Devetag, Giovanna; Di Guida, Sibilla

Working Paper

Feature-based choice and similarity in normal-form games: An experimental study

LEM Working Paper Series, No. 2010/18

Provided in Cooperation with:
Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies

Suggested Citation: Devetag, Giovanna; Di Guida, Sibilla (2010) : Feature-based choice and similarity in normal-form games: An experimental study, LEM Working Paper Series, No. 2010/18, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM), Pisa

This Version is available at:
http://hdl.handle.net/10419/89271

Standard-Nutzungsbedingungen:
Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen. Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Feature-based Choice and Similarity in Normal-form Games: An Experimental Study

Giovanna Devetag
Sibilla Di Guida

° University of Perugia, Italy
§ CIFREM- University of Trento, Italy

2010/18
October 2010
Feature-based Choice and Similarity in Normal-form Games: An Experimental Study*

Giovanna Devetag
University of Perugia

Sibilla Di Guida
CIFREM, University of Trento

October 2010

Abstract

In this paper we test the effect of descriptive “features” on initial strategic behavior in normal form games, where “descriptive” are all those features that can be modified without altering the (Nash) equilibrium structure of a game. We observe that our experimental subjects behave according to some simple heuristics based on descriptive features, and that these heuristics are stable even across strategically different games. This suggests that a categorization of games based on features may be more accurate in predicting agents' initial behavior than the standard categorization based on Nash equilibria, as shown by the analysis of individual behavior. Analysis of choice patterns and individual response times suggests that non-equilibrium choices may be due to the use of incorrect and simplified mental representations of the game structure, rather than to beliefs in other players' irrationality. Of the four stationary concepts analyzed (Nash equilibrium, QRE, action sampling, and payoff sampling), QRE results the best in fitting the data.

Keywords: normal form games, one-shot games, response times, dominance, similarity, categorization, focal points
JEL classification codes: C72, C91, C92

* We thank Christopher Gilbert, Paola Manzini, Marco Piovesan, Ondrej Rydval, Theodore L. Turocy, Enrico Zaninotto, participants at IMEBE 2010 in Bilbao, and seminar participants in Trento, for useful comments and suggestions. We thank the staff of the CEEL experimental lab for help in the implementation of the experiments. We gratefully acknowledge financial support from R.O.C.K. research group, of the University of Trento. The usual caveats apply.
1. Introduction

According to traditional game theory, strategic behavior is solely guided by a game equilibrium structure. As a consequence, players’ behavior should comply with equilibrium and should not be affected by modifications of a game that leave its equilibrium structure unaltered. This basic tenet of game theory has been repeatedly and convincingly proved wrong by a plethora of experimental studies on single-shot games in normal form, showing not only that subjects’ behavior is often out of equilibrium (Nash Equilibrium), but also that strategizing responds to features that are theoretically irrelevant (e.g., Bosch-Domènech and Vriend, 2008; Cooper and Van Huyck, 2003; Costa-Gomes et al., 2001; Crawford et al., 2008; Goeree and Holt, 2001, 2004). Some of these results have stimulated the development of new equilibrium concepts, in which agents’ behavior is explained by the “trembling hand” effect (as in the Quantal response equilibrium, McKelvey and Palfrey, 1995), by behavioral assumptions (Impulse Balance Equilibrium, Selten and Chmura, 2008), or by bounded rationality intended as a limited capacity to process information (Payoff-sampling equilibrium, Osborne and Rubinstein, 1998; Action-sampling equilibrium, Selten and Chmura, 2008).

But even these new stationary concepts fall short of capturing strategic behavior in a vast range of situations. Common to most experimental findings on single-shot games is, first, a high level of heterogeneity which cannot be accounted for by any model of choice; second, a large fraction of behavior that is either straightforwardly non-strategic in nature, or strategic in a non-standard sense. Behavioral models estimated using large data sets (Weizsacker 2003), and experiments that try to track down individual reasoning processes (Devetag and Warglien 2008; Rydval, Ortmann and Ostatnicky 2009) suggest that players reason through incomplete models of the strategic situation at hand, either tending to ignore their opponents’ incentives when making their choices, or treating them as mirror images of their own. Experiments that elicit both choices and beliefs about opponents’ play point at a general inconsistency between choices and beliefs, suggesting that in the absence of learning opportunities and feedback, the assumption that actions are driven by beliefs about the opponent must not be taken for granted (Costa-Gomes and Weizsacker 2008; Stahl and Haruvy 2008). Hence, more research is needed to investigate what drives choices in one-shot games, given that many strategic situations that people experience are unique, and that quite seldom repeated interaction on the same identical game with transparent feedback occur in the real world.
We submit that most of players’ behavior in one-shot games in normal form follows very simple choice criteria that are either non-strategic (in the sense that they do not seem to take the opponents’ incentives into account) or strategic in a naive sense, as will be explained later. As a consequence, players’ behavior can be influenced by manipulating a set of game features that do not alter the game pure strategy Nash equilibria.

In our experiment we employ 30 3x3 games in normal form that belong to five well-known game types. As we are interested in initial behavior only, players were rematched randomly at every round with no feedback, to avoid learning and “repeated game effects” as much as possible. We create, for each game type, six different versions through the manipulation of two features: the presence vs. absence of a “focal point” and the creation of three different levels of payoff variance for the strategy presenting the highest average payoff (which differs - except in one case - both from the strategy leading to the focal point and from the equilibrium strategy). Our definition of a focal point differs both from Schelling’s (1963) and from those previously used in all experimental games (Metha et al., 1994; Crawford et al., 2008; Bosch-Domène and Vriend, 2008), as we define “focal” any outcome that is Pareto efficient and yields identical payoffs to the players. It follows that focal points in our games need not be equilibrium outcomes. We also test the effect of payoff magnitude and position of the cell in the matrix in determining the attractiveness of a focal point.

We use the level of variance as an intuitive measure of a strategy riskiness. To the best of our knowledge, we are the first to examine the role of payoff variance in determining choices. Behavioral models that estimate a distribution of player “types” have focused on expected value (for “Level 1” types), on single payoffs (for Optimistic and Pessimistic or Maximin types), or on payoff sums (for Altruistic types). We show that the variance of a strategy, together with its expected value, must be taken into account when trying to explain players’ choices in single-shot games.

Our manipulations (that are mostly “economic” in nature, implying exclusively changes in payoffs and, for one game only, changes in the position of the focal point in the matrix) influence behavior significantly. Our main results can be summarized as follows: in the matrices in which a focal point is present, the large majority of our players (more than 83 per cent on average) pick either the strategy leading to the focal point or the strategy with the highest expected value, both strategies that are out of (Nash) equilibrium; in the matrices without the focal point (which is removed simply by introducing a slight payoff asymmetry), almost 74 per cent of players select the strategy with the highest expected value when its variance is low (still a non-equilibrium strategy), and the share of players who select this strategy declines almost linearly with the increase in its variance. We vary the attributes that characterize a focal point and find out that focality of an outcome is highest when
its payoffs are symmetric and “significantly” higher than the other payoffs that a player can obtain from the game. Pareto efficiency, cell position, and equilibrium property instead do not contribute to an outcome focality significantly. For the three game types for which the equilibrium strategy differs both from the strategy with the highest expected value and from the focal point strategy, the share of the equilibrium strategy averages 50 per cent only when the focal point is absent and the variance of the strategy with the highest expected value is high. Otherwise, the equilibrium strategy share does not exceed 20 per cent.

An analysis of subjects’ response times shows that matrices with a focal point take a shorter time on average to be processed, suggesting that focal points trigger forms of intuitive reasoning (Kuo et al. 2009); further, response times increase monotonically with the increase in the variance of the strategy with the highest average payoff.

We also show that players respond similarly to games that are “similar” in terms of the features specified above, even when these games belong to very different strategic types. Hence, a categorization based on the presence/absence of a series of features (e.g., an outcome with high and symmetric payoffs, a strategy with high expected value and low variance, etc.) may more useful in predicting initial behavior than a categorization based on a game equilibrium structure.

To try to explain our data, we hypothesize that a safe strategy with a high expected value and a focal point provide two solutions that “stand out” as compared to all other feasible game strategies. The choice of the first strategy is compatible with a “level 1” type (Costa-Gomes, 2001), and can derive either from diffuse priors on the opponent’s play or from a tendency to ignore the opponents’ incentives (Weizsacker 2003, Costa-Gomes and Weizsacker 2008), as observed in previous games; the choice of the focal point strategy is strategic in a naive sense, and we argue that the choice process leading to it may be similar to the choice process leading to a focal equilibrium in a game of coordination, implying some form of team reasoning. In our games, however, the choice of a focal point relies on an incomplete processing of all the matrix elements, since our focal points are not equilibria.

Our findings relate to previous studies in several ways: first, they provide evidence on behavior in single-shot normal form games that cannot be accounted for by any equilibrium concept, nor any behavioral model that assumes a distribution of player types and level-k thinking. Second, they point at the role of a strategy variance as a measure of the riskiness implicit in the choice of a strategy, and as a variable that mitigates the extent to which players may exhibit “Level 1” type of behavior. Third, it extends the notion of a “focal point” well beyond equilibrium outcomes in symmetric games, showing that focality may be a much more general property of single-shot game outcomes, both symmetric and asymmetric. More generally, our results show that mild payoff
changes induce quantitatively and statistically significant changes in behavior, suggesting that choices in these games may result from the interaction of game features that go well beyond their equilibrium structure and subjects’ limited and/or incorrect mental representations of the strategic situation at hand and of their opponents’ motivations.

More specifically, we argue that players in single-shot games, with no opportunity for learning and with no feedback, at first look for obvious ways to play. Picking a strategy with high payoff sum and low variance is one of such obvious ways; picking a focal Pareto-efficient outcome is another obvious solution. Only in the absence of such elements, may players start to reason strategically in a game-theoretic sense, and find their way through equilibrium.

Our results are in line with previous studies of mental models of games (Devetag and Warglien 2008), and add insight to the so-called “pre-game theory” (Camerer 2003), i.e., they contribute to the understanding of models of strategic interaction.

The rest of the paper is organized as follows: section 2 presents the games used in the experiment; section 3 describes the experimental design and implementation, and presents our behavioral hypotheses; section 4 presents the results: we first discuss aggregate results (section 4.1), and then analyze individual response times (section 4.2). In section 4.3 we test the predictive power of a series of non-standard equilibrium concepts (QRE, payoff sampling, action sampling, and impulse-balance). Finally, section 5 offers some concluding remarks.

2. The games

The payoff matrices used in the experiment are presented in Table 1.

We selected 5 3x3 different game types and created 6 versions of each game. In some cases new Nash equilibria emerged together with the original ones, which always remained.

The chosen base games are: a game with a strictly dominant strategy for the column player (the DomCol game henceforth), a game without pure strategy Nash Equilibria (noNE), a game with a unique pure strategy Nash Equilibrium but not solvable through iterated elimination of dominated strategies (UniqNE), a Prisoners’ Dilemma (PD), and a Weak Link coordination game (WL).

For each game we identify the strategy with the highest payoff sum or average payoff (HA), the equilibrium strategy (EQ, whenever a pure strategy Nash Equilibrium is present), and a strategy leading to a Focal Point (FP). A Focal Point is any cell containing Pareto Efficient and symmetric payoffs, located at the center of the matrix, except in the Weak Link game where all symmetric cells are positioned along the main diagonal from the highest to the lowest payoff. Except in the Weak Link game, our Focal Points are not equilibria. We also build different versions of our focal
points to test the relative contribution of Pareto efficiency, cell position, payoff magnitude, and payoff symmetry to an outcome focality.

Our analysis concerns almost entirely the behavior of the row players, since most of our games are not symmetric. Therefore, all the descriptions of strategies and matrices will regard the row player's perspective unless otherwise specified.

Our main goal is to investigate how the presence or absence of Focal Points affects subjects' strategic behavior, as well as the effect of increasing the variance of the HA strategy (we introduce three levels of variance: low, middle, high).

To this end, and in order to tease apart their separate as well as their joint effects, we create a matrix for every possible combination of features.

For each base game six matrices were created: FP and HA with low variance, FP and HA with middle variance, FP and HA with high variance, no FP and HA with low variance, no FP and HA with middle variance, no FP and HA with high variance. For ease of exposition, we name each matrix by the acronym identifying the game type, and by two acronyms identifying the features: “FP” indicates that the matrix has a focal point, “XFP” a matrix without focal point, “L”, “M”, “H” indicate the three levels of variance of the strategy with the highest payoff sum.

All the different versions of the same game were created modifying at minimum the content of cells and always maintaining the same equilibrium structure. In few cases these changes added new Nash equilibria in mixed strategies. In the extreme cases two matrices differ only for a single cell. Except in one matrix (WL_FP_L), the average payoff of the HA strategy was kept unchanged in the different versions of the same game, and only the payoff distribution was modified so as to change the value of payoff variance.

The matrices without FP were obtained by breaking the symmetry of payoffs, and by altering some “relevant attributes” of the FP outcome (see Hypothesis 4). In the case of the weak link game this was not possible without altering the game structure, so we obtained matrices without FP by moving the FP from the top-left cell to a less “focal” position. This allows us to investigate whether the position of the FP plays a role in the cell focality or whether focality is mainly due to Pareto efficiency, payoff symmetry, and payoff magnitude.

In order to measure the impact of every feature, we kept our three strategies of interest separated whenever possible. For example in the DomCol game, Row 1 identifies the HA strategy, Row 2 the FP strategy, and Row 3 the EQ strategy. This was not possible for the Prisoner's Dilemma, where the EQ and HA strategies coincide; therefore in that case a single row is simultaneously the EQ and the HA strategy.
To avoid spurious effects due to the position of the strategy in the matrix, we always kept the position of every strategy fixed in the different versions of the same game, with the only exception of the WL game.

The labels for the strategies used from now on are: EQ for the equilibrium strategy, FP for the strategy leading to the FP, XFP for the strategy in which the Focal Point has been removed, HA for the strategy with the highest average payoff. We then name COS a strategy that gives a constant payoff (which is present only in the weak link game), DOM a dominated (even weakly) strategy. Finally, we define QES a quasi-equilibrium strategy, in a sense that will be explained when discussing our results.

3. Experimental design and behavioral predictions

3.1 Experimental design and implementation

The experiment was conducted at the Computable and Experimental Economics Lab (CEEL) of the University of Trento, in 5 different sessions of 16 subjects each. In every session, 12 people were randomly assigned the role of row player, while 4 of them were assigned the role of column player, for a total of 60 observations for the row player and 20 for the column player. Roles were fixed throughout the experiment. This asymmetry is motivated by the fact (introduced above) that we were interested only in the behavior of the row player. Subjects made their choice as row or column player in the 30 matrices, being re-matched randomly at every round with a player of the opposite role. No feedback regarding the opponent's choice or the obtained payoff was revealed until the end of the experiment.

After entering the lab, subjects were assigned randomly to a pc cubicle and to the role of row or column player. A paper copy of the instructions was given to subjects and was also read aloud by the experimenter. Control questions were administered before the experiment started to assure that the rules of the experiment had been understood. Particular care was taken to make sure that subjects understood how to read a payoff matrix. In case of incorrect answers, instructions were repeated (a translated copy of the instructions and of the control questions in available by the authors upon request).

The experiment was computerized by using a Z-Tree based software (Fischbacher, 2007), especially developed for the purpose. The matrices were presented one at a time in random order, which differed from subject to subject. In each round, subjects had to select their preferred strategy by typing the corresponding row number (Figure 1 reports a sample of the software interface).
The strategies of all players were recorded and matched randomly, but no feedback was given until the end of the experiment. Subjects could use as much time as needed, but they were invited to use no more than 30 seconds. Nonetheless, in several occasions subjects used more than 60 seconds to decide, showing that the suggestion was not perceived as mandatory.

The final payment was determined by the outcomes of 5 matrices picked at random. The exchange rate was announced at the end and this was made explicit to subjects in the instructions. After the last matrix was displayed, one subject selected at random was asked to verify that some tags contained into a jar were each reporting the numeric code of one of the matrices played. Subsequently, a different, randomly selected subject, was asked to draw 5 tags from the jar, which determined the matrices that would be used to calculate subject payments. Then each subject was paid according to the choices she and her assigned opponent had made in those 5 matrices.

After the experiment and the selection of the payment matrices, some personality tests were administered to subjects together with general demographic questions. Before leaving the lab, subjects were administered the Holt and Laury lottery (Holt and Laury 2002), with real payments (in euros). Hence, players' final payment was the sum of their earnings from the five matrices selected, and the winnings from the lottery. The experimental sessions lasted no more than 1.5 hours and subjects earned on average 14 Euros for its completion. The minimum earning was equal to 10 Euros, while the maximum to 17.50 Euros.

### 3.2 Behavioral predictions

We formulate the following research hypotheses, around which the presentation of our results will be organized:

**Hypothesis 1** (relevance of FP): For each game type and for each variance level of HA, choice distributions in matrices with FP differ from choice distributions in the corresponding matrices without FP.

**Hypothesis 2** (relevance of FP and HA over EQ): when variance of HA is low, strategies FP and HA capture the majority of choices in games with a FP, and strategy HA captures the majority of choices in games without a FP.

**Hypothesis 3** (effect of variance): Keeping all other features fixed, when the variance of HA increases its share decreases.
Hypothesis 4 (nature of focality): the share of the FP strategy increases the more attributes defining a FP are present.

Attributes of FP:
1. payoff magnitude (“significantly” greater than the other payoffs for the row player)
2. symmetry of payoffs
3. centrality of the cell (or positioned in the main diagonal in the Weak Link)
4. Pareto-efficiency

Hypothesis 5 (Feature-based weak similarity hypothesis): a “key feature” has a similar effect in strategically different games, by influencing choice behavior in the same direction.

Hypothesis 6 (Feature-based strong similarity hypothesis): keeping all other features fixed, the choice distributions in matrices that are strategically different but similar with respect to the key features are closer - statistically - than the choice distributions of matrices that are strategically equivalent but differ with respect to the key features.

Hypothesis 7 (FP response times): the matrices with FP trigger intuitive reasoning while the matrices without FP trigger analytical reasoning: this difference appears in longer average response times for matrices without FP, ceteris paribus.

4. Results

4.1 Analysis of aggregate choices

Before discussing our hypotheses, we present an overview of the data. Figures 2 to Figure 4 report the observed frequencies, grouping the 30 games together, but analyzing matrix rows separately. Each figure reports two lines, one showing the frequencies of games with FP, the other those of games without FP. Since in the version with and without FP of the WL game the cells were the same, but the position in the matrix was changed, in these figures we have grouped together the cells according to their type, and not according to the row in which they were positioned.

Several facts emerge from the data: first, the choice distributions in the 6 versions of the same game look markedly different, showing that the presence vs. absence of the key features influences choices considerably. Second, some clear patterns can be recognized: specifically, the difference in observed frequencies between the same matrix with and without FP is evident in most cases, as are
the low share of the EQ strategy (except for the PD) and the effect of increasing the variance of HA. In particular, differences in the choice distributions of the matrix (FP, HA low var) and the matrix (XFP, HA high var) (the two extreme cases), are statistically significant in all games at least at the 0.01 level, according to a Chi-square test.

We now examine each of our hypotheses.

Hypothesis 1 (relevance of FP)
Recall that XFP is the strategy (i.e., the matrix row) corresponding to FP in the matrices in which the focal point has been removed. In our data the share of FP is always higher (and equal in only one case) than the share of XFP. The frequencies of FP, XFP, and the corresponding p-values are summarized in Table 2. As the table shows, the difference in most cases is not only statistically significant, but also quantitatively relevant. In the first three game categories - DomCol, noNE, and UniqNE - the average difference in share between FP and XFP is equal to 38%. In the case of PD and WL it lowers to 6.5%, and it equals 25.4% overall.

We made pairwise comparisons of the choice distributions by using a chi-square test. The hypothesis is confirmed for games DomCol, noNE and UniqNE: in all the 9 comparisons the difference is statistically significant (p-value < 0.01). In the PD too, the frequencies of XFP are always smaller or equal than the corresponding frequencies of FP, but the difference is statistically significant only in the pair with HA middle variance (chi-square test p-value<0.1, binomial test p-value<0.5, one-tailed). Two reasons can account for this difference: first and most importantly, the FP in game PD is weak (according to the attributes outlined in Hypothesis 4), consequently the related strategy is chosen by fewer subjects than in any other game. Second, in the PD the FP is eliminated only by breaking the symmetry, with a minimal change in payoff magnitude for the column player and no changes in the payoff of the row player.

In WL too, frequencies of FP are higher than those of XFP, but the differences are not statistically significant. A possible motivation (that will be explored in depth when discussing Hypothesis 4) is that in the WL, XFP is obtained by simply shifting the cell position without altering its content. This change apparently does not affect cell focality. We specify that the frequency for WL HA high variance is obtained by summing up the frequencies of FP and HA, since for structural reasons in that matrix two identical focal points appear, one in each of these strategies.

Concerning the relevance of the focal point, the behavior of the column players is particularly interesting. The DomCol game presents a strictly dominant strategy for the column player, while both noNE and UniqNE present a strategy yielding the highest payoff in 2 out of 3 cells and a
slightly lower payoff in the third cell: hence, a large share of FP on the part of column players indicates that its relevance is notable, given the alternatives available. The frequencies of FP, XFP and of the (quasi)-dominant strategies for the column players are presented in Table 3. When the FP is present the 100 percent of column players choose FP or the (Q)EQ strategy, while very few of the column players violate strict (or quasi) dominance when the focal point is absent, as shown by the values of the EQ shares reported in parentheses; hence, players do seem to understand the game and show compliance with basic principles of individual rationality. The choice of the FP strategy on the part of these players cannot therefore attributed to error or confusion. Since several strategies have frequency equal to 0, the chi-square test cannot be applied. We therefore only use the binomial, one-tailed test. The average difference between FP and XFP is equal to 32.8%, and in all but one case it is significant, with p-values ≤ 0.05. Altogether, our results confirm our hypothesis and show that, when the difference between the FP and XFP outcome is evident, the effect on subjects’ choice behavior is both quantitatively and statistically significant.

Hypothesis 2 (relevance of FP and HA over EQ)
We expect that, when some key features are present, players will be attracted to them more than to the equilibrium strategy. Key features provide, in players' perception, “salient” and “obvious” solutions to the game. Only when these features are absent, players may reason through the game more strategically and in some cases recognize the equilibrium strategy.
Table 4 summarizes our findings in relation to Hypothesis 2.
As hypothesized, in the case in which both key features are strong (FP, HA with low variance), these strategies capture the large majority of players’ choices, and when FP is eliminated, HA increases its attractive power leading to almost the same frequencies as in the previous case. Emblematic is the case of DomCol, where in DomCol_FP_L only 17% of players choose the Equilibrium Strategy even if it is the best response to a column player choosing a strictly dominant strategy, and in DomCol_L (where FP has been removed), HA is selected by 80% of the players. Looking at table 4, it is noteworthy that in noNE the pattern observed is similar to those of DomCol and UniqNE, although noNE does not have any pure strategy Nash equilibria. This finding is consistent with a “similarity judgment” approach (Rubinstein 1988, Leland 1994), according to which strategy C3 of noNE can be considered as an “almost-dominant” strategy since it yields the highest payoff in 2 out of 3 cases, and a not significantly lower payoff in the last case. Since choosing R3 is the best response to a column player choosing her “almost-dominant” strategy, (R3, C3) can be considered a “quasi-equilibrium” in pure strategies. This hypothesis is also supported by
the behavior of the column players, as the frequencies of column players' choices in DomCol and noNE are extremely similar, as shown in Table 3. PD and WL strongly support our hypothesis, as less than 5% of players fall outside the FP+HA combination, although in the PD HA=EQ by construction, and in WL the remaining strategy is weakly dominated. The only case that apparently contradicts our hypothesis is WL_L, where 48% choose HA and another 48% XFP. However, it has already been specified (and it will be clarified later) that in the WL game the XFP outcome has been created by simply locating the cell outside the main diagonal, with no change in payoffs. This is preliminary evidence that moving a FP cell from a central position does not reduce its focality, therefore the frequencies have to be interpreted as 96% of players choosing HA+FP, still in line with our hypothesis.

Hypothesis 3 (effect of variance)
It is reasonable to assume that a certain number of players will select the strategy with the highest expected value, assuming, more or less implicitly, that the opponent’s choices are equally likely. This behavior is relatively well known for normal form games and has been defined as “Level-1” or “Naive” (Stahl and Wilson, 1995; Costa-Gomes et al, 2001). What has not been taken into account so far is the role played by perceived risk in influencing “Level-1” types of reasoning. According to the literature, what matters for “Level-1” players is a strategy expected value. We instead assume, in line with previous findings (Warglien et al., 1999), that the attractiveness of the HA strategy (the highest expected value strategy) is also a function of its safety, therefore the higher the variance the lower the attractiveness, ceteris paribus. We first present the results for games DomCol, noNE, UniqNE and PD, and separately those for the WL. Table 5 reports data of the first four games.

The table shows that the share of HA always decreases monotonically when the variance of HA increases from low, to middle, to high, except in two cases where from middle to high it stays constant (noNE without FP and PD with FP). We test differences between matrices with HA-low variance and those with HA-high variance both by a chi-square test and a binomial one-tailed test. For games DomCol, noNE, and UniqNE both tests reveal that the differences are statistically significant (p≤0.1, except in two cases in which p≤0.5 ). Those for the PD without FP are likewise significant (p-value<0.01). PD with FP is the only case in which the difference is not significant, although the trend is the same as in the other games. The case of PD is particularly intriguing, since HA corresponds to EQ by construction, and it is weakly dominant. Hence, it is noteworthy that increasing the strategy variance without affecting its dominance induces a shift in behavior.
Figure 5 shows the frequency distribution of HA as a function of variance level. The downward slope is clearly visible. The highest share shifts from 92% to 80%, while the lowest from 43% to 20%. The average value passes from 0.68% to 0.43%.

A different approach must be used for the WL game. Here the effect of the variance cannot be observed directly, but it has to be inferred from the share of strategy COS (the strategy giving a constant payoff). Due to equilibrium constraints, while in HA low var and HA middle var, the strategies HA and FP are distinct, in HA high var two focal points appear: one in the former FP strategy and another in HA. Therefore, instead of testing whether increasing the variance of HA reduces its share, we verify whether it increases the share of COS. In WL with FP, the frequency of COS strategy passes from 2% in the low var matrix, to 8% in the middle var, to 18% in the high var matrix. In WL without FP instead, the frequency grows from 3% to 12%, to 23%. In both cases, the chi-square and the binomial tests show that the differences between low and high var matrices are statistically significant (p<0.1). We conclude that in WL too, our hypothesis is confirmed.

Hypothesis 4 (nature of focality)
While Hypothesis 2 simply postulates that the presence of focal points induces changes in behavior, this hypothesis measures the relative contribution of a series of attributes to an outcome focality. The point is relevant because it extends the notion of focal point and its properties well beyond the domain of equilibrium outcomes in (symmetric) coordination games. It has already been shown that the share of FP is always higher than that of XFP: it remains to be explained why some of the differences are more remarkable than others.

We identify 4 attributes of a game outcome that we judge relevant in determining focality:
1. payoff magnitude (“significantly” greater than the other payoffs)
2. symmetry of payoffs
3. centrality of the cell (or positioned in the main diagonal in WL)
4. Pareto-efficiency

“Payoff magnitude” refers to the magnitude of a cell payoff, when compared with other payoffs the same player can get elsewhere in the matrix. For example, in DomCol_FP_L the payoff of the focal point is “significantly” greater than the other payoffs, giving 80 experimental schillings against 40 of the second highest payoff. On the other hand in PD the payoff of the focal point is not significantly greater, as in PD_FP_L there are other 4 cells that can give the row player the same payoff as the FP cell (35 experimental schillings).
“Symmetry of payoffs” indicates that the payoffs of the two players are identical. “Centrality of the cell” refers to the position of the cell in the matrix. The FP was always located at the center of the matrix, except in the WL, where (given the presence of three symmetric cells with increasing magnitude) the symmetric cells have been positioned on the main diagonal in order of decreasing payoff magnitude.

The choice of “Pareto Efficiency” (henceforth PE) as an attribute instead of “Nash Equilibrium” differentiates our definition of a focal point from previous definitions used in the literature. We assume that players do not initially reason strategically in a game theoretic sense: therefore, we consider more relevant for the focality of an outcome to be Pareto efficient rather than an equilibrium.

A FP is an outcome (a cell) and not a strategy. Since only choices of strategies are observed and motivations for choices are not observed, the strategies yielding to a FP have been built in such a way that the outcomes other than the FP look particularly unattractive. In all games, one of the two remaining cells gives the lowest possible payoff to the row player, and in all games except the WL the remaining cell yields the second lowest payoff. Moreover, one of these two cells gives the highest possible payoff to the column player; hence, subjects should avoid picking FP if they imagine that the column player might go for her highest payoff (which in our games coincides with the equilibrium strategy for the column player).

In these games, 2 types of FP have been constructed: the first is a FP for the games DomCol, noNE, UniqNE, and WL, which satisfies the attributes of “payoff magnitude”, “symmetry of payoffs”, “centrality of the cell”, and “PE”. The second is the FP for PD, which satisfies “symmetry of payoffs”, “centrality of the cell”, and “PE”, but not “payoff magnitude”. Moreover 3 types of XFP outcomes have been constructed: the first is XFP for games DomCol, noNE, and UniqNE, which is obtained by breaking the symmetry of payoffs and by reducing their magnitude, so that the cell satisfies only the attribute of “centrality” and “PE”. The second XFP is that of WL, which is obtained simply by shifting the strategies so as to have all the cells with symmetric payoffs outside the main diagonal. Therefore this XFP outcome satisfies the attributes of “payoff magnitude”, “symmetry of payoffs”, and “PE”. The last XFP type is that of the PD, which is obtained by simply reducing the payoff of the column player. Since both payoffs were already relatively small, the payoff decrease in this case is slight. This XFP satisfies “centrality of the cell” and “PE” (in 2 out of 3 matrices).

Table 6 reports attributes and choice shares for a sample of the payoff matrices. The data show clearcut evidence that some of these attributes are an important source of focality, while others are not.
Let us first analyze PD_FP_L, where the FP strategy is not particularly successful, being chosen only by 10% of players. As the difference with PD_XFP_L is not significant, we infer that “payoff magnitude” determines cell focality, while the joint presence of “symmetry of payoffs”, “centrality of the cell”, and “PE” does not.

We then analyze games DomCol, noNE, and UniqNE. We treat them jointly since their FP and XFP cells share the same attributes. The FP strategy in these games is strongly attractive, obtaining a share that ranges from 32% to 47% in the low var case. Moreover, in all versions, differences between FP and XFP are always significant, suggesting that “symmetry of payoffs” and “payoff magnitude” (the attributes removed in XFP) are a key source of focality. On the contrary, since XFP is rarely selected, it appears that “PE” and “centrality of the cell” are two attributes of minor or no importance, as already suggested by the PD data.

In WL, FP has the strongest attractive power. In fact, when comparing the matrices with the same features, it reaches the highest frequencies. Although the share of FP is always higher than the share of XFP, the difference is never significant, again suggesting that “centrality of the cell” plays a minor role in determining focality.

Finally, we consider the separate effects of “symmetry of payoffs” and “payoff magnitude”: while the two attributes show a considerable attractive power when together, neither seems to create a focal point when present alone. In PD_XFP_L only 3% of subjects choose strategy DOM, although this contains a symmetric cell yielding an “acceptable” gain to both players. Similarly, in DomCol_XFP_L, only 2% of row players choose strategy XFP, which yields the highest (although not symmetric) gain compared with other matrix cells.

Hence, altogether these results suggest that a cell focality in a non-symmetric game is mainly due to the joint effect of “payoff magnitude” and “symmetry of payoffs”, while “centrality of the cell” and “PE” play a minor role. The two relevant attributes, when present in isolation, lose much of their attractive power.

Hypothesis 5 (Feature-based weak similarity hypothesis)

Our aim in the present study is not simply to show that Nash Equilibrium is a poor predictor of strategic behavior, but that the observed differences in choice of strategies between games sharing the same equilibrium structure follow predictable patterns governed by the presence vs. absence of the key features above defined.

Our data show that Nash Equilibrium is not able to explain observed frequencies, as we will explain in detail later on. For all our game types, the difference in choice shares between the matrix with all the key features and that without the key features is always significant with a p-value<0.01. A focal
point (as defined above) is one of such features, able to influence choices regardless of a game equilibrium structure. We have shown that even when FP is a strictly dominated strategy it is still able to attract a significant fraction of players' choices. This effect has been observed in several games, with different equilibrium structures, both symmetric and non-symmetric.

Another key feature that influences strategic behavior is HA when it is perceived as a “safe” option (low variance). In this case too, HA determines similar effects in different games, and the importance of the “safety” attribute is revealed by the emergence of an inverse relationship between the share of players choosing HA and its variance level.

Altogether, our results show that some features affect behavior in the same direction regardless of the game-theoretic properties of the strategic situation at hand. Therefore, it can be hypothesized that strategically different games may be perceived as similar when sharing some of the key features, may trigger similar emotions, similar reasoning types and similar considerations. These premises may constitute the basis for a behavioral model of cross-game similarity and game categorization. In this paper we limit ourselves to measure similarity between two games indirectly through the “property of two games to induce similar choice behaviors”. A complementary measurement method, which will be object of future research, would involve the elicitation of direct judgments of similarity on the part of players.

The next hypothesis goes further, pointing not only at the direction of the effects but also at their magnitude.

**Hypothesis 6 (Feature-based strong similarity hypothesis)**

It has been shown above that games with the same equilibrium structure that differ only for the key features generate different choice distributions. Here we propose that games with different equilibrium structures but with the same key features may generate choice distributions that are so similar to be statistically indistinguishable. This hypothesis refers to strong similarity, since it does not only concern the direction of the effects but also their magnitude.

Table 7 reports the p-values obtained by comparing games with same key features and different strategic structures, with the p-values<0.1 shaded in gray. We omit WL because its strategic structure is too different.

Data show that the key features do produce a strong similarity effect. As for games DomCol, noNE, and UniqNE, in the large majority of comparisons the frequency distributions appear indistinguishable among games sharing the same key features. While the frequencies are
significantly different when comparing the same game type with and without the key features, when these remain unaltered but the game structure changes, players’ strategic behavior remains invariant, suggesting that the difference is not perceived as such in the aggregate.

Moreover, in support to our hypothesis, it has to be noticed that the frequencies of DomCol, noNE, and UniqNE result all significantly different (according to a chi-square test) from one another only in the XFP_H case, when all the key features are removed and hence the real game structure is more visible.

These results can be interpreted in two ways: first, the key features are so salient to prevent players from perceiving a game inner strategic structure. Second, players base their strategic choices on features other than a game strategic structure (and expect other players to do as well), therefore when games share these key features they are perceived as similar and induce the same choice behavior.

4.2 Analysis of response times and correlations

In order to gain some insight into the choice process, we then turn to analyzing differences in response times. Figure 6 displays average response times, disaggregated by game class and matrix version.

Some recent studies in gaming behavior employ response time (henceforth RT) as a means to explore subjects’ decision making process, in opposition to more invasive and expensive methods based on the study of neural activity. Both Rubinstein (2007), and Piovesan and Wengstrom (2009) analyze the relationship between response times and social preferences; Rubinstein’s study finds out that fair decisions take a shorter RT than egoistic (more rational) ones, whereas Piovesan and Wengstrom (2009) seem to find the opposite relation, although the two experimental designs differ in many respects. In a recent fMRI study on gaming behavior, Kuo et al. (2009) found out that subjects took a much longer time, on average, to choose a strategy in dominance-solvable games as opposed to coordination games, and different areas of the brain activated when players faced instances of the two classes of games. On the basis of these findings, the authors suggest the existence of two different “strategizing” systems in the brain, one based on analytical reasoning and deliberation, the other based on intuition and a “meeting of the minds”.

As proposed by Kuo et al. (2009), we likewise hypothesize that matrices that present a focal point may trigger intuitive reasoning and hence require a shorter RT than matrices without focal point,
which instead are supposed to activate analytical reasoning. We expect the relation between RT and type of game not to be as notable as in Kuo et al. (2009), given that in their study the two game types were indeed strategically different, whereas in our case they only differ for the presence of a focal point, as defined earlier. Moreover, not all of our subjects chose the FP strategy, and those who did not presumably employed the same type of analytical reasoning used for games without FP. Nonetheless, the individual RT for matrices with FP is significantly shorter than the RT for matrices without FP, by a nonparametric Wilcoxon signed rank test (p=.003, one-tailed). Hence, the data support the hypothesis that matrices without focal point require more cognitive effort. Note that the significance results hold, despite the fact that several subjects did not select the focal point strategy in the matrices that contained it.

The second important finding is the increase in RT that can be observed when the variance of the HA strategy increases (from low, to middle, to high). The increasing pattern is clearly detectable in figure 6 and also in figure 7, which summarizes average response times by aggregating games according to variance level. The figure shows that increasing the variance leads to a large increase in RT. RT averages equal 17.71 in the low variance case, 20.98 in the middle variance case, and 23.66 in the high variance case. Pairwise differences of the individual RT are significant by a Wilcoxon signed rank test, one-tailed (p=0 for all cases: low var-middle var, low var-high var, and middle var-high var). We then compare the two “extreme” cases according to these findings, i.e. matrices with focal point and low variance - which should be fastest to process - and matrices without focal point and with high variance - which should require the highest cognitive effort instead. The difference in RT is indeed remarkable, increasing on average from 17.61 to 24.27 from the first to the second group. Also in this case, the difference in individual RT is significant (Wilcoxon signed rank test, p=0, one-tailed).

No significant correlations were found between individual RT, degree of risk aversion, and either number of FP choices or number of HA choices. Instead, a significant correlation was found between individual response times and number of EQ choices. The correlation coefficient is positive and equal to .273 (Spearman's rho coeff., p=.035, two-tailed) in the case choices from the modified PD (in which EQ=HA) are included, and it is equal to .331 (Spearman's rho coeff., p=.010, two-tailed) if choices from modified PD are excluded, leaving only “pure” EQ choices. This finding shows that players who were more likely to choose the equilibrium strategy EQ took longer to respond, similarly to what found by Kuo et al. (2009). These correlation results also suggest that choices of FP or HA may generally derive from an imperfect or simplified strategic reasoning rather than from beliefs in other players’ irrationality. In fact, if the latter was the case,
i.e., if players always identified correctly the equilibrium strategy even when they did not select it, we should not observe a higher response time for EQ choosers.

4.3 Equilibrium analysis

In the previous analysis we have used pure strategy Nash equilibria as a benchmark to evaluate observed frequencies. Any manipulation of the descriptive features was always referred to as strategically irrelevant as it was not altering the set of pure strategy Nash equilibria. Now we compare the descriptive power of other four stationary concepts, following Selten and Chmura (2008). We aim to understand which stationary concept best fits our data, and whether any of them is able to capture the effects due to changes in the key features.

Previous research has shown that Nash equilibrium can be a poor predictor of behavior as well as a good one, depending on many conditions (e.g., Goeree and Holt, 2001); nonetheless, in static games with complete information (like those analysed in this paper), Nash equilibrium is commonly judged a good estimator. We will show that with our data this is not the case.

In this section, we test Nash Equilibrium together with three alternative stationary concepts: Quantal Response Equilibrium (henceforth QRE; McKelvey and Palfrey 1995); action sampling equilibrium (Selten and Chmura, 2008); and payoff sampling equilibrium (Osborne and Rubinstein, 1998). Of these, only Nash is non-parametric and the others have one free parameter.

We provide a summary description of the parametric stationary concepts: according to QRE (McKelvey and Palfrey, 1995) agents make their choices based on relative expected utility and by using a quantal choice model. Moreover, players assume that other players apply the same strategies. The possibility of errors in the decision making process is taken into account.

Action sampling equilibrium is discussed at length in Selten and Chmura (2008). According to it, agents best respond to a sample (whose size is the unique parameter of the model) of observations of strategies played by their opponents. Generally the parameter is set equal to 7, which is why the model is often considered as non-parametric. By letting the parameter vary, we found the value yielding the most accurate fit of our data.

Payoff sampling (Osborne and Rubinstein 1998) is similar to action sampling. In this model, agents take one sample of actions for each pure strategy available, and then play the strategy with the highest average payoff. This model too has one parameter, since the samples have the same size.

First, we calculate the estimations with sample sizes ranging from 1 to 10 for action sampling, and
(due to computability restrictions) from 1 to 9 for payoff sampling. We then compare estimated and observed frequencies using the mean square deviation (MSD) and find the parameter value that minimises it. We found optimal sample size parameter values of 9 and 1, for action sampling and payoff sampling respectively. Similarly, we calculate QRE with values of lambda in the interval (0.01, 3). For QRE, the parameter value that best fits the data is 0.1. For QRE estimations we have used a specifically developed software: GAMBIT (McKelvey et al., 2010).

Figures 8, 9, and 10 report, divided by row, the observed and estimated frequencies. In the analysis, together with the stationary concepts, we also include the random choice model.

At first sight, Nash and action sampling seem to perform poorly. Generally, they underestimate the frequency of row 1 (the one corresponding to strategy HA) and of row 2 in the matrices with FP. On the other hand, they overestimate the frequency of row 3, generally corresponding to the equilibrium strategy. In particular, they do not seem to capture the effects of changes in the variance of HA, while Nash is unable to capture the effect of FP. Emblematic is the case of DomCol, where both Nash and action sampling give the same estimates in all six versions of the game. Often, action sampling coincides with one of the game Nash Equilibria. When more than one is available, action sampling oscillates between them, and small changes in payoffs are able to change the expected frequency from 0 to 100%.

Payoff sampling performs clearly better than both Nash and action sampling. Even small changes in the payoffs affect it, but the reactions are smoother than those observed in action sampling. Nonetheless, the estimations are not precise, and often the difference between estimated and observed frequencies exceeds 20%.

Of all the stationary concepts, QRE seems the best estimator.

Figure 11 reports the MSD scores for the four stationary concepts and the uniformly distributed random choice model. Since in several games Nash selects more than one prediction, we selected the one closest to the observed frequencies. However, results show that NE is the worst predictor.

Figure 11 confirms this finding. There is a clearcut difference in the accuracy of fit: Nash equilibrium and action sampling equilibrium perform poorly, whereas payoff sampling and QRE perform significantly better. Random choice is in between the two groups, outperforming Nash and action sampling. However, the trend of the data presented in figures 8, 9, and 10, suggests that the first is probably the result of a statistical artifact.

Differences in performances were tested by using a two sided Wilcoxon signed rank test. We compared the observed frequencies for each matrix row with the estimations of the four stationary
concepts and of the uniformly distributed random choice model. The statistical analysis confirms our previous results: QRE is significantly better than Nash, random choice, action sampling ($p \leq 0.01$), and payoff sampling ($p \leq 0.05$). The second best model is payoff sampling, which performs better than any other except QRE. Random choice performs better only than Nash ($p \leq 0.1$), while Nash and action sampling are statistically undistinguishable.

Concluding, as suggested by the analysis of aggregate choices, Nash equilibrium performs poorly and captures almost none of the effects of the descriptive features. Of all the others stationary concepts analyzed, QRE is the best estimator. This result is quite interesting given that in previous studies (Selten and Chmura, 2008) QRE is the second worst performer, better only than Nash. With the features we are taking into consideration, QRE is able to capture even minute modifications, avoiding overreactions.

5. Discussion and conclusions

We have shown that initial behavior in normal form games can be explained by a set of very simple behavioral rules that eschew optimization and that are triggered by the presence of salient features: two of such features are a “focal point” and a strategy with high expected value and low variance. These features also influence cross-game similarity perception, in such a way that subjects treat strategically equivalent games differently when these games differ with respect to the salient features, and, symmetrically, treat different games equally when these games share the same features.

More specifically, we show that the attractive power of focal points extends to asymmetric games and to non-equilibrium outcomes, and we identify two attributes (“payoff symmetry” and “payoff magnitude”) that, when jointly present, are most responsible of making an outcome focal. We also show that the presence of a strategy with high expected value and low variance (a “safe”, attractive strategy) is a strong choice attractor. Together, the strategy yielding the focal point and the safe strategy explain most of players’ choices. Subjects react in similar ways to games that present the same features, regardless of their game-theoretic category, and treat formally equivalent games differently when these differ with respect to the descriptive features.

Analysis of response times shows that matrices with focal point are faster to process than matrices without focal point, and that there is a direct relationship between variance level of the HA strategy and average response time. Equilibrium choices take longer than other choices, suggesting that out-
of-equilibrium choices are not due to beliefs in others' irrationality, but rather to the use of simplified/incorrect mental representations of the strategic situation at hand (Devetag and Warglien 2008). Finally, we explore the predictive power of Nash equilibrium and other non-standard stationary concepts: QRE performs best, followed by payoff sampling equilibrium, random choice and Nash equilibrium. None of the stationary concepts considered, despite their differing ability to capture our data, is able to fully reproduce the magnitude of feature-based changes in behavior. Future research will have to explore subjects’ categorization and deliberation process more in depth, through the use of eye-tracking techniques, and by eliciting direct similarity judgments.
References


Figure 1
Snapshot of the game interface
Figure 2

Figure 3
Figure 4

Frequency of the third row

Figure 5

Frequency distribution of HA as a function of variance level
Average response times by aggregating games according to HA variance level
Figure 8

Row 1: observed and estimated frequencies

Figure 9

Row 2: observed and estimated frequencies
Average of the sum of the squared distances between observed and estimated frequencies, for the four stationary concepts, plus the uniformly distributed random choice.
Table 1

Summary of all matrices, grouped by type of game, by level of HA variance, and by presence of FP.

Shaded the pure strategy Nash Equilibria
<table>
<thead>
<tr>
<th>Row player</th>
<th>Freq. FP</th>
<th>Freq. XFP</th>
<th>P-value chi-square</th>
<th>P-value one-tail binomial</th>
</tr>
</thead>
<tbody>
<tr>
<td>DomCol HA low</td>
<td>38%</td>
<td>2%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>DomCol HA middle</td>
<td>42%</td>
<td>7%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>DomCol HA high</td>
<td>43%</td>
<td>5%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>noNE HA low</td>
<td>32%</td>
<td>7%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>noNE HA middle</td>
<td>50%</td>
<td>7%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>noNE HA high</td>
<td>58%</td>
<td>0%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>UniqNE HA low</td>
<td>47%</td>
<td>13%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>UniqNE HA middle</td>
<td>45%</td>
<td>3%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>UniqNE HA high</td>
<td>43%</td>
<td>12%</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PD HA low</td>
<td>10%</td>
<td>5%</td>
<td>0.58</td>
<td>0.24</td>
</tr>
<tr>
<td>PD HA middle</td>
<td>17%</td>
<td>5%</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>PD HA high</td>
<td>10%</td>
<td>10%</td>
<td>0.20</td>
<td>0.50</td>
</tr>
<tr>
<td>WL HA low</td>
<td>57%</td>
<td>48%</td>
<td>0.60</td>
<td>0.46</td>
</tr>
<tr>
<td>WL HA middle</td>
<td>58%</td>
<td>50%</td>
<td>0.62</td>
<td>0.46</td>
</tr>
<tr>
<td>WL HA high</td>
<td>82%</td>
<td>77%</td>
<td>0.73</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table 2

Frequencies of FP strategies for row players, and p.values.
<table>
<thead>
<tr>
<th>Column player</th>
<th>Freq. FP (EQ)</th>
<th>Freq. XFP (EQ)</th>
<th>P-value one-tail binomial</th>
</tr>
</thead>
<tbody>
<tr>
<td>DomCol HA low</td>
<td>30% (70%)</td>
<td>5% (95%)</td>
<td>0.05</td>
</tr>
<tr>
<td>DomCol HA middle</td>
<td>50% (50%)</td>
<td>0% (100%)</td>
<td>0.00</td>
</tr>
<tr>
<td>DomCol HA high</td>
<td>35% (65%)</td>
<td>5% (95%)</td>
<td>0.02</td>
</tr>
<tr>
<td>noNE HA low</td>
<td>25% (75%)</td>
<td>0% (100%)</td>
<td>0.03</td>
</tr>
<tr>
<td>noNE HA middle</td>
<td>45% (55%)</td>
<td>0% (100%)</td>
<td>0.00</td>
</tr>
<tr>
<td>noNE HA high</td>
<td>30% (70%)</td>
<td>5% (90%)</td>
<td>0.05</td>
</tr>
<tr>
<td>UniqNE HA low</td>
<td>60% (40%)</td>
<td>15% (70%)</td>
<td>0.05</td>
</tr>
<tr>
<td>UniqNE HA middle</td>
<td>45% (55%)</td>
<td>30% (70%)</td>
<td>0.26</td>
</tr>
<tr>
<td>UniqNE HA high</td>
<td>60% (40%)</td>
<td>25% (70%)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 3

Frequencies of FP strategies for column players, and p-values. In parenthesis the frequencies of EQ and QEQ strategies in the corresponding matrices

<table>
<thead>
<tr>
<th>Game</th>
<th>Frequencies of FP + HA low var</th>
<th>Frequencies of HA with low var in matrices XFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DomCol</td>
<td>83%</td>
<td>80%</td>
</tr>
<tr>
<td>noNE</td>
<td>83%</td>
<td>73%</td>
</tr>
<tr>
<td>UniqNE</td>
<td>90%</td>
<td>75%</td>
</tr>
<tr>
<td>PD</td>
<td>97%</td>
<td>92%</td>
</tr>
<tr>
<td>WL</td>
<td>99%</td>
<td>48% (+48%)</td>
</tr>
</tbody>
</table>

Table 4
<table>
<thead>
<tr>
<th>Strategy (matrix)</th>
<th>DomCol FP</th>
<th>DomCol XFP</th>
<th>NoNE FP</th>
<th>NoNE XFP</th>
<th>UniqNE FP</th>
<th>UniqNE XFP</th>
<th>PD FP</th>
<th>PD XFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA low variance</td>
<td>45%</td>
<td>80%</td>
<td>52%</td>
<td>73%</td>
<td>43%</td>
<td>75%</td>
<td>87%</td>
<td>92%</td>
</tr>
<tr>
<td>HA middle variance</td>
<td>27%</td>
<td>48%</td>
<td>37%</td>
<td>53%</td>
<td>28%</td>
<td>68%</td>
<td>80%</td>
<td>87%</td>
</tr>
<tr>
<td>HA high variance</td>
<td>23%</td>
<td>43%</td>
<td>20%</td>
<td>53%</td>
<td>20%</td>
<td>47%</td>
<td>80%</td>
<td>68%</td>
</tr>
<tr>
<td>Chi-square test</td>
<td>0.02</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.34</td>
<td>0.00</td>
</tr>
<tr>
<td>Binomial test one-tail</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.23</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 5

Frequencies of HA strategies for row players, and p.values of the comparison between low and high variance frequencies

<table>
<thead>
<tr>
<th>Strategy (matrix)</th>
<th>PD</th>
<th>DomCol, noNE, UniqNE</th>
<th>WL</th>
<th>PD</th>
<th>DomCol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payoff magnitude</td>
<td>FP</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Symmetry of payoff</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Centrality of the cell</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Pareto efficiency</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Frequency</td>
<td>10%</td>
<td>5%</td>
<td>42%</td>
<td>7%</td>
<td>57%</td>
</tr>
</tbody>
</table>

Table 6
Attributes and choice frequencies for a sample of cells
Table 7

Comparison of games with same key features and different strategic structures. Shaded p-values ≤ 0.1.
Appendix A

Experiment instructions

INSTRUCTIONS

Welcome!
You are about to participate in an experiment on interactive decision-making funded by the R.O.C.K. (Research on Organizations, Coordination and Knowledge) research group of the University of Trento. Your privacy will be guaranteed: the results will be used and published anonymously. All your earnings during the experiment will be expressed in Experimental Currency Units (ECU). Your earnings will depend upon your performance in the experiment, according to the rules that we will explain to you shortly. You will be paid privately and in cash at the end of the experimental session. The other participants will not be informed about your earning. The experiment is divided in two, unrelated parts. The instructions for the second part will be distributed at the end of the first part. Your behavior and the earnings you obtain in the first part do not affect your earning in the second part in any way. The maximum earnings you can obtain in the experiment equals 20 Euros.

FIRST PART

The experiment consists of 30 rounds; in each round you will face an interactive decision-making situation. The word “interactive” means that the outcome of your decision will be determined by your choice and by the choice of another participant, randomly chosen. More specifically, your earnings in each decision-making situation will be determined by the combination of your choice and the choice of the participant with whom you will be paired in that round.

THE EXPERIMENTAL STRUCTURE

The structure of each interactive decision problem, henceforth GAME, will be represented by a table like the one below:

\[
\begin{array}{c|c|c}
\text{YOUR ACTIONS} & \text{THE OTHER PLAYER’S ACTIONS} \\
\text{(Row Player)} & \text{(Column Player)} \\
\hline
\text{R1} & (6,4) & (4,7) \\
\text{R2} & (3,4) & (5,6) \\
\end{array}
\]

The table has to be read as follows: you and the participant with whom you are paired will have the role, respectively, of ROW PLAYER and COLUMN PLAYER, or the other way around. The available choices of the ROW PLAYER are represented by the rows of the table (in the example R1 and R2), while the available choices of the COLUMN PLAYER are represented by the columns of the table (in the example, C1 and C2).

If your role in a round is that of ROW Player, the participant with whom you are paired will have the complementary role of COLUMN Player, and vice-versa. You will learn your role by reading the labels on the table. The label “YOUR ACTIONS” will be placed close to your role, while the label “THE OTHER PLAYER’S ACTIONS” will be close to the role of the player you are paired
with. For example, in the table like the one presented above, you have the role of ROW player, while the player with whom you are paired has the role of COLUMN player, therefore for him/her the labels are inverted.

IMPORTANT: you will keep the same role (ROW or COLUMN) in all the decisional tables of the experiment, although the participant with whom you are paired will be picked randomly (and therefore will be probably different) in each round.

Each possible combination of choices of row and column player (i.e. each possible combination of rows and columns of the table) identifies one cell in the matrix. Each cell reports two numerical values in parenthesis. These values indicate the earnings (in Experimental Currency Units) of each participant associated with that combination of choices. Conventionally, the first number represents the earnings of the ROW PLAYER (regardless of whether it is you or the other), while the second number represents the earnings of the COLUMN PLAYER.

For example: in the table below, if YOU, the ROW PLAYER, choose the row R1, and the OTHER chooses the column C2, then your earnings will be those in the cell at the intersection between row R1 and column C2: YOU (ROW Player) earn 4 ECUs and the OTHER (COLUMN Player) 7 ECUs.

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>(6,4)</td>
<td>(4,7)</td>
</tr>
<tr>
<td>R2</td>
<td>(3,4)</td>
<td>(5,6)</td>
</tr>
</tbody>
</table>

Keep in mind that you cannot choose directly the cell of the table, but only one of the rows or columns, depending on your role. Only the combination of both choices will select one and only one cell, corresponding to your earnings and to those of the other participant.

MATCHING RULES

For each decisional table, the participant with whom you are paired is randomly selected by the software. Obviously, being the matching rule random and being the number of decisional tables larger than the number of participants in the session it will happen that during the experiment you will be paired more than once with the same subject. However, you will never know the identity of the participant you are matched with, nor will you know his/her choice in a table after you have taken yours.

INFORMATION

In each of the 30 rounds the screen will show the decisional table (see Appendix B) for that round, and you will be invited to make a decision. Each table is marked by a numerical code, which will be used for the final payment. The code appears in the top-left corner of each decisional table. The top-right corner of the screen specifies the remaining time for your decision. You have to communicate
your decision by typing 1, 2, or 3 in the space “I choose row/column number”, and by clicking with the mouse the “confirm” button.

In order for the next round to start, ALL the participants must have entered their decision for the current round, therefore we ask you not to take more than 30 seconds to choose; after 30 seconds a message text in the top-right corner of the screen will invite you to write down your decision. If you delay your decision considerably you will oblige the other players to wait.

You will face 30 decisional matrices, corresponding to 30 different interactive situations. There is no relation among your choices in the different games, each game is independent from the others. At the end of the 30th round, the first part of the experiment will be completed and your earnings for this part will be determined.

THE PAYMENT

Each matrix is identified by a code. Some tags have been placed in a box, each reporting the code of one of the matrices. The experimenter will ask to one of you selected randomly to verify that the box contains 30 tags, and to verify that the codes on the tags are really different from one another. Subsequently, the experimenter will ask a different participant, selected randomly, to pick 5 of these tags from the box. Each of you will be paid according to the earnings obtained in the tables corresponding to the extracted codes. The earnings in each of the 5 selected tables will be determined matching your choice with the choice of the participant with whom you were matched in that table. Since each of 30 decisional tables of the experiment has a positive probability to be selected for the payment, we ask you to devote the same attention to all of them.

Before the experiment starts, we will ask you to answer a simple anonymous questionnaire (see Appendix C), in order to make sure that the instructions have been perfectly understood or whether some clarifications are needed. If there are incorrect answers, the relevant part of the instructions will be repeated. After the questionnaire phase is completed, the experiment will start.

It is very important that during the experiment you remain silent, and that you never communicate with the other participants, neither verbally, nor in any other way. For any doubt or problem you might have, limit yourself to raise your hand and the experimenter will approach you. If you do not remain silent or if you behave in any way that could potentially disturb the experiment, you will be asked to leave the experimental laboratory, and you will not be paid.

Thank you for your kind participation!
Appendix B

This image was printed and presented to participants as an example of the graphical interface that they would use in the experiment.
Appendix C

QUESTIONNAIRE

Dear participant,

The following questionnaire is anonymous and it has the only purpose of verifying your comprehension of the experiment rules.

We ask you to answer to the following questions. In case you are uncertain on how to respond, feel free to consult the instructions sheet.

When you are done, please raise your hand and a member of the staff will check that all your answers are correct.

Thank you for your collaboration!

Suppose you are assigned the role of ROW Player:

- If the COLUMN Player chooses strategy C2 and you choose strategy R2, how many ECU will you gain? ........... And the other player? ...........
- If you choose strategy R2, and COLUMN Player chooses strategy C3, how many ECU will he/she gain? ........... And what about you? ...........
- If the other player chooses C1, your earnings will be:
  - If you choose R1: ...........
  - If you choose R2: ...........
  - If you choose R3: ...........

Suppose you are assigned the role of COLUMN Player:

- If the ROW Player chooses strategy R2 and you choose strategy C1, how many experimental points will you gain? ........... And the other player? ...........
- If the other player chooses R1, your earnings will be:
  - If you choose C1: ...........
  - If you choose C2: ...........
  - If you choose C3: ...........

Your role (of ROW or COLUMN player) in the rounds of the experiment will change:

TRUE or FALSE

The participant with whom you will be paired will be determined randomly in each round, and you will never be matched more than once with the same participant.

TRUE or FALSE

After you have taken your decision in a table, you will be able to observe the choice of the participant with whom you were paired.

TRUE or FALSE

COLUMN Player

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>10,20</td>
<td>30,40</td>
<td>50,40</td>
</tr>
<tr>
<td>R2</td>
<td>1,2</td>
<td>3,4</td>
<td>6,3</td>
</tr>
<tr>
<td>R3</td>
<td>15,30</td>
<td>5,9</td>
<td>15,7</td>
</tr>
</tbody>
</table>
Appendix D

Experiment Instructions (Phase 2)

The sheet that was given to you shows 10 numbered ROWS, each ROW presents 2 OPTIONS: L and R. We ask you to choose one and only one of the two options in each row. Your earnings will be determined in the following way.

This is a box that contains 10 numbers, from 1 to 10, which will be used to determine your earnings. After you have made your choices, we will extract 2 numbers: the first number will determine the ROW that will be used to calculate your earnings, the second number will determine your earnings given the OPTION, L or R, that you have chosen for that ROW. Obviously, each ROW has the same probability of being chosen, equal to 1/10.

Now, pay attention to ROW 1. OPTION L pays 2 Euros if the number drawn is 1, and 1.60 Euros if the number drawn is a number that goes from 2 to 10 (extremes included). OPTION R pays 3.85 Euros if the number drawn is 1, and 0.1 Euros if the number drawn is a number that goes from 2 to 10 (extremes included). All the ROWS are similar, meaning that the earnings for both OPTIONS remain the same. The only difference is that moving towards the bottom of the table, the possibility of winning the larger amount increases for both OPTIONS; consequently, the possibility of winning the lower amount decreases. If ROW 10 is selected, there will be no need to extract the second number, because each OPTION will pay for sure the larger amount, that is, 2 Euros for OPTION L and 3.85 Euros for OPTION R.

L is the default option for all ROWS, but you can choose to switch to OPTION R by simply marking the corresponding ROW. Please note that you can switch from L to R only once and that the switch is irreversible; therefore, you have to mark just ONE ROW, which indicates that in all the ROWS above you prefer OPTION L, while in the marked ROW and in all ROWS below you prefer OPTION R. If you do not want to change, i.e., if you prefer OPTION L in all ROWS, you will not mark anything. If you always prefer OPTION R, you have to mark the first ROW. You can choose any of the 10 ROWS, but you can pass from L to R just once, therefore you can at most put 1 mark.

Once you have finished we will collect your sheet. When all participants have completed their choices, one of you will draw the two numbers from the box. Remember, the first extraction will determine the ROW that will be used to calculate everybody’s earnings, the second number will determine your earnings; the first number will be re-inserted in the box before the second number is extracted. Your earnings in this choice task will be summed up to those obtained in the first part of the experiment and the total amount will be paid to you privately at the end of the experiment.

EXAMPLE
Suppose that the ROW drawn randomly is ROW 3, and suppose you have marked one of the rows below ROW 3. Since ROW 3 is above your mark, this indicates that you prefer OPTION L for ROW 3. Then if the second drawn number is (let’s say) 5, your earning equal 1.6 Euros.

Please, answer the questions at the end of the sheet. We only need this information for statistical purposes.
<table>
<thead>
<tr>
<th>ROW</th>
<th>Option L</th>
<th>Switch from L to R</th>
<th>Option R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 € with 1 or 1.6 € with 2-10</td>
<td>□</td>
<td>3.85 € with 1 or 0.1 € with 2-10</td>
</tr>
<tr>
<td>2</td>
<td>2 € with 1-2 or 1.6 € with 3-10</td>
<td>□</td>
<td>3.85 € with 1-2 or 0.1 € with 3-10</td>
</tr>
<tr>
<td>3</td>
<td>2 € with 1-3 or 1.6 € with 4-10</td>
<td>□</td>
<td>3.85 € with 1-3 or 0.1 € with 4-10</td>
</tr>
<tr>
<td>4</td>
<td>2 € with 1-4 or 1.6 € with 5-10</td>
<td>□</td>
<td>3.85 € with 1-4 or 0.1 € with 5-10</td>
</tr>
<tr>
<td>5</td>
<td>2 € with 1-5 or 1.6 € with 6-10</td>
<td>□</td>
<td>3.85 € with 1-5 or 0.1 € with 6-10</td>
</tr>
<tr>
<td>6</td>
<td>2 € with 1-6 or 1.6 € with 7-10</td>
<td>□</td>
<td>3.85 € with 1-6 or 0.1 € with 7-10</td>
</tr>
<tr>
<td>7</td>
<td>2 € with 1-7 or 1.6 € with 8-10</td>
<td>□</td>
<td>3.85 € with 1-7 or 0.1 € with 8-10</td>
</tr>
<tr>
<td>8</td>
<td>2 € with 1-8 or 1.6 € with 9-10</td>
<td>□</td>
<td>3.85 € with 1-8 or 0.1 € with 9-10</td>
</tr>
<tr>
<td>9</td>
<td>2 € with 1-9 or 1.6 € with 10</td>
<td>□</td>
<td>3.85 € with 1-9 or 0.1 € with 10</td>
</tr>
<tr>
<td>10</td>
<td>2 € with 1-10</td>
<td>□</td>
<td>3.85 € with 1-10</td>
</tr>
</tbody>
</table>

Answer to the following questions:

In which faculty are you enrolled? ________________________________________________
In which year did you enroll? ____________________________________________________
When were you born? _______/_______/_______
Please, specify where you were born and your nationality _______________________________
______________________________________________________________________________
Specify M or F
Did you attend courses on Game Theory? __________________________________________
If yes, which courses? __________________________________________________________
Do you know what a Nash Equilibrium is? _________________________________________
If yes, in which courses have you studied it? ______________________________________