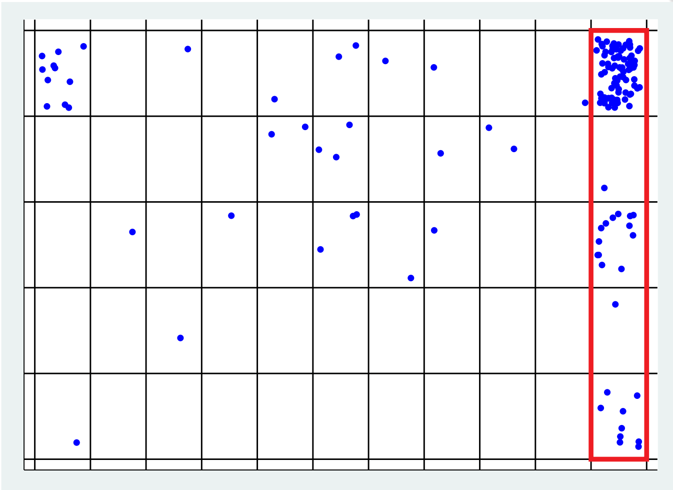


Figure A.23: Choice Distribution of Game 1st-3R with Targets $(0, 2)$ (own) and $(2, 0)$ (computer) on a 11×5 map

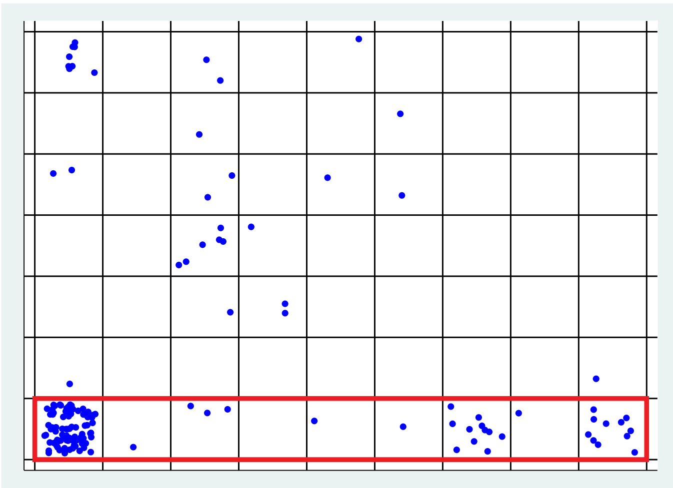


Figure A.24: Choice Distribution of Game 1st-4 with Targets $(-2, 0)$ (own) and $(0, -2)$ (computer) on a 9×7 map

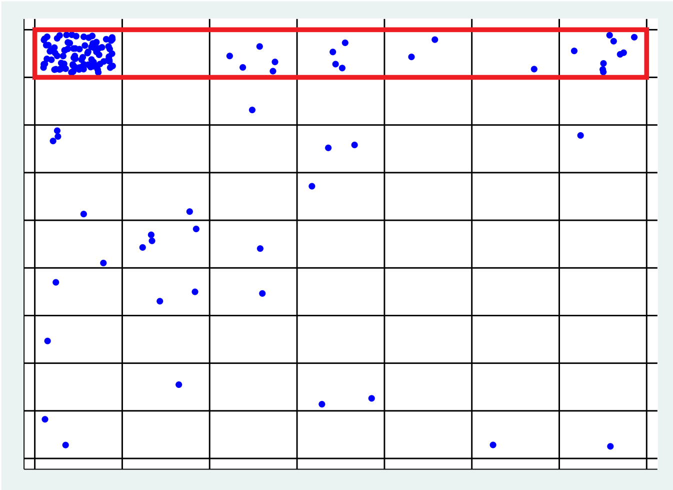


Figure A.25: Choice Distribution of Game 1st-5 with Targets $(-4, 0)$ (own) and $(0, 2)$ (computer) on a 7×9 map

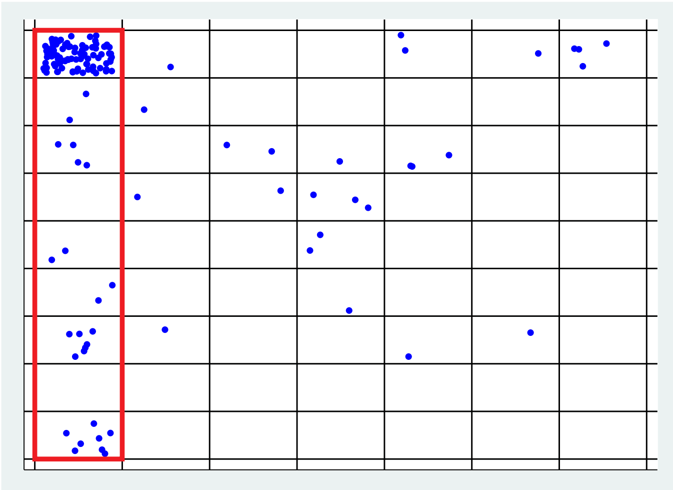


Figure A.26: Choice Distribution of Game 1st-5R with Targets $(0, 2)$ (own) and $(-4, 0)$ (computer) on a 7×9 map

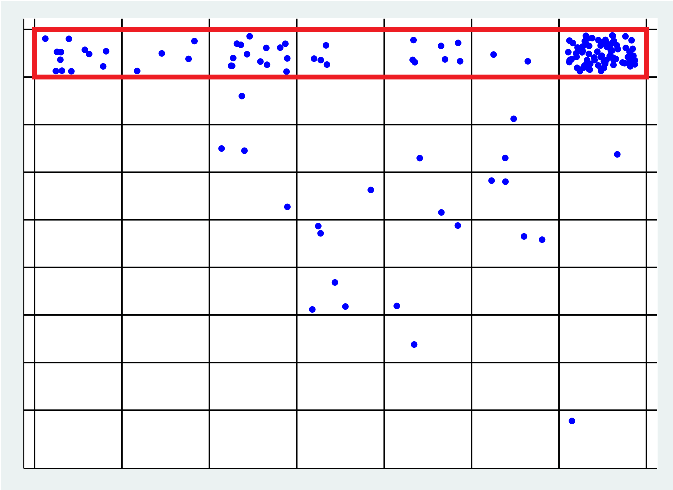


Figure A.27: Choice Distribution of Game 1st-6 with Targets $(2, 0)$ (own) and $(0, 2)$ (computer) on a 7×9 map

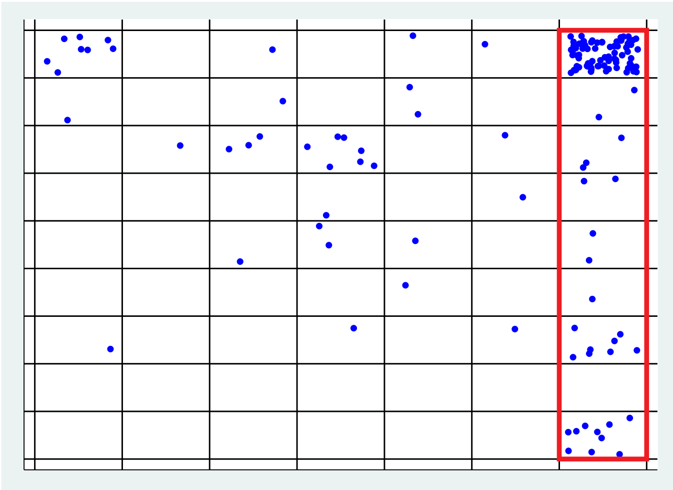


Figure A.28: Choice Distribution of Game 1st-6R with Targets $(0, 2)$ (own) and $(2, 0)$ (computer) on a 7×9 map

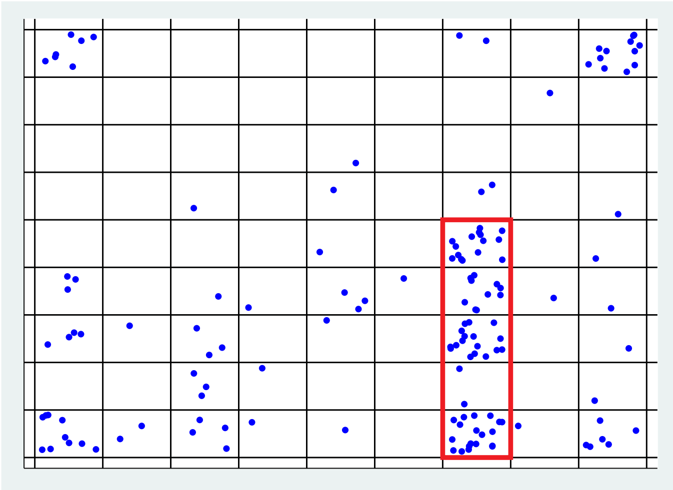


Figure A.29: Choice Distribution of Game 1st-7 with Targets $(-2, -6)$ (own) and $(4, 4)$ (computer) on a 9×9 map

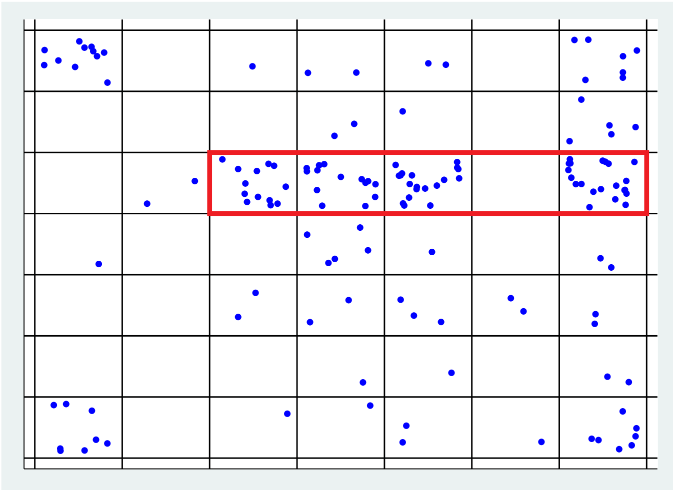


Figure A.30: Choice Distribution of Game 1st-8 with Targets $(4, -2)$ (own) and $(-2, 4)$ (computer) on a 7×7 map

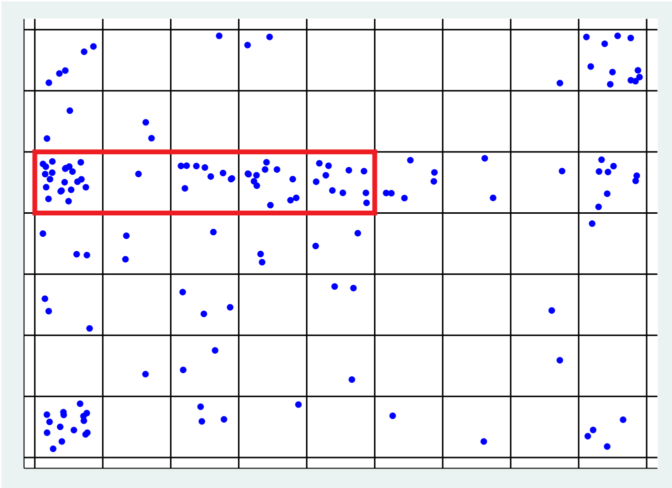


Figure A.31: Choice Distribution of Game 1st-9 with Targets $(-6, -2)$ (own) and $(4, 4)$ (computer) on a 9×7 map

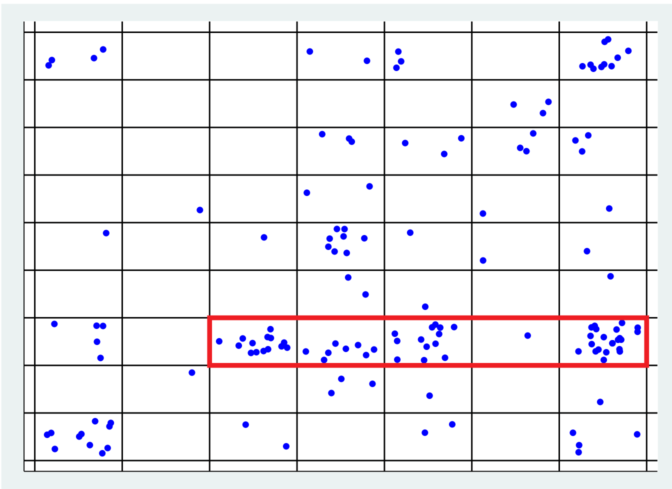


Figure A.32: Choice Distribution of Game 1st-10 with Targets $(4, 2)$ (own) and $(-2, -4)$ (computer) on a 7×9 map

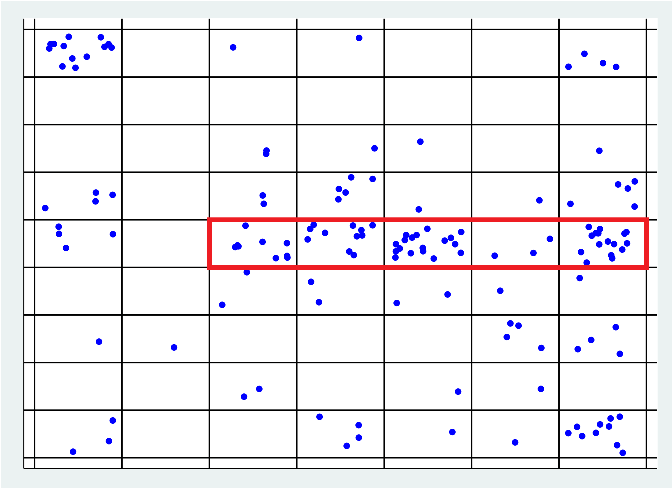


Figure A.33: Choice Distribution of Game 1st-11 with Targets $(4, -4)$ (own) and $(-2, 6)$ (computer) on a 7×9 map

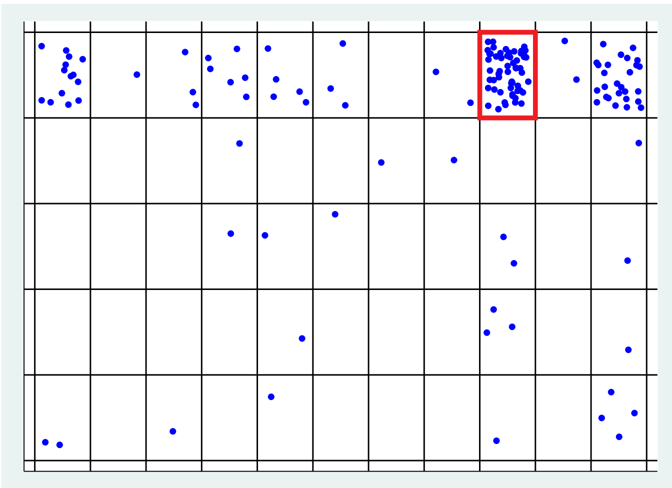


Figure A.34: Choice Distribution of Game 1st-12 with Targets $(-2, 4)$ (own) and $(6, 2)$ (computer) on a 11×5 map

B.4 Parallel Analysis

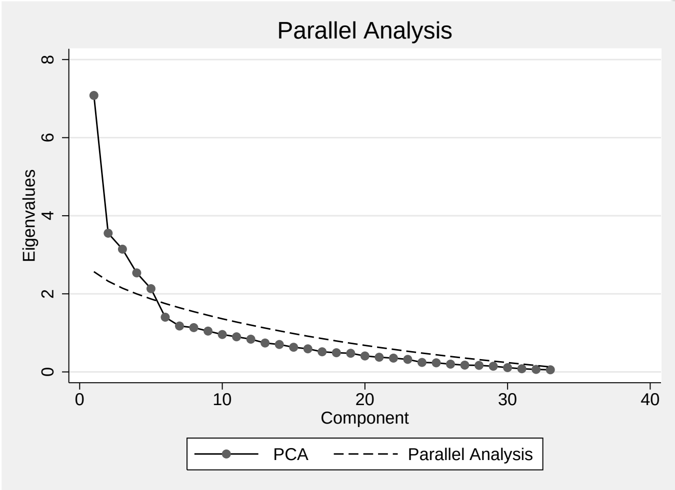


Figure A.35: Plot of Eigenvalues of the Actual Data and Plot of Eigenvalues from Parallel Analysis

C Instructions (Slides Used in the Experiment)

Experimental Instructions

Experiment 1 (Practice) - 1

- Each round you pair with another person
- For Practice, the other person is Computerized
  - Programmed to act in a pre-set way
- You choose the option **LEFT** or **RIGHT**; the other person will choose **UP** or **DOWN**
  - The other person's choice matters ONLY if you choose **RIGHT**

Experiment 1 (Practice)

- Your earnings will then be determined by the **BLUE** numbers next to each box
  - Numbers (how much you earn) vary across rounds
  - The other person's earnings are in **GREY**
- Results WILL NOT count toward final earnings
  - This is just practice to make sure you understand

Experiment 1 (Real) - 1

- Same as Practice:
  - Each round you pair with another person
- The other person is a **fellow UCLA Student**
- You choose the option **LEFT** or **RIGHT**; the other person will choose **UP** or **DOWN**
  - The other person's choice matters ONLY if you choose **RIGHT**

Experiment 1 (Real)

- Your earnings will then be determined by the **BLUE** numbers next to each box
  - Numbers (how much you earn) vary across rounds
  - The other person's earnings are in **GREY**
- **You will not see the other person's decisions**

- Results WILL count toward your final earnings
  - Earnings from **one round** will be randomly drawn

Experiment 2 – Participant 1

- Each round you pair with another person
- The other person is a **fellow UCLA Student**
- You are given 10 CHIPS to be allocated between you and the other person
- Each CHIP assigned to you gives you **$1**
- Each CHIP assigned to the other person gives him/her **$0.50, $1 or $2** (differs across rounds)

Experiment 2 – Participant 1

- The other person can only accept your allocation
- Results WILL count toward your final earnings
  - Earnings from **one round** will be randomly drawn

Experiment 2 – Participant 1

- Some rounds have a third person: Participant 3
  - Allocate **0-5 deduction POINTS** depending on your allocation of CHIPS for you and Participant 2
- Each deduction POINT assigned to you
  - Reduces $1 from You
  - Reduces $0.25, $0.50 or $1 from him/her (differs)
- No feedback on rounds with Participant 3
  - Don't know allocation of deduction POINTS

Experiment 3 (Practice)

- Each round you pair with another person
- For Practice, the other person is Computerized
  - Programmed to act in a pre-set way
- Both of you will place **markers** on a grid
  - Markers may overlap
- The other person will go first
  - You will see other's marker before you decide
- Results WILL NOT count toward final earnings
  - This is just practice to make sure you understand

Experiment 3 (Practice)

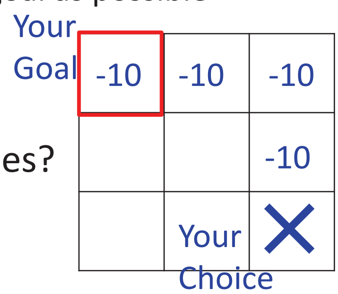
- Each round you have a **goal** where you want your marker to be located, compared to the other person's marker.
  - Ideal location is not fixed, but relative to where the other person puts their marker

Example: "1 ABOVE" means your goal is for your marker to be one square above the other's

Example: Other's goal "2 LEFT" means their goal is to place a marker 2 squares to the left of yours

Experiment 3 (Practice)

- Both of you will see both of your goals.
- Start with $10; lose $0.50 for each square between your marker and the ideal one
  - Want to be as close to your goal as possible



- Any questions about the rules?

Experiment 3 (Practice)

- Now you will go through some Practice Rounds
- For Practice, the other person is Computerized
  - Programmed to act in a pre-set way
- Please ask questions as you go, and let us know if there is anything that is confusing
- Results WILL NOT count toward final earnings
  - This is just practice to make sure you understand

Experiment 3 (Part A)

- Now you will go though Part A
  - You and the other person choose simultaneously
  - Nobody will see other's marker
  - You need to think about (and guess) where the other person might place the marker
- The other person is a **fellow UCLA Student**
- Results WILL count toward final earnings
  - Earnings from **one round** will be randomly drawn

Experiment 3 (Part B)

- Now you will go though Part B
  - You go first
  - The other person will see your marker
- The other person is a Computerized Person
  - Programmed to **earn the most** for himself
- Results WILL count toward final earnings
  - Earnings from **one round** will be randomly drawn