Year : 2022

Winner-Take-All Markets, Multigame Contact and Cooperation: Three Essays in Experimental Economics

Laferrière Vincent Mickael

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

WINNER-TAKE-ALL MARKETS, MULTIGAME CONTACT AND COOPERATION: THREE ESSAYS IN EXPERIMENTAL ECONOMICS

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteur en Économie

par

Vincent Mickael LAFERRIÈRE

Directeur de thèse
Prof. Joao Montez

Co-directeur de thèse
Christian Thoeni

Jury

Prof. Rafael Lalive, président
Prof. Rustamdjan Hakimov, expert interne
Prof. Maria Bigoni, experte externe

LAUSANNE
2022

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

WINNER-TAKE-ALL MARKETS, MULTIGAME CONTACT AND COOPERATION: THREE ESSAYS IN EXPERIMENTAL ECONOMICS

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteur en Économie

par

Vincent Mickael LAFERRIÈRE

Directeur de thèse
Prof. Joao Montez

Co-directeur de thèse
Christian Thoeni

Jury

Prof. Rafael Lalive, président
Prof. Rustamdjan Hakimov, expert interne
Prof. Maria Bigoni, experte externe

LAUSANNE
2022

UNIL | Université de Lausanne
HEC Lausanne
Le Décanat
Bâtiment Internef
CH-1015 Lausanne

IMPRIMATUR

Sans se prononcer sur les opinions de l'auteur, la Faculté des Hautes Etudes Commerciales de l'Université de Lausanne autorise l'impression de la thèse de Monsieur Vincent Mickael LAFERRIÈRE, titulaire d'un bachelor et d'un master en Économie politique de l'Université de Lausanne, en vue de l'obtention du grade de docteur en Économie.

La thèse est intitulée :

WINNER-TAKE-ALL MARKETS, MULTIGAME CONTACT AND COOPERATION: THREE ESSAYS IN EXPERIMENTAL ECONOMICS

Lausanne, le 12 décembre 2022

La Doyenne

Marianne SCHMID MAST

Members of the doctoral committee

Professor Joao MONTEZ
Professor of Economics, HEC, University of Lausanne
Thesis supervisor

Professor Christian THOENI
Professor of Economics, FDCA, University of Lausanne
Thesis co-supervisor

Professor Rustamdjan HAKIMOV
Associate Professor of Economics, HEC, University of Lausanne
Internal member of the doctoral committee

Professor Maria BIGONI
Professor of Economics, Department of Economics, University of Bologna
External member of the doctoral committee

University of Lausanne
Faculty of Business and Economics

PhD in Economics

I hereby certify that I have examined the doctoral thesis of

Vincent LAFERRIERE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: ______________________________ Date: 14.11.22

Prof. Joao MONTEZ
Thesis supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics

I hereby certify that I have examined the doctoral thesis of

Vincent LAFERRIERE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: ______________________ Date: November 14, 2022

Prof. Christian THOENI
Thesis co-supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics

I hereby certify that I have examined the doctoral thesis of

Vincent LAFERRIERE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: ______________________ Date: 7. 11. 2022

Prof. Rustamdjan HAKIMOV
Internal member of the doctoral committee

I hereby certify that I have examined the doctoral thesis of

Vincent LAFERRIERE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: Maria Bigoni Date: November 6, 2022

Prof. Maria BIGONI
External member of the doctoral committee

Acknowledgments

Foremost, I am deeply indebted to my two thesis supervisors, Profs. Joao Montez and Christian Thöni, for their guidance. Thanks to our previous collaboration, Prof. Christian Thöni was the person who convinced me to start this adventure. Working with him daily for all these years has been nothing but an enriching and exalting experience. His open-mindedness and passion for research make him the example every PhD student needs. I would like to thank Prof. Joao Montez for providing me with guidance throughout the process. He taught me the rigor and attention to detail required to conduct a research project. I realize how lucky I am to have had such great mentors and I will benefit from their advice for the rest of my career.

For their helpful and constructive comments, I would like to thank the two other members of the doctoral committee of this thesis: Profs. Maria Bigoni and Rustamdjan Hakimov. This thesis greatly benefited from their suggestions.

I do not want to miss the opportunity to thank two other co-authors, Prof. Catherine Roux and Dr. David Staubli. Without them, our research projects would not have been as successful and enjoyable. I am grateful to have been able to collaborate with them.

Throughout these years at the University of Lausanne, I also received support from many others. In particular, I would like to thank Prof. Christian Zehnder for allowing me to be a Lecturer, the HEC-LABEX team for their support when running all these experiments, members of the CPhD for running the doctoral program, and members of the Centre de droit public.

My family played an essential role in this achievement. I would like to thank my parents for their unconditional support throughout my graduate studies and my wife, Laetitia, for her tremendous patience and kindness.

Finally, I had the chance to share this journey with many colleagues. Especially, I would like to thank Clémence and Pauline with whom I shared unforgivable memories and will continue to do so. Hoping not to forget others, I would also like to thank the present and past, but never to be forgotten, members of the room 409 team: Alexandre, Amandine, Andrea, Camille, Dominique, Eléa, Florian, Jason, Julia, Julien, Léo, Maïté, Manon, Mélanie, Natalia, Nathan B., Nathan P., Niclas, Sarah, and Stéphanie.

Contents

Introduction

This PhD thesis contains three chapters on competitive and cooperative behavior. Every chapter makes use of laboratory experiments to study individuals' behavior in strategic situations relevant to the field of Economics. The three chapters offer experimental results, new insights, and methodological improvements; all contributing towards a better understanding of prominent topics in experimental economics.

In Chapter 1, we look at market entry decisions, and more specifically, we provide an explanation for people's tendency to enter excessively in markets with winner-take-all characteristics. In such markets, considerable payoffs are achievable but they are skewed towards a handful of firms or individuals at the top, whereas others are usually left with very little. Examples of the latter abound: network externalities create winner-take-all markets, such as social networks or software in the IT industry; highly competitive markets with few slots, such as markets for CEOs, politicians, art performers, or athletes often have these characteristics too. When entering such markets—for example, through entrepreneurship—one usually forgoes steadier earnings in a standard market. In our laboratory experiment, we tackle this question by comparing entry decisions between standard market entry games and winner-take-all games. Both conditions have equivalent expected payoffs, but payoffs are vastly more volatile in the winner-take-all condition. In this situation, traditional decision-making models with risk aversion predict more entrants in the standard market entry condition than the winner-take-all condition. However, in our experiment, we observe the opposite with the winner-take-all condition resulting in substantially more entries. We explore three candidate explanations for excess entry: blind spot, illusion of control, and joy of winning, none of which receive empirical support. We provide a novel theoretical explanation for excess entry based on Cumulative Prospect Theory and test it empirically. Our results suggest that excess entry into highly competitive environments is not caused by a genuine preference for competing, but is instead driven by probability weighting. Market entrants overweigh the small probabilities associated with the high-payoff outcomes in winner-take-all markets, while they underweigh probable failures.

In Chapters 2 and 3, we shift our attention towards cooperation, another widely studied topic in experimental economics. In particular, we investigate cooperation in situations where players repeatedly interact across multiple games; what we call "multigame contact". Although multigame contact offers great theoretical implications for cooperation and finds applications in many real-life situations, multigame contact has received surprisingly little attention. We propose and run a novel experimental design to study subjects' behavior in the presence or absence of multigame contact.

Chapter 2 reports on an experiment where subjects play a pair of indefinitely repeated prisoner's dilemma games either with the same partner, or with two different partners. In contrast to our theoretical prediction, we find no evidence that multigame contact increases overall cooperation. Nonetheless, we observe that multigame contact systematically affects behavior: subjects link their decisions across games when playing with the same partner. Multigame contact proves to be a double-edged sword, as simultaneous cooperation and simultaneous defection in the two games are both more likely under multigame contact. This suggests that the effects of multigame contact on cooperation are more complicated in practice than theory would predict.

These results call for further evidence to better understand the mechanisms at play. Chapter 3 builds on Chapter 2 and reports on a new experiment examining the effect of multigame contact on cooperation. Subjects play two indefinitely repeated games, a prisoner's dilemma and a stag hunt game, either with the same partner, or with two different partners. The second treatment dimension is the order of play within a round: either the two games are played simultaneously, or the prisoner's dilemma is played before the stag hunt game. Contrary to the theoretical predictions, multigame contact does not improve cooperation in the prisoner's dilemma. When games are played simultaneously, multigame contact even leads to less efficient outcomes. Non-credible threats can explain why multigame contact does not help, or even harms cooperation, in our experiment.

Chapter 1

Explaining Excess Entry in Winner-Take-All Markets

VINCENT LAFERRIÈRE†, DAVID STAUBLI AND CHRISTIAN THÖNI

We report experimental data from standard market entry games and winner-take-all games. At odds with traditional decision-making models with risk aversion, the winner-take-all condition results in substantially more entry than the expected-payoff-equivalent market entry game. We explore three candidate explanations for excess entry: blind spot, illusion of control, and joy of winning, none of which receive empirical support. We provide a novel theoretical explanation for excess entry based on Cumulative Prospect Theory and test it empirically. Our results suggest that excess entry into highly competitive environments is not caused by a genuine preference for competing, but is instead driven by probability weighting. Market entrants overweight the small probabilities associated with the high payoff outcomes in winner-take-all markets, while they underweight probable failures.

1.1 Introduction

Empirical research finds that most start-ups fail within a few years. Using data on the U.S. manufacturing sector, Dunne et al. (1988) report that 61.5% of newly established plants exit the market after five years and 79.6% after 10 years. Mata and Portugal (1994) confirm the high failure rates based on data of the Portuguese manufacturing industry. Hamilton (2000) finds that the majority of people who enter self-employment face lower expected earnings but higher variance than in a paid job. More recent evidence is barely more encouraging. According to the U.S. Small Business Administration Office of Advocacy's 2020 Frequently Asked Questions, the 10-year survival rate of new businesses is 33.6%, and the 15-year survival rate is 25.7%.[1]

This excess entry is at odds with early findings in the experimental literature. Standard experimental market entry games fail to find excess entry (see Rapoport et al., 1998;

†This chapter originates from a project that David Staubli and Christian Thöni started. I joined the project later and was in charge, at first, of designing and running the new experiment (Section 1.5). Once the data collected, I performed a large part of the new analysis and was heavily involved in writing Section 1.5 and rewriting the rest of the article. This chapter is a reproduction of the published article: Laferrière, V., Staubli, D., & Thöni, C. (2022). Explaining excess entry in winner-take-all markets. *Management Science*.

[1]https://cdn.advocacy.sba.gov/wp-content/uploads/2020/11/05122043/Small-Business-FAQ-2020.pdf

Sundali et al., 1995). Instead, they find the number of market entrants to be in line with Nash equilibrium predictions. As Kahneman (1988, p. 12) puts it, this tacit coordination toward the Nash equilibrium "looked almost like magic".

How come there is such a disconnect between empirical evidence from the laboratory and from the field? Entrepreneurship in the real world deviates in two main respects from the laboratory market entry game setting. First, in the real world, successful market entry depends on entrepreneurial skills. Second, many markets in the real world feature winner-take-all characteristics; that is, few competitors capture a very large share of the rewards and the remaining competitors are left with very little. Examples for the latter abound: network externalities create winner-take-all markets, namely in the IT industry, such as for social networks (e.g. Facebook, Twitter) or software (e.g. operating systems). Also, industries with one or only few slots display winner-take-all characteristics. Examples are highly competitive environments like markets for CEOs or politicians. Finally, we see a particularly high concentration of rewards for a few competitors in the entertainment business, such as performing arts (music, dance, etc.) or sports (tennis, football, etc.).

More recent research on market entry in competitive environments has focused on the role of beliefs about relative skill. Camerer and Lovallo (1999) report results from laboratory experiments showing that overconfidence can drive excess entry in a setting where market returns depend on skill. Moore and Cain (2007) build on Camerer and Lovallo (1999) and manipulate the difficulty of the task to determine skill levels. They find that skill-dependent market returns lead to excess entry only in the easy-task condition, but not in the difficult-task condition. Moore et al. (2007), Dorfman et al. (2013), and Cain et al. (2015) confirm the finding that competitions over easy tasks are most likely to produce excess entry. Wu and Knott (2006) study aggregated market data on entry decisions of entrepreneurs and find excess entry in markets where there is a high degree of uncertainty about ability and low demand uncertainty. Comparing the evolution of overconfidence between strategic and nonstrategic environments, Brilon et al. (2019) find that overconfidence persists when subjects can strategically send signals about their skill and opt out of competition (strategic). Overconfidence vanishes, however, when subjects are forced to compete with each other and cannot strategically send signals (nonstrategic). Morgan et al. (2016) allow subjects to make postentry investment decisions to investigate the role of natural and strategic risk. They show that adding natural risk to market entry leads to a slight increase in the frequency of market entry and to a much higher postentry investment.[2]

In Danz (2020), subjects have to choose between a piece-rate compensation and a competitive tournament after observing their competitors' past performances. By manipulating ex-post information about their competitors' previous tasks, Danz shows that hindsight bias generates overplacement and increases subjects' valuations of tournament participation.

So far, the significant body of literature drawing on insight from behavioral economics to explore the roots of entrepreneurship is inconclusive. There is no clear "smoking gun" to account for the excess entry in the real world documented in the empirical literature (see Åstebro et al., 2014, for a review).

Closest to our work is Fischbacher and Thöni (2008), who modify the payout scheme

[2]Experimental research has also found overbidding in rent-seeking contests (see Dechenaux et al., 2015, for a review). Cason et al. (2020) show that efforts are consistently higher in winner-take-all contests compared to contests with proportional prizes. Sheremeta (2018) explores multiple behavioral explanations and finds impulsivity to be the primary driver of overbidding.

to incorporate the winner-take-all characteristics of many real world markets. Deviating from standard market entry games, they do not distribute the market evenly among the entrants. Instead, they randomly determine a single winner who takes the entire market. In this winner-take-all setting, they find excess entry well beyond Nash equilibrium predictions, absent any role for skill. However, since they do not contrast their results with an expected-payoff-equivalent standard market entry game, their setting cannot single out the winner-take-all feature as the cause of excess entry.

In this paper, we fill this gap by running both a winner-take-all game (WTA) and an expected-payoff-equivalent standard market entry game (MEG). This allows us to isolate the effect of the winner-take-all feature. We find a large and statistically significant treatment effect: the WTA condition creates more market entry than the MEG condition. This treatment effect is puzzling. As market entry in the WTA condition offers the same expected payoff (henceforth, "expected payoff" always refers to "expected monetary payoff") but more variance than in the MEG condition, we would expect just the contrary based on standard expected utility models with risk aversion.

We investigate a number of explanations for excess entry found in the literature. First, we rule out utility curvature as an explanation. To account for excess entry into WTA markets, convexity of utility functions would be required, which is incompatible with findings in the empirical literature. Then we find that neither wrong beliefs about other players' entry probability (henceforth, "blind spot", following Camerer and Lovallo, 1999), nor erroneous beliefs that one can influence random processes ("illusion of control", see Langer, 1975) can explain the treatment effect. Furthermore, we address the possibility that the competition-against-others structure of the WTA game provides an extra utility associated with successful market entry ("joy of winning"). Running two additional conditions that preserve the payoff consequences but eliminate the competitive element, we find that the treatment effect does not disappear when the competitive element is removed.

We finally investigate a new approach to understand excess entry through the lens of Cumulative Prospect Theory (CPT; Kahneman and Tversky, 1979, and Tversky and Kahneman, 1992). We show that—depending on the parameter specifications—CPT can be a powerful explanation for excess entry in winner-take-all markets. To explore the predictive power of CPT, we ran an additional experiment where we estimate subjects' CPT parameters and relate them to their entry behavior. Our parameter estimates are in line with previous findings from the literature. We show that our pooled parameter estimates are consistent with the treatment effect. The driver of the effect is the inverted S-shaped weighting of cumulative probabilities. This translates into overweighting of small probabilities of winning large payoffs in the winner-take-all game. Further, we classify subjects as expected utility theory (EUT) types or CPT types using a finite mixture model approach. Relating this classification to market entry behavior, we are able to show that excess entry in the winner-take-all condition is primarily driven by CPT types.

Explaining excess entry by the biased perception of probabilities as described by the weighting of cumulative probabilities (henceforth "probability weighting") is promising for mainly two reasons. First, the inverted S-shaped weighting function is a well established finding of the behavioral economics literature (see Abdellaoui, 2000; Bruhin et al., 2010a; Bruhin et al., 2018; Fehr-Duda & Epper, 2012; Fehr-Duda et al., 2011; Fennema & Wakker, 1997; Harrison & Rutström, 2009; Wakker, 2010; G. Wu & Gonzalez, 1996). Second, probability weighting offers a complementary explanation to overconfidence as a driving force behind excess entry. This is important because there are environments in which overconfidence is unlikely to arise (e.g. in competitions on difficult tasks).[3] We further show that the predictive success of probability weighting in a CPT framework extends to strategic situations. Whereas there is a large number of experimental studies on CPT applied to individual choice problems, empirical research on CPT preferences in strategic contexts is surprisingly sparse.[4]

The remainder of the paper is organized as follows. The next section presents the experimental design. Section 1.3 presents the main results. In Section 1.4 we investigate explanations for the treatment effect: blind spot, illusion of control, and joy of winning. In Section 1.5 we derive predictions for CPT and put the explanation to an empirical test. Section 1.6 concludes.

1.2 Experimental design

1.2.1 The game

In our game, n players decide simultaneously whether to enter a market ($Entry_i = 1$) or to stay out ($Entry_i = 0$). If a player decides to stay out, then she gets a fixed payoff I. Staying out can be interpreted as earning a fixed payment for regular employment. In our setting, n is equal to 14 and I is equal to 45 experimental currency units (ECUs). If a subject enters the market, then her payoff depends on the number of market entrants ($E = \sum_{j=1}^{n} Entry_j$). The payoffs for entering the market follow the function described in Table 1.1, where E is the number market entrants, $\Pi(E)$ is the market payoff, and $\pi(E)$ is the market payoff divided by the number of entrants ($= \Pi(E)/E$).

Table 1.1: Payoff function for entrants, dependent on the number of entrants E

E	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\Pi(E)$	100	130	155	175	190	200	205	209	212	214	215	215	215	215
$\pi(E)$	100	65	51.7	43.8	38	33.3	29.3	26.1	23.6	21.4	19.5	17.9	16.5	15.4

In the market entry game (MEG) condition, each entrant earns a payoff $\pi(E)$. To illustrate, if five players enter the market, then the market payoff is 190 and each entrant earns 38. In the winner-take-all (WTA) condition, the market payoff is randomly assigned to one entrant only. That is, one entrant earns $\Pi(E)$, while all other entrants earn zero. Each entrant receives the payoff of $\Pi(E)$ with probability $1/E$ and zero otherwise. To illustrate, if five players enter the market, then one randomly selected entrant gets the total market payoff of 190 and the four other players get a payoff of 0. Each entrant has a probability of 0.2 to be assigned the market payoff of 190. Consequently, given the number of entrants, the *expected* payoff for market entry is the same in both conditions.

Similarly to the payoff function in Fischbacher and Thöni (2008), $\Pi(E)$ is increasing and concave in E. In our case, however, $\Pi(E)$ flattens out at $E = 11$. In reality, total market returns may or may not be increasing in the number of competitors. As argued

[3]Overconfidence about skill is far from being a ubiquitous phenomenon. It depends on subjects' characteristics (Schulz & Thöni, 2016), and whether the task is familiar or non-familiar to the subjects (Hoelzl & Rustichini, 2005).

[4]Ernst and Thöni (2013) show that behavior in all-pay auctions is consistent with reference-dependent utilities as proposed by CPT; Brünner et al. (2019) find evidence for CPT preferences in bidding behavior in online auctions; Nguyen et al. (2016) relate prospect theory preferences to trust.

by Frank and Cook (1995), more competitors may increase the market value for several reasons. First, the winner in a competition with a large number of competitors will likely perform better than the winner in a competition with a small number of competitors. Thus, if not only relative but also absolute performance is rewarded, the winner's prize increases in E. This is plausible, for instance, in the music business or in sports with absolute performance measures (e.g., world records in athletics). Second, particularly in sports or the performing arts, more contestants increase public attention and media coverage, which in turn increases the rewards for the top performers. The marginal effect of an additional entry on market volume, however, is assumed to be decreasing. This gives rise to strictly decreasing expected payoffs *per entrant*, $\Pi(E)/E$, as the number of entrants increases. Due to limited potential market demand, an entrant therefore imposes an externality on other entrants by lowering their expected payoffs.

1.2.2 Nash equilibria

In this section, we derive Nash equilibria in pure and in mixed strategies. We start with the simplest case, assuming that players maximize their expected payoff. As the MEG and WTA conditions are equivalent under this assumption, the following analysis holds for both games. The pure-strategy Nash equilibria are straightforward. As long as there are less than three entrants, it is a best response to enter. When three other players enter the market, entry is no longer profitable, as a fourth entrant would forgo 45 and earn $\pi(4) = 43.8$. Any permutation of three players entering the market and the other 11 players staying out constitutes a Nash equilibrium in pure strategies. For the symmetric mixed-strategy Nash equilibrium, a player's entry probability, p^*, must satisfy:

$$\sum_{E=1}^{n} \binom{n-1}{E-1} {p^*}^{E-1} (1-p^*)^{n-E} v(E) = u(I), \tag{1.1}$$

where $v(E) = u(\pi(E))$ in the MEG condition and $v(E) = [1/E]u(\Pi(E)) + [(E-1)/E]u(0)$ in the WTA condition. The condition equates the value of staying out (right-hand side) to the value of entering the market if all *other* players enter the market with probability p^* (left-hand side). Assuming expected payoff maximization (linear $u()$, risk neutrality) and inserting the parameters of our experimental setting for n, I, $\Pi(E)$, and $\pi(E)$ yields $p^* = 0.25$. If all players enter with probability p^* and stay out with probability $1-p^*$, then the expected number of market entrants is 3.52. Besides the symmetric equilibrium, there are numerous asymmetric equilibria in mixed strategies. They feature some players who always enter, some who always stay out, and some who randomize. In any Nash equilibrium, the average number of entrants is between 3 and 3.80. The former corresponds to the equilibrium in pure strategies and the latter corresponds to an asymmetric Nash equilibrium in mixed strategies where four players enter with probability 0.95 and the other ten player stay out. Any mixed-strategy Nash equilibrium implies more market entrants than the pure-strategy equilibrium. This is because, in our parameterization, the third entrant in the pure-strategy equilibrium is clearly better off entering than staying out and the fourth entrant is almost indifferent between entering and staying out. For a detailed analysis of the Nash equilibria, we refer the reader to Fischbacher and Thöni (2008).

The effect of risk preferences as explained by expected utility theory is straightforward. If we assume risk aversion (concave $u()$), then entry becomes less attractive in both the MEG and WTA conditions and the equilibrium is characterized by fewer expected entrants. The effect is, however, stronger for the WTA condition and the theory predicts more entrants in the MEG condition than in the WTA condition. The opposite is the case for convex utility. We will analyze the equilibria under richer preference assumptions in more detail in Section 1.5.

The asymmetric equilibria mentioned above are indicative for what happens if we assume heterogeneous agents. We restrict our attention to a model with two types of players and complete information. Consider a group with z risk-neutral players (linear $u()$) and $14 - z$ risk-averse players (concave $u()$). Because Equation (1.1) cannot be satisfied for both types of players at the same time, only one of the two types is willing to play a mixed strategy in equilibrium. If $z > 3$, then there is an equilibrium in which the z risk neutral players play a mixed strategy while the risk averse players do not enter. The expected number of entrants would then be equivalent to the game with only z players, and we can use Equation (1.1) to derive the entry probabilities. It turns out that the expected number of entrants is fairly insensitive to changes in group size (as long as $z > 3$), gradually decreasing from 3.80 for $z = 4$ to 3.52 for $z = 14$. If $z = 3$ then there is no mixed-strategy equilibrium, while for $z < 3$ there are mixed-strategy equilibria in which the risk neutral players always enter and (some of) the risk averse players use a mixed strategy. The same logic applies if the group contains z players with risk appetite (convex $u()$) and $14 - z$ risk-neutral/risk-averse players.

1.2.3 Experimental procedures

We repeat the game for 20 periods in groups of size 14. We use a partner matching; that is, subjects are allocated to groups of 14 at the beginning of the session and interact only within this group during the session.[5] For all groups, there are two phases: a phase of 20 periods in the WTA condition, and a phase of 20 periods in the MEG condition. To control for order effects, about half of the groups played the WTA condition first and the MEG condition second; for the remaining groups we reversed the order. In total, we observe 6,160 entry decisions from 154 subjects.

At the beginning of every period, we elicit subjects' beliefs about how many other subjects would enter the market (hereafter, we will refer to the entrants apart from the subject considered as "other entrants"). We incentivize truthful revelation of beliefs by rewarding a correct guess with five ECUs. Experimental research shows that belief accuracy increases when correct beliefs are rewarded (see Gächter & Renner, 2010; Wright & Aboul-Ezz, 1988). However, in order to keep subjects' focus on the main game, the reward for a correct belief is rather small compared with the payoffs of the game.[6] In the WTA condition, we use the computer to generate a uniformly distributed random number between 0 and 100 for each subject. Among the subjects who chose to enter the market, the one with the highest random number wins the winner-take-all competition and the corresponding payoff. All competitors learn their random number and the random number of the winner in their group at the end of each period. In both conditions, subjects are also informed about the number of other entrants at the end of each period. The strategies

[5] In the first session with 28 subjects we ran a stranger matching with two groups. The results are almost identical for the two matching protocols.

[6] The design of our belief elicitation stage is deliberately kept simple. Eliciting full probability distributions would require that subjects allocate probability mass across 14 outcomes (0–13 other entrants) in each round. This would distract subjects from the main task. The downside of our method is that we elicit only the mode of the distribution, which is consistent with a wide range of probability distributions. For a discussion of elicitation methods and their biases see Armantier and Treich (2013).

“market entry” and “staying out” are neutrally denoted as “Alternative A” and “Alternative B”.

Subjects were paid out the sum of the payoffs of the 40 market entry decisions as well as the rewards for the correct guesses on the number of other entrants. The accumulated earnings in ECUs were converted to real money at the end of the experiment at an exchange rate of CHF 1.5 (≈ USD 1.5) per 100 ECUs.

The sessions were run in the laboratory of the University of St. Gallen with undergraduate students recruited with ORSEE (Greiner, 2015). The experiment was programmed in z-Tree (Fischbacher, 2007). Subjects were randomly allocated to the sessions so that they could not infer with whom they would interact. Sessions lasted between 70 and 80 minutes with an average pay per subject of CHF 37 (≈ USD 37).

1.3 Results

Figure 1.1 shows that the WTA condition results in more market entry compared with the MEG condition. In phase 1 (phase 2) the average number of entrants across the 20 periods is 6.42 (5.67) in the WTA condition and 4.41 (4.38) in the MEG condition. The treatment effect is highly significant in both phases ($p = 0.008$, exact Wilcoxon-Mann-Whitney test using group averages as observations). Additionally, we can analyze the treatment effect within group. Pooling all groups across the two phases, the average number of entrants is 4.39 in the MEG condition and 6.08 in the WTA condition. An exact Wilcoxon signed-rank test on the within-group treatment effect yields a p-value of 0.002. Compared with the symmetric Nash equilibrium in mixed strategies (depicted by the horizontal lines in Figure 1.1), we observe excess entry in both conditions and in both phases.

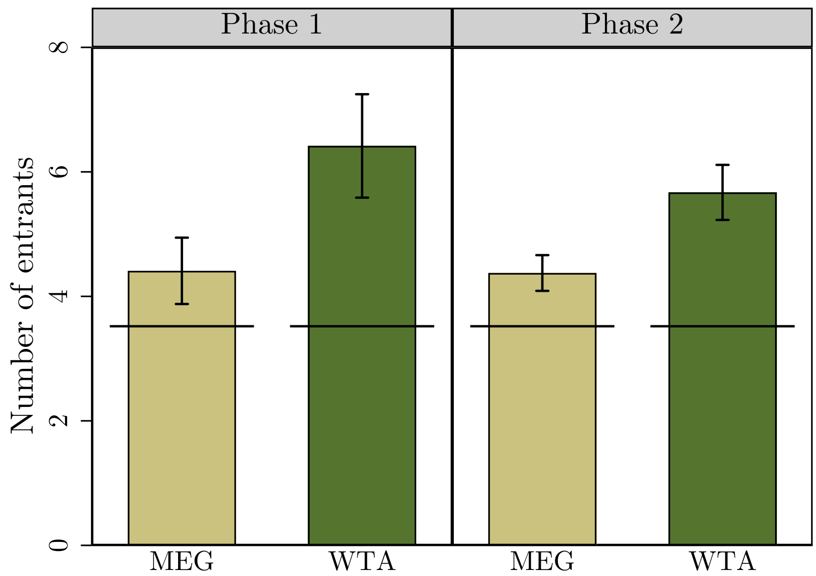


Figure 1.1: Average number of entrants in the two conditions per phase. The horizontal lines show the number of entrants predicted by the symmetric mixed-strategy Nash equilibrium with risk neutral players. Spikes show 95% confidence intervals, clustered on group.

Figure 1.2 shows market entry over time. It displays the average entry frequency across the 20 periods and both phases. The entry frequencies in both conditions become

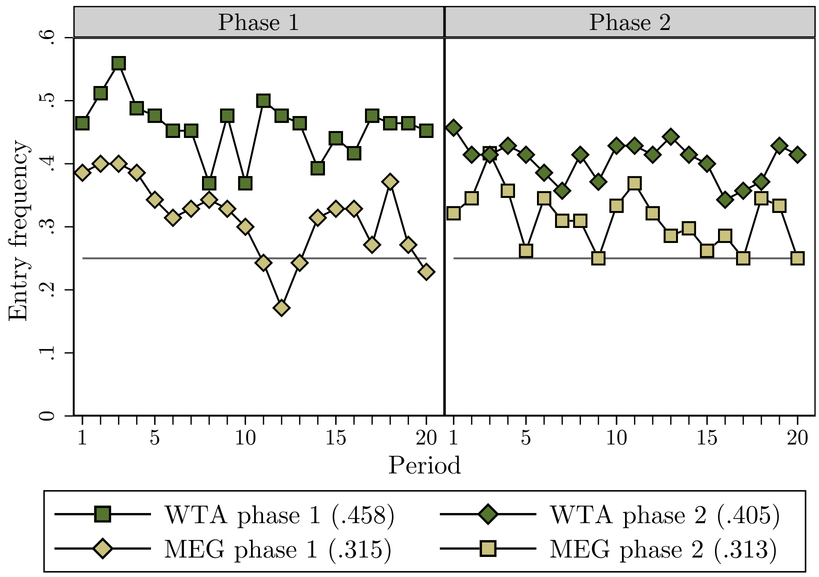


Figure 1.2: Frequencies of market entry in the two conditions per period across phase 1 and phase 2 (average over all periods in parentheses). The horizontal lines show the entry probability predicted by the symmetric mixed-strategy Nash equilibrium with risk-neutral players.

somewhat lower over time. While there is some indication that the MEG converges towards the predicted entry frequency, excess entry remains substantial in the WTA condition. In Table 1.2 we report linear probability models to test for time and order effects. Model (1) shows that both within phase and between phases we observe a significantly negative effect on entry. In Model (2) we interact period and the dummy for the second phase with the treatment dummy. The results suggest that the negative trend across the 20 periods within a phase is mainly driven by the MEG condition, whereas the reduction in entry frequency across phases is mainly driven by the WTA condition.

There is considerable heterogeneity in the entry behavior across subjects. The standard deviation of the individual number of entries over the 20 periods in the WTA condition is 6.41 (average individual number of entries is 8.68; data from both phases). About 25% of the subjects enter two or fewer times during the 20 periods, while 10% of the subjects enter in every period. In the MEG condition, the standard deviation is somewhat lower (5.48), and 30% of the subjects enter two or fewer times, while only 2.6% enter in every period. For histograms of individual entry frequency see Figure 1.A.2 in the appendix (p. 30).

Taken together, we observe excess entry relative to the Nash equilibrium in both conditions. In the MEG condition, we observe a negative trend over time such that, in the last periods, the average number of entrants is no longer significantly different from the prediction ($p = 0.453$ in the final two periods, both phases). In the WTA condition, however, the corresponding difference remains highly significant ($p = 0.004$).

Our results on the WTA condition replicate and confirm the findings of Fischbacher and Thöni (2008), and our comparison with the expected-payoff-equivalent MEG condition allows us to identify the winner-take-all characteristics as the causal source of excess entry. This is puzzling because—for any number of other entrants—market entry in the MEG and WTA conditions offer the same expected payoff, but the latter comes with more

Table 1.2: Linear probability models for entry

	Dependent variable: Entry	
	(1)	(2)
MEG	-0.118^{**}	-0.116^{*}
	(0.011)	(0.042)
Period	-0.003^{**}	-0.002
	(0.001)	(0.002)
Phase 2	-0.028^{*}	-0.053^{*}
	(0.011)	(0.023)
MEG × Period		-0.003
		(0.003)
MEG × Phase 2		0.051
		(0.031)
Constant	0.482^{**}	0.480^{**}
	(0.021)	(0.036)
F-test	38.0	44.3
Prob $> F$	0.000	0.000
R^2	0.018	0.019
N	6,160	6,160

Notes: OLS estimates. Dependent variable is an individual entry decision; independent variables are dummies for the treatment (MEG) and phase 2, period, and interactions. Robust standard errors, clustered on matching group, in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

variance. Thus, the results are at odds with legions of studies showing that subjects generally display risk-averse behavior. Fischbacher and Thöni (2008) speculated that subjects either have "illusion of control" or that they gain extra utility from the thrill of competition in the WTA condition relative to the MEG condition. In the following, we present additional data analyses and new experimental measures to investigate candidate explanations for the excessive entry behavior in the WTA condition relative to the MEG condition.

1.4 Candidate explanations for excess entry

In this section, we test the three hypotheses: blind spot, illusion of control, and joy of winning. For the first, we make use of the beliefs elicited in the main experiment; for the second and third, we present additional experimental measures.

1.4.1 Blind spot

In this section, we address the blind spot explanation, that is, excess entry due to miscalibrated beliefs about other players' entry probability in the two conditions. Both conditions are strategic games: The fewer other players a particular player expects to enter, the more attractive market entry becomes. The treatment effect could be caused by biased beliefs. If players mistakenly believe that other players are less likely to enter in the WTA relative to the MEG condition, then the optimal response may be to enter more frequently in the WTA than in the MEG condition. This could arise from wrong beliefs about other players' risk preferences. As the WTA condition is subject to risk, overestimation of other players' risk aversion leads to underestimation of their entry probability.

To address the blind spot explanation we elicited the subjects' beliefs about the number of other entrants. We define the variable *Belief error* as the actual number of other entrants minus a subject's belief about the number of other entrants. It is thus a measure for the *underestimation* of the number of other entrants. Over the course of the 20 periods, beliefs become more accurate and there is no systematic difference across conditions in terms of belief accuracy.[7] The blind spot explanation requires systematic underestimation in the WTA condition. Figure 1.3 shows the mean belief error across the 20 periods for the MEG and the WTA conditions. The mean belief error fluctuates around zero. That is, on average subjects' beliefs do not seem to be systematically biased.

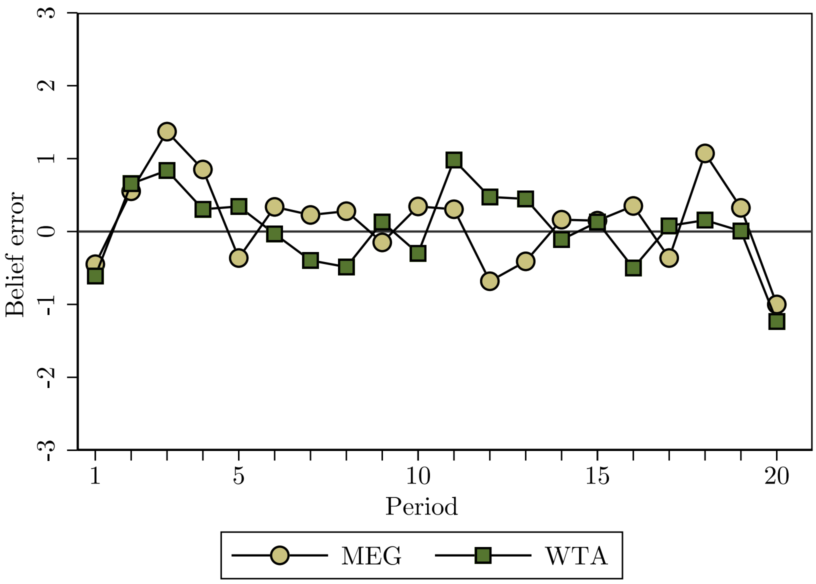


Figure 1.3: The figure shows the mean belief error in the MEG and the WTA conditions across the 20 periods (both phases). Positive (negative) values indicate that subjects underestimate (overestimate) the number of other entrants.

Table 1.3 shows ordinary least squares (OLS) regressions with the *Belief error* as the dependent variable. The estimate of the constant in Model (1) reflects the average *Belief error* across all subjects and all periods. The estimate indicates that subjects underestimate the number of other entrants by 0.095, which is not significantly different from zero. On average, beliefs are therefore very accurate. The result in Model (2) indicates that subjects underestimate the number of other entrants somewhat less—by 0.102—in the WTA condition compared with the MEG condition (statistically insignificant). To provide evidence for the blind spot explanation, we would need subjects to underestimate

[7] In the first five periods, the absolute difference between the realized number of entrants and the belief is 2.08 on average, while in the last five periods the corresponding value is 1.64. If we regress the absolute belief error on period, a dummy for the WTA condition, and a dummy for the second phase, then we observe a negative coefficient for period ($\beta = -0.028; p = 0.007$), while the other two coefficients are not significantly different from zero.

the number of other entrants more strongly in the WTA condition. The negative sign of the coefficient estimate goes therefore against the blind spot explanation.

Table 1.3: Explaining belief errors

	Dependent variable: Belief error		
	(1)	(2)	(3)
WTA (β_1)		-0.102	0.265^{*}
		(0.060)	(0.084)
Entry (β_2)			1.606^{**}
			(0.177)
Entry × WTA (β_3)			-1.290^{**}
			(0.263)
Constant (β_0)	0.095	0.145^{*}	-0.358^{**}
	(0.055)	(0.060)	(0.054)
F-test	–	2.9	66.6
Prob $> F$	–	0.124	0.000
R^2	–	0.001	0.060
N	6,160	6,160	6,160

Notes: OLS estimates. The dependent variable is the belief error, number of other entrants – beliefs. Robust standard errors, clustered on matching group, are in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

This does not suffice to rule out the blind spot explanation. The result in Model (2) is consistent with the following beliefs in the WTA condition: Those who enter frequently underestimate the number of other entrants, and those who stay out overestimate the number of other entrants (thereby keeping average beliefs unbiased). We investigate this possibility by adding an interaction of the treatment dummy with the individual entry decision in Model (3).

Table 1.4: Mean belief error as a function of the WTA dummy and the Entry dummy

	WTA = 0	WTA = 1
$Entry = 0$	$\hat{\beta}_0 = -0.358^{**}$	$\hat{\beta}_0 + \hat{\beta}_1 = -0.093$
$Entry = 1$	$\hat{\beta}_0 + \hat{\beta}_2 = 1.247^{**}$	$\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 = 0.222^{+}$

Notes: $\hat{\beta}_j$ is the estimate of coefficient β_j from Model (3) of Table 1.3. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Model (3) allows us to differentiate the effect of the beliefs by entry decision and condition. Table 1.4 provides the respective coefficient estimates. We find that subjects who enter in a given period display a positive belief error and subjects who do not enter in a given period display a negative belief error. To tackle the blind spot explanation, we look at the difference in belief errors between the MEG and WTA conditions only for observations with $Entry = 1$. That is, we compare the lower-left cell and the lower-right cell of Table 1.4. The difference is equal to $\hat{\beta}_1 + \hat{\beta}_3 = -1.025$. This indicates that entrants underestimate the number of other entrants more strongly by 1.025 in the MEG condition compared to the WTA condition.

Again, the negative sign of the estimate, $\hat{\beta}_1 + \hat{\beta}_3$, goes against the blind spot explanation. To substantiate blind spot as an explanation for the treatment effect, we would need subjects to underestimate more strongly (or overestimate less strongly) the number of other entrants in the WTA condition compared to the MEG condition. Thus, the treatment effect cannot be explained by subjects erroneously underestimating the number of entrants in the WTA condition. On the contrary, if anything, subjects' beliefs strengthen the puzzle.

1.4.2 Illusion of control

Subjects might believe they have some influence over the random process of the market payoff in the WTA condition. Such an "illusion of control" (IOC; Langer, 1975) might increase perceived profitability of market entry in the WTA condition. To investigate whether IOC drives the treatment effect, we introduce a measure for IOC and relate individual differences to market entry behavior.

We measure IOC with all participants after they completed the MEG and WTA phases. We identify IOC by the willingness to pay for a particular lottery ticket over other lottery tickets with the same objective winning probabilities. The 28 subjects in a session were presented with 28 symbols such as "%", "$", or "@". In the session, we had an urn containing the complete set of symbols (28 cards), and subjects were told that two of the 28 symbols would be drawn at random. The two subjects with the respective symbols won additional CHF 50 (which is significantly more than the average earnings per subject in the main experiment).

After all subjects picked a symbol from the 28 symbols, they were informed that in order to have a one-to-one matching between symbols and subjects we would have to allocate the unchosen symbols to some of the subjects who chose the same symbols as others. Before subjects knew whether their preferred symbol was chosen more than once, all subjects had to indicate their willingness to pay to keep their symbol. Subjects were endowed with 800 ECUs. We explained to the subjects that there would be a sealed-bid second-price auction for their symbol and they were asked to submit a bid. We explained to the subjects that, in this auction format, it would be optimal to bid exactly their willingness to pay.

Objectively, every symbol has the same probability to be chosen at random, and subjects are aware of that. Consequently, there is no objective reason to pay to keep a particular symbol instead of being allocated another as-yet-unchosen symbol. If subjects nevertheless indicate a willingness to pay for their symbol, we interpret this as IOC.

On average subjects bid 114 ECUs with a standard deviation of 218. This implies that there is substantial variance: 43% of the subjects display no willingness to pay at all, 56% bid at most 5 ECUs, and 4% are willing to bid their full endowment of 800. Relative to the extra pay for winning the lottery (CHF 50 = 3333 ECUs), the vast majority of subjects have a rather low willingness to pay (see Figure 1.A.3 in the appendix, p. 31, for a histogram of the bids).

What matters for our purpose, however, is whether our measure for IOC has predictive power for excess entry. In Table 1.5, we use OLS regressions to investigate the effect on (1) the individual entry frequency in the WTA condition, and (2) the treatment effect (individual entry frequency in the WTA condition minus individual entry frequency in the

MEG condition), both measured in percent. The results show that estimated coefficients in both models are close to zero and far from significant. These results suggest that illusion of control does not offer an explanation for excess entry.

Table 1.5: Excess entry and illusion of control

	Dependent variable: Entry frequency (in %)	
	WTA	WTA–MEG
Illusion of control	0.010	0.008
	(0.016)	(0.011)
Constant	42.290**	11.192**
	(1.899)	(1.568)
F-test	0.4	0.4
Prob $> F$	0.555	0.524
R^2	0.004	0.003
N	154	154

Notes: OLS estimates. The dependent variable is the individual entry frequency in WTA, or the treatment effect (frequency of entry in WTA minus frequency of entry in MEG). Both measures in percent. The independent variable is the bid in the auction for the symbol. Robust standard errors, clustered on matching group, are in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

1.4.3 Joy of winning

A main conjecture that Fischbacher and Thöni (2008) outline in their conclusion to explain excess entry is the thrill of winning a competition against others provided by the WTA condition. In the literature, this phenomenon is often referred to as "joy of winning". The concept posits that winning a prize in a competitive environment generates more utility than winning the same prize in a simple lottery. Researchers have argued for the existence of joy of winning in contests and auctions. Studies include Amaldoss and Rapoport (2009), Brookins and Ryvkin (2014), Dohmen et al. (2011), Goeree et al. (2002), Herbst (2016), Herrmann and Orzen (2008), and Sheremeta (2010).

In our setting, the idea can be formalized by adding an extra payoff for the winner. This payoff represents the nonpecuniary joy of winning against the other market entrants. The mixed-strategy Nash equilibrium condition is then written as follows:

$$\sum_{E=1}^{n} \binom{n-1}{E-1} {p^*}^{E-1} (1-p^*)^{n-E} \frac{1}{E} \left[\Pi(E) + jow\right] = I. \tag{1.2}$$

The main difference to Equation (1.1) is the term jow which is added to the winning payoff in the WTA condition. For simplicity, Equation (1.2) assumes linear utility functions. Predicting WTA entry probabilities in the range of 0.405 to 0.458 (see Figure 1.2) would require a jow term of about 80.

To address this potential explanation for excess entry, we run additional experimental sessions with two alternative conditions. The conditions are designed to preserve the monetary incentives of the WTA condition while eliminating the competitive element. We replace the strategic WTA game by a non-strategic individual choice problem; that is, in both conditions, a subject's payoff is independent of the decisions of the other subjects in the room. Instead, they depend on the frequency distribution of the number of market entrants as observed in the WTA condition in our main experiment (hereafter, "WTA condition" always refers to the WTA condition of the main experiment as described in Section 1.2.1).

The first treatment presents subjects with a situation where they have to form beliefs about the market entry behavior of the participants in the WTA condition in the past. We denote this treatment as AMB, for ambiguous winning probabilities. The second treatment goes one step further and presents the subjects with objective winning probabilities (OBJ) taken from the WTA condition.

In the AMB condition subjects decide in each of the 20 periods between Alternative A and Alternative B. The payoff function is identical to the WTA condition: Alternative B pays 45 ECUs for sure, while Alternative A pays either zero or a payoff which depends on the number of entrants according to Table 1.1. The number of entrants, however, is not decided in the game but is instead taken from the data of the WTA condition from the main experiment. More precisely, we allocate each subject to a particular group of 14 subjects from a session with the WTA condition. In each period in which the subject chooses Alternative A, we replace one of the entrants in the original data by the subject in question to calculate the payoff in case of winning.[8] The feedback about the random numbers and the entry decisions is identical to the WTA condition; that is, the AMB condition offers the same learning stimuli as the WTA condition. In addition, we also elicited beliefs. Because the subject in question is not an entrant, we elicit beliefs about the behavior of the group of 14 subjects from the WTA condition.

In the OBJ condition we go one step further towards individual choice and remove the belief element. Subjects again have to decide between Alternative A and Alternative B in 20 rounds, and the payoff for Alternative B remains 45 ECUs. The payoffs for Alternative A are still determined by entry behavior in the WTA condition, but we communicate the probability distribution of the entry behavior to the subjects. More precisely, in the OBJ condition, we inform subjects about the absolute and relative entry frequencies in the WTA condition. The information shown in Table 1.6 is presented to subjects in the experimental instructions. It shows the absolute and relative frequencies of every possible number of market entrants observed in the WTA condition. The first row of the table shows the number of market entrants; the second row replicates the market payoff for each number of market entrants from Table 1.1. The interpretation is unchanged: if, for instance, five players enter the market, the market payoff is 190. The third row shows the number of times the respective number of market entrants was observed in the WTA condition. For instance, in 35 rounds the number of market entrants was five.[9]

The forth row shows the relative frequency with which the respective number of market entrants was observed in the WTA condition. For instance, in 19.4% of the rounds (=35/180) the number of market entrants was five. In the instructions the table was accompanied by a histogram showing the relative frequencies as a visual aid.

[8]In one group from the WTA condition data there was one period with no entry at all. We treated this case similar to the case with one entrant, i.e., the subjects who were allocated to this situation and chose Alternative A received 100 ECUs. This holds also for the OBJ condition.

[9]The sum across all cells of the third row is 180 group/period outcomes. These stem from all sessions of the main WTA experiment except the one session where we used a stranger matching. We dropped the session with stranger matching because the instructions in the AMB and OBJ conditions explain that the data stems from experiments using a partner matching. See the appendix (pp. 43–49) for the experimental instructions.

Table 1.6: Information about the entry frequencies in the OBJ condition.

Number of participants who chose A	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Prize to win in ECU	100	100	130	155	175	190	200	205	209	212	214	215	215	215	215
Absolute number of rounds	1	0	1	8	20	35	51	30	15	13	2	4	0	0	0
Relative frequency in %	0.6	0	0.6	4.4	11.1	19.4	28.3	16.7	8.3	7.2	1.1	2.2	0	0	0

Notes. Information about the winning probabilities offered to the subjects in the OBJ condition. The table shows absolute and relative frequency of the number of entrants based on the entry decisions in the WTA condition in the main experiment. Alternative "A" stands for market entry.

Thus, not only are payoffs for Alternative A independent of the decisions of the other subjects in the session, but, in the OBJ condition, subjects know the probability of each possible number of entrants and the probability of winning the market payoff given any number of entrants. This transforms the decision problem into a risky choice with complete information about all contingencies.

We conducted two sessions with a total of 53 subjects, none of which participated in the main experiment. All subjects had to read the original instructions of the WTA condition in order to have the same information about the game as the subjects from the main experiment. In addition, subjects received instructions detailing that they would not be in a strategic situation with the other subjects currently present in the laboratory. All subjects played 20 periods of the AMB condition followed by 20 periods of the OBJ condition. The average payoff per subject in the additional experimental sessions was CHF 31 ($\approx$ USD 31).

The left panel of Figure 1.4 shows the frequency of Alternative A decisions in the AMB condition compared with the entry frequency observed in phase 1 of the WTA condition. The frequencies are very similar, with 0.458 in the WTA condition and 0.475 in the AMB condition ($p = 0.741$). The right panel of Figure 1.4 shows the frequency of Alternative A decisions in the OBJ condition compared with the number of entrants in Phase 2 of the WTA condition. The entry frequency in the objective probability condition is 0.425, which is even somewhat higher than the 0.405 of the WTA condition ($p = 0.741$).

To sum up, we find no evidence that the risky alternative becomes less attractive if we eliminate the competitive element, and we conclude that "joy of winning" is unlikely to be a driver of excess entry.

1.5 Explaining excess entry with Cumulative Prospect Theory

After our unsuccessful empirical quest for the causes of excess entry, we redirected our efforts to look for theoretical arguments that might help us understand why—contrary to our intuition—the winner-take-all feature renders market entry *more* attractive compared to the market entry game. This time we were successful: it turns out that Cumulative Prospect Theory (CPT; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) captures the differences between the MEG and WTA conditions surprisingly well. As a first step, we derive the Nash equilibria for CPT players. Next, we run new experimental sessions in which we elicit subjects' certainty equivalents for 40 two-outcome lotteries after they completed the MEG and WTA phases. This additional task allows us to estimate pooled CPT parameters. Finally, we compare actual entry decisions in both the MEG

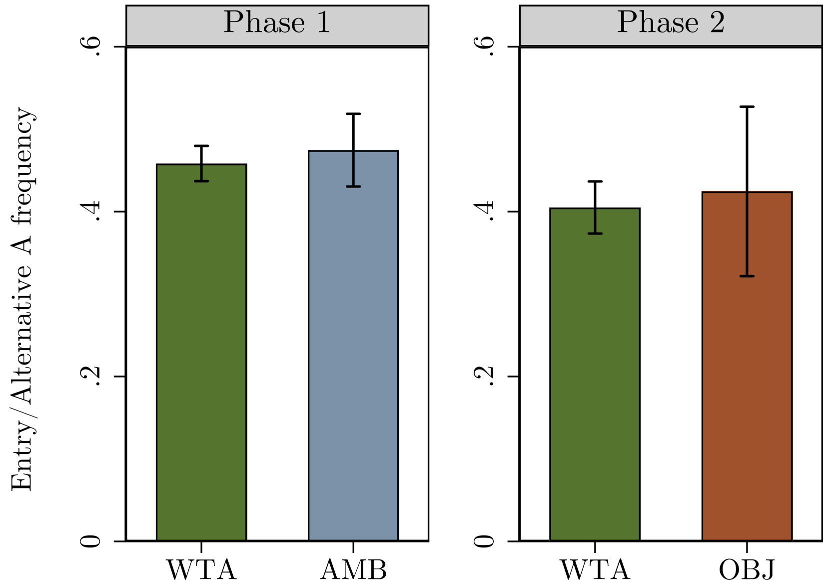


Figure 1.4: Left panel: Average entry frequency in phase 1 of the WTA condition and frequency of Alternative A decisions in the ambiguous probabilities condition (AMB). Right panel: Entry frequency in phase 2 of the WTA condition and frequency of Alternative A decisions in the objective probabilities condition (OBJ). Spikes show the 95% confidence intervals. In the WTA condition standard errors are clustered on group.

and the WTA conditions to theoretical predictions using our CPT parameter estimates. We show that our parameter estimates for the curvature of the utility function and the function to transform objective cumulative probabilities into subjective cumulative probabilities (henceforth "weighting function") can predict excess entry in the WTA condition. The decisive component is probability weighting.

The adequacy of linear probability weights has already been questioned by Allais (1953). Since then, a variety of non-linear functional forms for probability weights have been suggested in the literature (Goldstein & Einhorn, 1987; Lattimore et al., 1992; Prelec, 1998; Quiggin, 1982; Tversky & Kahneman, 1992). We use probability weighting in the framework of CPT to make Nash equilibrium predictions in the MEG and WTA conditions.[10] To compute Nash equilibrium predictions under CPT we rewrite Equation (1.1) as

$$\sum_{k=1}^{m} \omega_k u(x_k) = u(I). \tag{1.3}$$

The left-hand side describes the subjective value of market entry, where ω_k is the decision weight on outcome x_k, $u(x_k)$ is the utility associated with outcome x_k, and m is the number of possible outcomes for an entrant. In a mixed-strategy Nash equilibrium the expression on the left-hand side must be equal to the utility provided by staying out, $u(I)$. We use a two-parameter weighting function proposed by Goldstein and Einhorn (1987)

[10]For analysis of Nash equilibrium properties with (Cumulative) Prospect Theory decision making, see Goeree et al. (2003), Keskin (2016), or Metzger and Rieger (2019).

and Lattimore et al. (1992):[11]

$$h(Q, \delta, \gamma) = \frac{\delta Q^{\gamma}}{\delta Q^{\gamma} + (1-Q)^{\gamma}}, \tag{1.4}$$

where Q is the objective cumulative probability and the parameter γ governs likelihood sensitivity. If $\gamma \in (0, 1)$, the function displays an inverted S-shaped pattern where small cumulative probabilities are upweighted and large cumulative probabilities are downweighted (likelihood insensitivity). If $\gamma > 1$, then the function displays an S-shaped pattern where low cumulative probabilities are downweighted and high cumulative probabilities are upweighted (likelihood oversensitivity). Upweighting (downweighting) of small cumulative probabilities leads to overweighting (underweighting) of small probability extreme outcomes. The parameter δ governs the elevation of the weighting function (optimism/pessimism). If $\delta \in (0, 1)$, then the individual is pessimistic as she downweights the probability of high payoff outcomes and upweights the probability of low payoff outcomes. Conversely, if $\delta > 1$, then the individual is optimistic as she upweights the probability of high payoff outcomes and downweights the probability of low payoff outcomes. When $\gamma = \delta = 1$ the weighting function is linear and the model boils down to expected utility maximization with $h(Q, 1, 1) = Q$.

The decision weights in Equation (1.3) are weighted cumulative probabilities on outcomes that are ordered by their payoff. If $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_m$ holds for the set of possible realizations of x, then the subjective decision weights are:

$$\begin{aligned}
\omega_1 &= h(q_1) - h(0) \\
\omega_2 &= h(q_1 + q_2) - h(q_1) \\
&\vdots \\
\omega_k &= h(Q_k) - h(Q_{k-1}) \\
&\vdots \\
\omega_m &= h(1) - h(Q_{m-1}),
\end{aligned}$$

where q_k denotes the probability that x is equal to x_k and $Q_k = \sum_{j=1}^{k} q_j$ denotes the cumulative probability, that is the probability that x is greater than or equal to x_k; $h(Q_k)$ is the weighting function as defined in Equation (1.4) applied to the cumulative probability Q_k. By setting $h(1)$ equal to one and $h(0)$ equal to zero we have the subjective decision weights sum up to one. To model utility curvature we use a constant relative risk aversion (CRRA) utility function:

$$u(x, \eta) = \begin{cases} \frac{x^{(1-\eta)}-1}{1-\eta}, & \text{for } \eta \neq 1 \\ \ln(x), & \text{for } \eta = 1 \end{cases} \tag{1.5}$$

where $x \geqslant 0$ is the payoff and η measures CRRA. The parameter η determines the curvature of the utility function: $\eta > 0$ corresponds to concave utility functions, $\eta = 0$ to linear utility functions, and $\eta < 0$ to convex utility functions.[12] Because negative payoffs

[11]The literature offers a wide range of specifications for the weighting function. We follow Gonzalez and Wu (1999), who argue that the evidence in the domain of gains favors a weighting function with two parameters.

[12]In standard expected utility models, risk preferences are directly related to the curvature of the utility function. In a CPT framework risk averse behavior can be attributed to the utility function and/or the probability weighting function. In what follows we will avoid the term risk aversion. Instead, we directly refer to the curvature of the utility function or to probability weighting. For a discussion on risk aversion in a CPT framework see e.g. Schmidt and Zank (2008).

are not possible in our games we only formulate the utility in the domain of gains.[13]

What does this imply for Nash equilibrium predictions in the MEG and WTA conditions? As market entry in the two conditions is subject to different levels of risk, the two conditions cease to be equivalent in terms of Nash equilibrium predictions. In the following we derive the symmetric mixed-strategy equilibria for the MEG and WTA conditions separately.

1.5.1 MEG condition

In the MEG condition, the payoffs for market entry are determined by $\pi(E)$ according to Table 1.1. To compute the value of market entry, we order all possible payoffs associated with market entry from the highest to the lowest:

$$\begin{aligned} X &= \{x_1, x_2, \ldots, x_{14}\} = \{\pi(1), \pi(2), \ldots, \pi(14)\} \\ &= \{100, 65, 51.7, 43.8, 38, 33.3, 29.3, 26.1, 23.6, 21.4, 19.5, 17.9, 16.5, 15.4\}. \end{aligned}$$

Here, x_k is the payoff from entering the market if $k-1$ other players enter the market and $13-(k-1)$ other players stay out. Accordingly, q_k denotes the probability of winning the payoff associated with $k-1$ other players entering and $13-(k-1)$ other players staying out. Given the number of entrants, the payoff is non-random in the MEG condition, i.e., there is only strategic risk.

The probabilities $\{q_1, q_2, ..., q_{14}\}$ depend on the entry probability p of the other players. The probability that $k-1$ other players enter the market is:

$$q_k = \binom{13}{k-1} p^{k-1}(1-p)^{13-(k-1)}, \qquad k = 1, \ldots, 14. \tag{1.6}$$

To obtain the Nash equilibrium predictions, we compute the decision weights ω_k associated with the probabilities q_k. The symmetric Nash equilibrium in mixed strategies is then the entry probability p^* that solves Equation (1.3). The predicted number of market entrants is obtained by multiplying the probability p^* with the number of players.

The first part of Table 1.7 shows the expected number of entrants for selected values of the three preference parameters. At the top left, we start with the benchmark case of linear utility and linear probability weights, which results in 3.52 expected entrants (see Section 1.2.2). Increasing concavity of the utility function (η) has hardly any effect on entry, whereas increasing likelihood insensitivity (lowering γ) increases entry towards four entrants. Finally, lowering optimism (δ) reduces entry somewhat.

1.5.2 WTA condition

In the WTA condition, the payoff of market entry is subject to two kinds of risk: (i) strategic risk, originating from the behavior of other players, and (ii) natural risk, originating from the random draw of the winner among the entrants. In our main model, we will assume that players do not distinguish between natural and strategic risk. Consequently,

[13] In the instructions, we frame the decision situation clearly as choice between a secure and a risky gain. Nevertheless, one could argue that the payoff from staying out (45) serves as a reference point, which would introduce the possibility for losses in both games. In Appendix 1.A.1.1 (p. 27), we discuss the equilibria for a reference point of 45 and loss aversion. It turns out that loss aversion reduces the expected number of entrants very similarly to increasing concavity of the utility function. In the following we will restrict our attention to the gains only version in the main text.

Table 1.7: Predicted number of entrants in MEG and WTA

		MEG			WTA		
		$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$
$\gamma = 1$	$\delta = 1$	3.52	3.49	3.47	3.52	2.79	2.22
	$\delta = 0.8$	3.31	3.29	3.26	2.59	2.09	1.69
$\gamma = 0.75$	$\delta = 1$	3.74	3.70	3.66	6.14	4.38	3.05
	$\delta = 0.8$	3.47	3.43	3.39	4.08	2.84	1.99
$\gamma = 0.5$	$\delta = 1$	4.18	4.10	4.03	14.00	10.44	6.64
	$\delta = 0.8$	3.78	3.71	3.64	10.01	6.39	3.55

Notes. Expected number of entrants in the symmetric mixed strategy equilibrium with CPT preferences, for selected values of the three preference parameters η, γ, and δ.

we will treat the two sources similarly and apply probability weighting to the compound objective probabilities.[14] Some support for combining the two sources of risk comes from our results from Section 1.4.3, where we report similar entry behavior in the OBJ condition compared to the WTA condition. We start by ordering the possible payoffs from the highest to the lowest to compute the value of market entry. The set of possible payoffs is:

$$X = \{x_1, x_2, \ldots, x_{15}\} = \{\Pi(14), \Pi(13), \ldots, \Pi(1), 0\}$$
$$= \{215, 215, 215, 215, 214, 212, 209, 205, 200, 190, 175, 155, 130, 100, 0\}.$$

Thereby, 215 is the winner's payoff for entering the market if 10, 11, 12, or 13 other players enter; 214 is the winner's payoff for entering the market if 9 other players enter. Zero is the payoff for entering the market with any number of other players if the player does not win the market payoff. The probabilities $\{q_1, q_2, ..., q_{15}\}$ associated with the set of payoffs are:

$$q_k = \begin{cases} \binom{13}{13-(k-1)} p^{13-(k-1)}(1-p)^{k-1}\frac{1}{14-(k-1)}, & \text{for } k = 1, \ldots, 14 \\ 1 - \sum_{j=1}^{14} q_j, & \text{for } k = 15. \end{cases} \tag{1.7}$$

Again, the symmetric Nash equilibrium in mixed strategies is the entry probability p^* that solves Equation (1.3). The right part of Table 1.7 shows the expected number of entrants in the WTA condition. Starting with the 3.52 entrants in the benchmark case ($\eta = 0, \gamma = 1, \delta = 1$), we observe that the qualitative effects of the parameters are the same as in the MEG condition. However, the magnitude of the changes is much larger. For example, concavity of the utility function ($\eta > 0$) and pessimism ($\delta < 1$) reduce the number of entrants substantially. At the same time, likelihood insensitivity ($\gamma < 1$) results in substantial excess entry. Table 1.7 gives us a feel of the parameters we need to explain higher number of entrants in the WTA condition relative to the MEG condition. Without probability weighting ($\gamma = 1, \delta = 1$) the model reduces to standard expected

[14] Alternatively, one could assume that probability weighting applies to natural risk only. Our theoretical results would not differ much, because most of the predicted treatment effect originates from the natural risk part. Finally, probability weighting could only affect strategic risk. In this case (assuming linear utility) the model would not predict any difference between the WTA and the MEG conditions. Appendix 1.A.1.2 (p. 28) provides a formal discussion on different variants of probability weighting.

utility and does not permit to explain the treatment effect (unless, of course, we were to assume convex utility, $\eta < 0$). Comparing the expected number of entrants across conditions shows that the model can explain the treatment effect if γ is sufficiently low, η is moderate, and δ is not too low. If we rule out (as most of the literature does) convex utility, then our theoretical results show that the explanation for excess entry within CPT must stem from probability weighting. Of the two parameters of the weighting function, the main driver is γ. If $\gamma < 1$, then small probability events with very high and very low payoffs gain weight. In the WTA condition, these are the high payoffs if one wins against many competitors. The worst outcome in the WTA condition (losing against any number of competitors) is underweighted, because it occurs with high probability. In the MEG condition, overweighting occurs for both the best and the worst outcomes of market entry, as both are small probability events.[15]

Thus, CPT can be a powerful explanation for excess entry into winner-take-all markets. Its predictive power, however, depends on the preference parameters. In order to explore the predictive power of CPT, we ran new experiments. The experiments replicate the design of WTA and MEG conditions, followed by an elicitation of a series of certainty equivalents, which permits us to estimate CPT parameters. We compare our estimated parameters to previous estimates from the literature and use our estimates to evaluate the predictive power of CPT.

1.5.3 Experimental design

In a new experiment, we elicit subjects' certainty equivalents for two-outcome lotteries (CPT task) after they completed the MEG and WTA phases. We will label these new observations as MEG_{CPT} and WTA_{CPT}, respectively. Both the MEG_{CPT} and WTA_{CPT} conditions are identical to the main experiment. Again, we run the MEG condition first in half of the sessions, reversing the order for the other half. We elicit the subjects' certainty equivalents for 40 lotteries. Each lottery offers an amount x_1 with probability p and an amount $x_2 < x_1$ with probability $1 - p$. Half of the lotteries were in the gain domain ($x_1 > x_2 \geq 0$), and the other half were corresponding lotteries in the loss domain ($0 \geq x_1 > x_2$). For each lottery in the loss domain, we endowed subjects with $|x_1| + |x_2|$. This ensures that subjects do not end up with a negative payoff, and it equalizes the expected payoff of a lottery in the loss domain to its counterpart in the gain domain. At each screen, subjects are presented with a lottery on the left and with 20 rows of equally spaced guaranteed outcomes ranging from x_1 to x_2 on the right. For each row, we ask the subject if she prefers the lottery or the corresponding guaranteed outcome. We enforce monotonicity; that is, subjects only choose the minimum certainty equivalent they prefer to a lottery and the computer fills in the remaining rows accordingly. A subject's elicited certainty equivalent for a lottery is the arithmetic mean of the lowest guaranteed amount preferred to the lottery and the highest guaranteed amount not preferred to the lottery.[16] Once a subject took the decisions on the 40 lotteries (they are allowed to go back and forth), a lottery and a row were randomly chosen by the computer and the subject's payoff is determined based on her decision for the randomly chosen lottery and row.

[15] Alternatively, it would be possible to predict the treatment effect without inverted S-shaped probability weighting. In a model with $\gamma = 1$ combined with optimism ($\delta > 1$) players upweight the outcomes with high payoffs and downweight the outcomes with low payoffs. In the main text we emphasize the model with $\gamma < 1$ and $\delta < 1$, because there is substantial empirical support for these parameter ranges (see discussion in Section 1.5.4.1).

[16] See Table 1.A.3 in the appendix (p. 30) for the set of lotteries and pp. 50–54 for screenshots.

In addition to the show-up fee of CHF 5, subjects were paid for both phases, as in the main experiment. Furthermore, subjects received the payment for the CPT task. Subjects received information about their payoff only after the CPT task. The exchange rate for this task was CHF 1 per 3 ECUs.

The sessions were run online with undergraduate students from the University of Lausanne and the EPFL (recruitment: ORSEE, Greiner, 2015; programming: oTree, Chen et al., 2016). We ran ten sessions with a total of 134 subjects.[17] This gives us 5,360 entry decisions and 5,360 certainty equivalent elicitations. Sessions lasted between 47 and 100 minutes,[18] with an average payoff per subject of CHF 43.

1.5.4 Results

Our new sessions replicate the findings on market entry behavior of the main experiment: in the first (second) phase, we observe on average 6.49 (4.94) entrants in the WTA_{CPT} condition and 4.86 (4.18) entrants in the MEG_{CPT} condition. The treatment effect is significant ($p = 0.022$, exact Wilcoxon signed-rank test).

We use the CPT task to estimate the preference parameters. Following the estimation procedure for the single group case by Bruhin et al. (2010a), we estimate the CPT parameters based on the lottery decisions in the gain domain. This procedure uses maximum likelihood to estimate the parameters based on the elicited certainty equivalents. We account for two sources of heteroskedasticity in the error variances. First, the error is proportional to the lottery's range $|x_1 - x_2|$, since subjects faced 20 equally spaced guaranteed outcomes for each lottery. Second, we account for heteroskedasticity between individuals, as subjects may differ with respect to previous knowledge, attention span, or ability.

We use the data elicited from the CPT task to perform parameter estimations. In a first step, we perform pooled estimations of CPT parameters. Assuming homogeneity across subjects, we confront market entry predictions using the parameters from the pooled estimations with the data on entry behavior in the MEG_{CPT} and the WTA_{CPT} conditions. In the second step, we relax the homogeneity assumption. We perform finite mixture model estimations to group subjects into types and test whether entry behavior is related to the type.

1.5.4.1 Pooled CPT estimations

Using the full sample, we find the following point estimates (standard errors) for the CPT parameters: $\hat{\eta} = 0.131$ (0.031); $\hat{\gamma} = 0.517$ (0.023); $\hat{\delta} = 0.929$ (0.049). Our results are in line with previous estimates from the literature that use the same two-parameter specification for the weighting function (Bruhin et al., 2010a; Bruhin et al., 2018; Fehr-Duda et al., 2006; Fehr-Duda et al., 2011). Overall the estimates in these studies point to a significantly inverted S-shaped weighting function and moderate pessimism. In terms of

[17] Due to no-shows and connection problems we did not always manage to get the desired group size. We had one session each with 11, 12, and 13, while the remaining seven sessions had 14 subjects. Such variations in group size are not a problem for our analysis because the equilibrium predictions of the expected number of entrants barely changes for group sizes between 11 and 14.

[18] The sessions' durations varied substantively because subjects could go at their own pace in the CPT task. Once a subject was done with this task, she would directly move to the final payment page. The payment was done by bank transfer within a week.

the utility function, most estimates indicate a curvature close to linearity, with moderate deviations in both directions.

Given our parameter estimates, CPT predicts 3.92 entrants in the MEG condition and 7.28 in the WTA condition in the symmetric mixed-strategy Nash equilibrium. We can use our estimates to illustrate why probability weighting produces such strong excess entry in the WTA condition, even with a concave utility function. In the equilibrium of the WTA game an entrant's probability of getting the highest prize (215, winning among eleven or more entrants) is 0.0055. The weight of these outcomes in the decision process is ten times higher (0.059). Conversely, the probability of getting zero is 0.863, while the decision weight attached to this outcome is downweighted to 0.736. In the MEG condition it is also the case that the best outcome (100, being the sole entrant) is strongly overweighted. However, in contrast to the WTA condition, there is also substantial overweighting of the unprofitable outcomes of market entry (entry with six or more competitors). See Figure 1.A.1 in the appendix (p. 29) for an illustration.

To gain an idea on the error margin of the predicted number of entries, we use a bootstrap approach. Based on 1000 bootstrap samples, we estimate a distribution of CPT parameters, which we use to predict the corresponding entry probabilities. Using these probabilities allows us to derive a distribution of the expected number of entrants in the MEG and WTA conditions. The left panel of Figure 1.5 shows the results. The horizontal lines indicate the mean number of predicted entrants in the MEG_{CPT} and the WTA_{CPT} conditions, with the gray bands indicating the 95% confidence intervals. In the WTA_{CPT} condition, we observe a number of entrants below the point prediction, but still within the confidence interval. The low number of entrants in the WTA_{CPT} condition relative to the Nash equilibrium prediction is mainly due to the low entry frequency when played in the second phase. If we restrict the sample to the first phase, then the observed number of entrants in the WTA_{CPT} condition is substantially higher (see Figure 1.A.4 in the appendix, p. 33). In the MEG_{CPT}, on the other hand, the observed number of entrants is above the Nash equilibrium prediction.

1.5.4.2 Finite mixture model CPT estimations

The empirical literature on CPT preferences often finds evidence for substantial heterogeneity between subjects. However, CPT elicitation procedures like the one we use are too imprecise to provide reliable estimates on an individual level (Monroe, 2020).[19] To account for heterogeneity, we therefore identify types using a finite mixture model (FMM). Again, we follow Bruhin et al. (2010a) estimation procedure for the two-group case. As the authors explain, the idea is to assign an individual's risk taking choices to either the EUT group or the CPT group, each of the two groups being characterized by a distinct vector of parameters. The assignment of each subject to one of the two groups is based on the posterior probability of type-membership.[20]

[19]Table 1.A.4 in the appendix (p. 31) provides summary statistics of the individual CPT parameter estimates and Table 1.A.5 (p. 32) shows OLS estimates of entry frequency in WTA on the individual preference parameter estimates. All coefficient estimates have the expected sign: increasing utility curvature (η) or likelihood sensitivity (γ) leads to less entry in WTA and increasing optimism (δ) leads to more entry in WTA. However, none of the coefficients reach statistical significance.

[20]We refer to Bruhin et al. (2010a, 2010b) for a detailed explanation of the estimation procedure. Figure 1.A.5 in the appendix (p. 33) shows an histogram of individuals' posterior probability of assignment and Table 1.A.6 (p. 32) shows the FMM estimates. We elicited certainty equivalents in both the gain and loss domains to keep the experimental design as close as possible to Bruhin et al. (2010a). For the param-

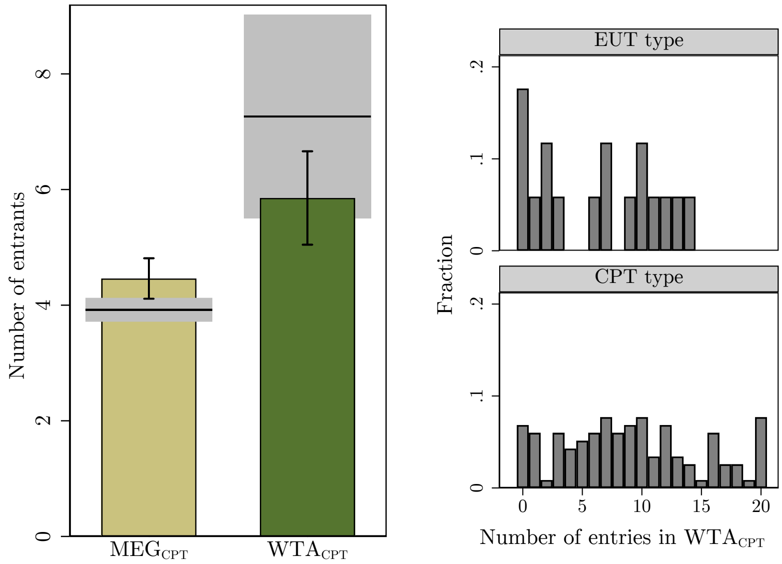


Figure 1.5: Left panel: Average number of entrants in the two conditions (bars with spikes indicating 95% confidence intervals); predicted number of entrants (horizontal bar) and corresponding 95% confidence interval (gray band). Right panel: Histogram of the number of entries during the 20 periods for subjects classified as expected utility maximizer (EUT) and Cumulative Prospect Theory (CPT) types.

According to our FMM estimates, 117 out of the 134 subjects are classified as CPT types and the remaining 17 as EUT types. With 13% of the subjects classified as EUT types, the estimated share of EUT subjects in our subject pool is somewhat lower than in Bruhin et al. (2010a) where the estimated share in the Zurich 2006 experiment is 22%.

The right panel of Figure 1.5 shows histograms of the number of times subjects of the EUT type (upper part) and the CPT type (lower part) enter in the WTA_{CPT} condition. The CPT types are less likely to enter very rarely (two or fewer times out of 20) and more likely to enter very often (15 times or more). In order to test whether the CPT types differ significantly from the EUT types, we regress the entry frequency (in percent) across the 20 periods of the WTA_{CPT} on the type dummy. Table 1.8 shows the two main results. First, even for EUT types we observe entry frequencies above Nash equilibrium predictions under the expected payoff maximizer assumption. Recall that while the predicted entry frequency under the expected payoff maximizer assumption is 25 percent, players of the EUT type who anticipate that the players of the CPT type enter more frequently should respond by refraining from entering in the winner-take-all game. Our results do not confirm this: The estimated entry frequency for EUT types (31.4 percent in Model 1) is even above the 25 percent. Second, the frequency at which CPT types enter is about 15 percentage points higher than for EUT types in the WTA_{CPT} condition. The effect is highly significant. In Model (2) we control for order effects, confirming the result that entry is less frequent when the WTA_{CPT} condition is played after the MEG_{CPT} condition. The dummy for CPT types remains highly significant. These results suggest that, in line with theory, excess entry is mainly driven by CPT types.

eter estimations, we only use the lotteries in the gain domain since we are interested in decision situations between a secure and a risky gain (see Footnote 13).

Table 1.8: Entry frequency and type.

	Dependent variable: Entry frequency (in %)	
CPT type	14.170^{**}	15.087^{**}
	(4.314)	(4.591)
WTA in phase 2		-13.913^{*}
		(4.558)
Constant	31.471^{**}	36.381^{**}
	(4.926)	(3.590)
F-test	10.8	6.4
Prob $> F$	0.009	0.019
R^2	0.027	0.083
N	134	134

Notes: OLS estimates. Dependent variable is the individual entry frequency in the WTA condition (in %). Independent variables are a dummy for the subject's type according to the finite mixture model, and a dummy for the WTA played after the MEG condition. Baseline case is the EUT type. Robust standard errors, clustered on group, in parentheses. $^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$.

1.6 Conclusion

In this study, we systematically explore potential causes of excess entry into markets. Early experimental research showed that this phenomenon does not occur in standard market entry games. More recently, researchers have shown that overconfidence can create excess entry in games where the returns on market entry depend on skills. We provide experimental evidence that a market modeled as a winner-take-all game with a purely random determination of the winner creates strong excess entry relative to an expected-payoff-equivalent market entry game. This means that excess entry can occur even in environments where overconfidence in skills is irrelevant. We provide evidence that the explanation for excess entry is probability weighting. In environments where overconfidence about skill plays a role, the two biases—overconfidence and probability weighting—presumably reinforce each other.

It is clear that utility curvature, which underlies standard decision-making models, is unlikely to explain excess entry. If subjects display concave utility functions, then we would expect less rather than more entry in the winner-take-all condition compared to the market entry game. With additional data, we explore and discard a number of explanations discussed in the previous literature. First, we show that wrong beliefs about the number of other players entering the market (blind spot) cannot account for the treatment effect. Second, we find no evidence that the treatment effect is driven by subjects who fall victim to the erroneous belief that they can influence random processes (illusion of control). Third, we find no evidence for explanations positing extra utility of winning in a competitive environment (joy of winning). To address this explanation, we ran an additional nonstrategic condition which preserves expected payoffs and risk of the winner-take-all condition, but does not contain the competitive element. We find that the number of market entrants remains virtually unchanged. In an additional treatment, we even provide objective probabilities associated to market entry and still find similar levels of excess entry. This latter result further strengthens the argument that it is not the additional complexity that comes with a strategic interaction that drives excess entry.

We identify probability weighting in accordance with Cumulative Prospect Theory as a powerful explanation for the treatment effect. We compute Nash equilibria of the standard market entry game and the winner-take-all condition. Under a broad range of realistic parameters, the model predicts more entry in the winner-take-all condition than in the market entry game. The effect is driven by the fact that the winner-take-all competition offers high stakes with low probabilities, while losing the competition is very likely. Typical parameters for the probability weighting function lead to overweighting of the profitable outcomes and underweighting of the bad outcome.

We ran an additional experiment, which allows us to estimate CPT parameters and relate these estimates to entry behavior. The calibrated predictions are consistent with observed entry behavior. In addition, the results of our finite mixture estimates suggests that excess entry is mainly caused by subjects that can be classified as CPT type. Taken together, these results suggest that CPT is a useful descriptive theory not only for individual choice problems, but also in situations involving strategic risk. In addition, the fact that we observe very similar entry frequencies in our non-strategic conditions provides indicative evidence that CPT probability weighting may be similarly applied to both natural risk and strategic risk. In other words, moving from subjective belief-based risk to objective probabilities (as in our OBJ condition) seems not to matter much for entry behavior and may be explained with the same underlying probability distortion.

Should regulators or nudgers care about excess entry in winner-take-all games? Our results provide a suggestive answer to that question. If excess entry was caused by non-standard preferences, such as high-risk tolerance or strong non-monetary benefits from winning the competition against others (joy of winning), then there would be little ground for paternalistic interventions. Our results suggest that excess entry is not caused by such preferences but is instead caused by probability weighting. It is conceivable that subjects in our lab would not consciously choose to have decision weights that differ from objective probabilities and, if made somehow aware of the discrepancy, would prefer to use objective probabilities in their decision process. Thus, upcoming entrepreneurs or hopeful young athletes might enter winner-take-all markets not only because of overconfidence about their abilities but also due to a systematic overestimation of the small winning probabilities. Consequently, a nudge to "objectify" probabilities might benefit would-be competitors. We leave for future research to investigate the design and effectiveness of such nudges. A much simpler policy conclusion is drawn by Frank and Cook (1995) who argue that progressive consumption taxation might be an effective measure to reduce the attractiveness of entry into winner-take-all markets.

For business executives who take investment decisions, our findings can provide guidance. To gauge return prospects of an investment, anticipating the behavior of potential competitors is crucial. Biased probability perception can account for skew-seeking behavior and thus attraction to winner-take-all markets through optimism and likelihood-insensitivity. Åstebro et al. (2015) have documented optimism and likelihood insensitivity both among college students and business executives. Based on our findings, markets with winner-take-all characteristics are likely to attract more competitors than the market volume warrants. As formalized in the Porter's Five Forces Framework (Porter, 1989), the threat of entry is an important determinant of the attractiveness of an industry in terms of its profitability. Our results suggest that an investment in a market where few competitors capture a large share of the rewards is less attractive compared to markets with a more equal share of rewards but the same expected return.

Finally, our results provide us with a testable conjecture for further research on market entry decisions. We compare two extreme payoff distributions—winner-take-all versus equal payoffs for all entrants—and show that, counterintuitively, the riskier market attracts more entry. Consequently, in real market entry data one should find more excess entry in markets where the profits are skewed towards the top performers. In particular, excessive entry of new firms and high failure rates should be less prevalent in sectors where payoffs are relatively equal, arguably, for example, in hospitality, and more prevalent in sectors like information technology.

1.A Appendix

1.A.1 Cumulative Prospect Theory: Extensions

1.A.1.1 Loss aversion

Our main treatment of the WTA and MEG conditions with Cumulative Prospect Theory (CPT) assumes that players' reference points are zero and all payoffs in the game are perceived as gains. Alternatively, we could assume that players have some reference point r, relative to which they evaluate the monetary payoff x using the utility function:

$$u(x,\eta) = \begin{cases} \frac{1}{1-\eta}\left((x-r)^{(1-\eta)} - 1\right), & \text{for } x \geqslant r \\ \frac{1}{1-\eta}\left(-\lambda(r-x)^{(1-\eta)} - 1\right). & \text{else} \end{cases} \tag{1.8}$$

For $0 < \eta < 1$ this function is concave in the domain of gains and convex in the domain of losses. We assume the CRRA parameter (η) is the same in both domains, and we introduce a parameter for loss aversion (λ). Allowing for a different CRRA parameter in the domain of losses has very similar effects as variations in λ. In contrast to the solutions discussed in the main text, we now assume that the reference point is $r = 45$. Table 1.A.1 shows the predicted number of entrants for selected combinations of η, γ, and λ (with δ=1). The results for linear utility and $\lambda = 1$ are identical to the ones shown in Table 1.7. Increasing the CRRA parameter (η) while keeping $\lambda = 1$ reduces predicted entry in the WTA condition, but not as much as in Table 1.7. This is due to the fact that the utility function is convex in the domain of losses. Similarly to increasing the CRRA parameter, loss aversion substantially reduces the attractiveness of the WTA market, providing further support for the argument that probability weighting is the driver for the treatment effect.

Table 1.A.1: Predicted number of entrants, reference point at 45.

		MEG			WTA		
		$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$
$\gamma = 1$	$\lambda = 1$	3.52	3.47	3.42	3.52	3.12	2.81
	$\lambda = 1.2$	3.38	3.33	3.27	2.76	2.49	2.26
	$\lambda = 1.4$	3.27	3.21	3.14	2.27	2.07	1.90
$\gamma = 0.75$	$\lambda = 1$	3.74	3.67	3.60	6.14	5.02	4.14
	$\lambda = 1.2$	3.56	3.49	3.41	4.43	3.61	3.02
	$\lambda = 1.4$	3.42	3.34	3.25	3.28	2.72	2.31
$\gamma = 0.5$	$\lambda = 1$	4.18	4.08	3.97	14.00	11.93	9.22
	$\lambda = 1.2$	3.92	3.81	3.69	10.77	8.20	6.10
	$\lambda = 1.4$	3.71	3.59	3.46	7.77	5.67	4.07

Notes. Expected number of entrants in the symmetric mixed strategy equilibrium with CPT preferences with a reference point at 45, and for three levels of loss aversion (λ). We plot the entries for $\delta = 1$ and selected levels of η and γ.

1.A.1.2 Natural vs. strategic risk

In the main text, we derive the CPT equilibria under the assumption that probability weighting applies to the combination of natural and strategic risk. Applying probability weighting only to one of the two sources of risk would change the prediction. As the empirical literature on CPT almost exclusively deals with natural risk, it seems natural to apply probability weighting only to this source of uncertainty. In the following, we discuss two methods:

WTA, N1: Weighting natural risk conditional on number of entrants. Here we simply assume that the player separates the strategic risk (ending up with $E-1$ competitors) from the natural risk conditional on the number of competitors ($q_E = 1/E$). For $E = \{2, 3, \dots, 14\}$ we use Equation (1.4) to transform the objective winning probabilities q_E into weights. The low-probability high payoffs in the WTA markets have higher weights than the objective probabilities, which produces higher number of expected entrants relative to the case with linear probabilities. The left part of Table 1.A.2 shows the results. For comparison we also show the results of the MEG condition, which are not affected by probability weighting, because there is only strategic risk. Comparing the results to our model in the main text (Table 1.7) shows that the predicted number of entrants is very similar, suggesting that the inclusion of strategic risk is not a decisive factor when we explain the treatment effect.

WTA, N2: Decomposing natural and strategic risk. From Equation (1.7), we can write the cumulative probabilities of achieving an outcome at least equal to x_k:

$$Q_k = \sum_{j=1}^{k} \binom{13}{13-(j-1)} p^{13-(j-1)}(1-p)^{j-1} \frac{1}{14-(j-1)}, \quad \text{for } k = 1, \dots, 14$$

and $Q_{15} = 1$ where $x_{15} = 0$ is the outcome when a player enters and loses.

If we multiply and divide Q_k by the objective probabilities of achieving an outcome at least equal to x_k, we can decompose natural and strategic risks in the following way:

$$Q_k = \left(\sum_{j=1}^{k} \binom{13}{13-(j-1)} p^{13-(j-1)}(1-p)^{j-1} \right) \left(\frac{\sum_{j=1}^{k} \binom{13}{13-(j-1)} \frac{p^{13-(j-1)}(1-p)^{j-1}}{14-(j-1)}}{\sum_{j=1}^{k} \binom{13}{13-(j-1)} p^{13-(j-1)}(1-p)^{j-1}} \right).$$

The first parenthesis on the right-hand side is the objective cumulative probability of facing at least $k-1$ competitors when entering. The second parenthesis is a weighted average of the risk component of each outcome greater than or equal to x_k. If we apply equation Equation (1.4) only on the second parenthesis, then we can derive the decision weights only applying probability weighting on the risk components. The right part of Table 1.A.2 shows the results. Using this method, we find again very similar number of entrants compared to our model in the main text.

Table 1.A.2: Predicted number of entrants in WTA, weighting natural risk only.

		WTA, N1			WTA, N2		
		$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$
$\gamma = 1$	$\delta = 1$	3.52	2.79	2.22	3.52	2.79	2.22
	$\delta = 0.8$	2.71	2.21	1.80	2.61	2.10	1.70
$\gamma = 0.75$	$\delta = 1$	5.83	4.03	2.83	5.97	4.14	2.82
	$\delta = 0.8$	3.81	2.76	2.09	3.82	2.61	1.83
$\gamma = 0.5$	$\delta = 1$	14.00	10.26	5.83	14.00	10.36	6.11
	$\delta = 0.8$	9.78	5.51	2.84	9.89	5.77	2.45
MEG		3.52	3.49	3.47	3.52	3.49	3.47

Notes. Expected number of entrants in the symmetric mixed strategy equilibrium with CPT preferences and probability weighting applied to natural risk only, for selected values of the three preference parameters η, γ, and δ. WTA, N1 and WTA, N2 refer to the two variants of separating natural from strategic risk. If we weigh only strategic risk, the MEG predictions do not vary in γ and δ.

1.A.2 Additional tables and figures

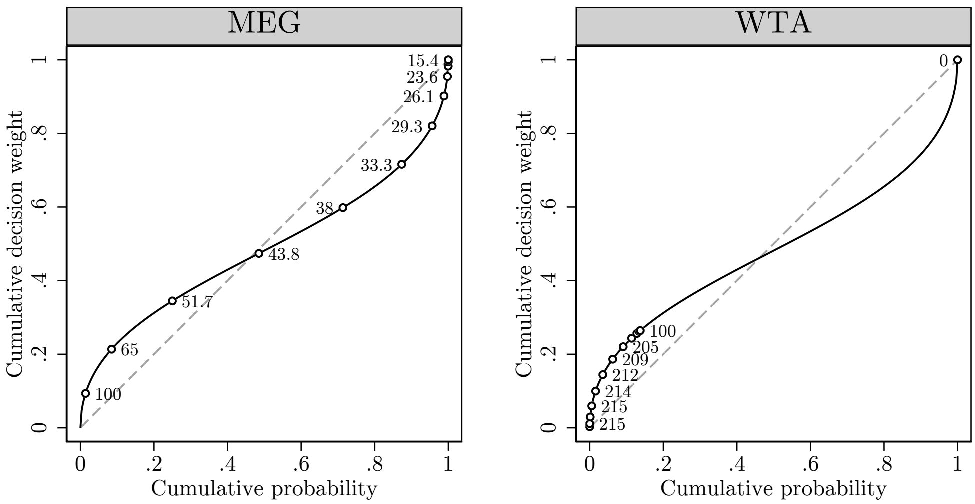


Figure 1.A.1: Probability weighting function and outcomes in the MEG and WTA conditions. The solid line shows the CPT probability weighting function (see Equation 1.4), calibrated with our estimated parameters $\hat{\gamma} = 0.517$ and $\hat{\delta} = 0.929$. Dots indicate all possible outcomes for an entrant in the MEG condition (left panel) and the WTA condition (right panel). Due to space constraints not all dots are labelled with the respective payoff.

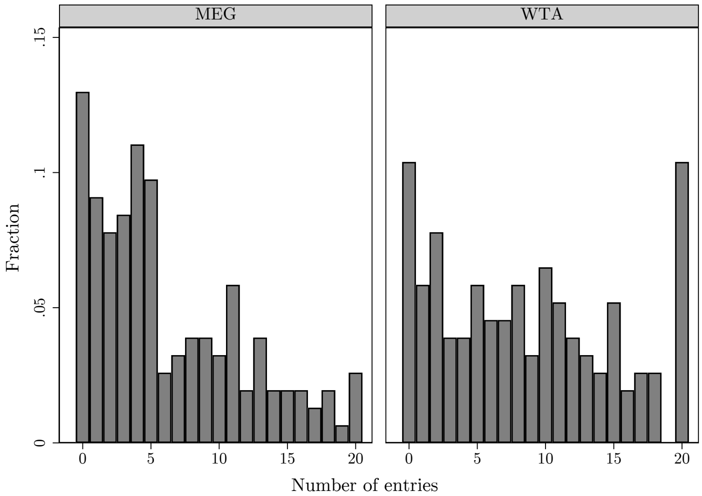


Figure 1.A.2: Histograms of individual entry frequency. Number of times a subject entered the respective market during the 20 periods, left panel for the MEG condition, right panel for the WTA condition. Data of both phases combined.

Table 1.A.3: Lotteries in the gain domain $(x_1, p; x_2)$

p	x_1	x_2	p	x_1	x_2	p	x_1	x_2
0.10	20	10	0.95	40	10	0.50	10	0
0.50	20	10	0.05	50	20	0.50	20	0
0.90	20	10	0.25	50	20	0.05	40	0
0.05	40	10	0.50	50	20	0.25	40	0
0.25	40	10	0.75	50	20	0.95	50	0
0.50	40	10	0.95	50	20	0.10	150	0
0.75	40	10	0.05	150	50			

Notes. The outcomes are denominated in CHF. Each lottery has its counterpart in the loss domain. For example, the counterpart to the first lottery $(20, 0.10; 10)$ is $(-10, 0.10; -20; e = 30)$ where e is the endowment, which covers any losses and equalizes the expected payoffs in both domains. These lotteries are the same as used in Bruhin et al. (2010a) for the Zurich 2006 sessions.

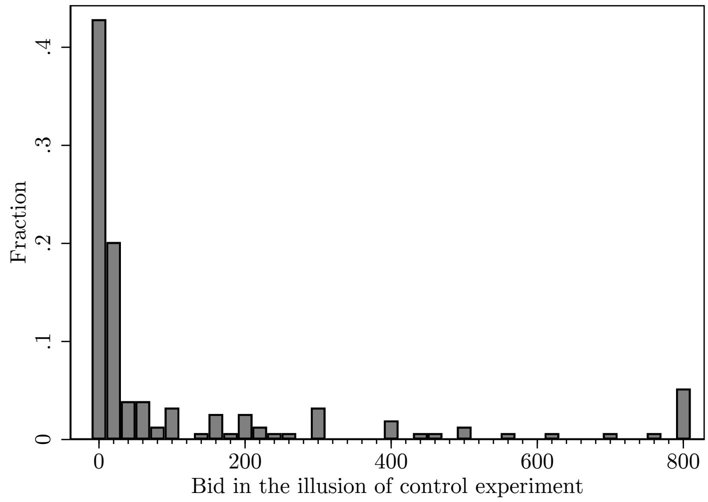


Figure 1.A.3: Histogram of individual bids in the second price auction of the illusion of control experiment. Bids are rounded up to multiples of 20.

Table 1.A.4: Summary statistics of the individual CPT parameter estimates.

	Mean	Median	Std Dev	Min	Max
CRRA $(\hat{\eta}_i)$	-0.106	0.071	1.483	-8.998	4.947
Likelihood sensitivity $(\hat{\gamma}_i)$	0.685	0.588	0.565	0.000	5.251
Optimism/pessimism $(\hat{\delta}_i)$	1.115	0.915	1.109	0.000	9.984

Number of observations = 134

Table 1.A.5: Entry frequency and individual CPT parameters.

	Dependent variable: Entry frequency (in %)	
WTA in phase 2	-14.020^{**}	-13.946^{**}
	(4.076)	(4.034)
$\hat{\eta}_i$	-1.682^{*}	-2.274
	(0.557)	(2.211)
$\hat{\gamma}_i$		-1.393
		(5.949)
$\hat{\delta}_i$		0.913
		(3.216)
Constant	49.419^{**}	49.263^{**}
	(3.833)	(4.888)
F-test	6.5	4.9
Prob $> F$	0.018	0.022
R^2	0.060	0.061
N	134	134

Notes: OLS estimates. Dependent variable is the individual entry frequency in the WTA condition (in %). Independent variables are a dummy for the WTA played after the MEG condition and the individual estimates for the three cumulative prospect theory parameters. Robust standard errors, clustered on group, in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table 1.A.6: Finite mixture model estimates.

	EUT type	CPT type
Relative size	0.127	0.873
	(0.029)	(0.029)
CRRA ($\hat{\eta}$)	0.080	0.141
	(0.013)	(0.034)
Likelihood sensitivity ($\hat{\gamma}$)		0.471
		(0.021)
Optimism/pessimism ($\hat{\delta}$)		0.930
		(0.055)
Number of subjects	134	
Number of observations	2680	
Log likelihood	-7875.8	
AIC	15765.6	
BIC	15806.9	

Notes: Robust standard errors, clustered on subjects, in parentheses.

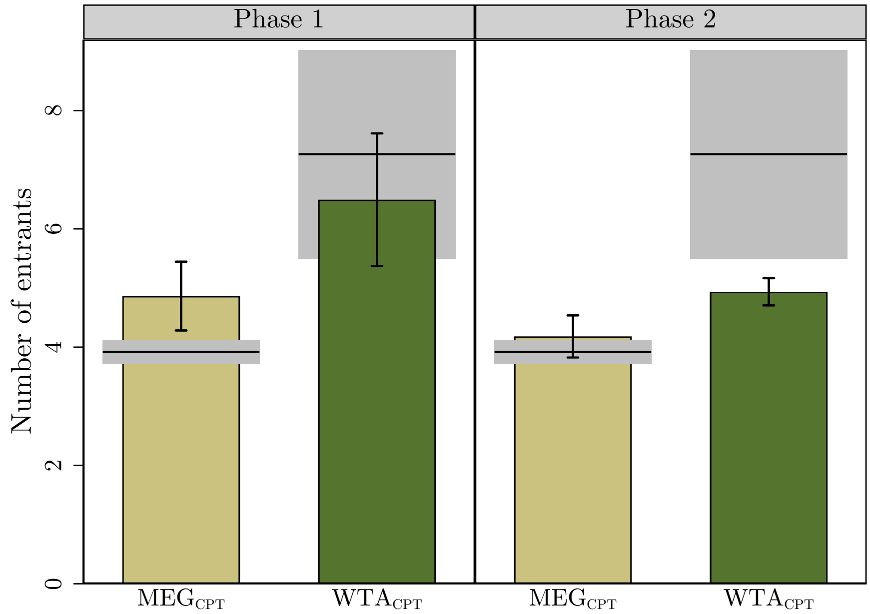


Figure 1.A.4: Average number of entrants by phase (bars with spikes indicating 95% confidence intervals) and 95% confidence interval of the predicted number of entrants (gray band).

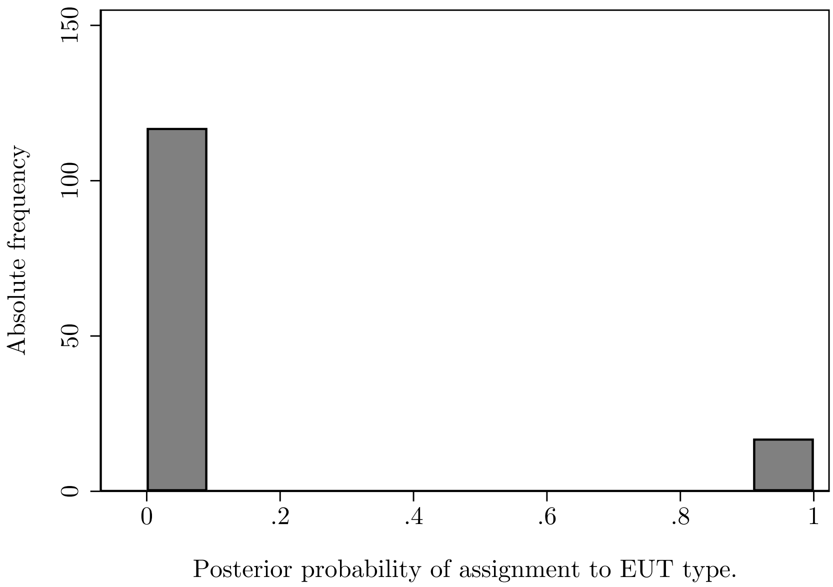


Figure 1.A.5: Histogram of posterior probability of assignment to EUT type.

1.A.3 Experimental instructions

This appendix section shows a translated version of the experimental instructions that were handed out to subjects on paper (original instructions were in German or French). First we show the instructions of the main experiments (pages 34 to 42), followed by the instructions for the additional experiments described in section 1.4.3 (pages 43 to 49) and in section 1.5.3 (pages 50 to 54).

General instructions for the participants

You are about to participate in multiple economic experiments. **The experiments are independent of each other.** If you read the instructions carefully, you can, depending on your decisions, earn more or less money. It is therefore important to read the following instructions carefully.

The instructions you received from us are solely for your private information. Communication with other participants is **strictly forbidden** during the course of the experiment. Please ask an instructor in case of any questions. If you don't comply with those rules we will have to exclude you from the experiment including from any payments.

During the experiment, your income will not be computed in Swiss francs but in **points**. The points that you earn during the experiment will be converted to Swiss francs and paid out in cash. The following exchange rate applies:

1 point = 1.5 Swiss cents.

Instructions for experiment 1

For this experiment you receive an endowment of 800 points, which corresponds to 12 Swiss francs.

The experiment in which you are participating consists mainly of a lottery. The experimenter has a deck of 28 cards. On each of these cards there is one of the following 28 symbols:

#	E	*	+	€	J	%	$	O	?	Q	ö	¬	[
|	¦	§	=	¢	7	!	(	A	£	@	Y	S	1

The 28 symbols are now distributed to the participants. At the end of the experiment two cards will be drawn at random. The two participants with the corresponding symbols on their cards will receive **50 Swiss francs** in addition to the earned points that will be converted to Swiss francs.

To begin, you are prompted on the screen to choose one of those 28 symbols. Insert your choice and confirm with the OK-button.

The probability is pretty high, though, that more than one person chooses the same symbol. This problem will be addressed on the second screen.

In the cases where a particular symbol is chosen twice or more times we run an **auction**. The participant who wins the auction receives his or her desired symbol. To the other participants the computer will randomly assign another symbol which is not yet taken. For you to be able to participate in the auction we endowed you with 800 points. Though, the auction is not a "normal" auction where the participants progressively increase their bids. It works a bit differently:

The auction we perform is a so-called **second-price auction**. In this type of auction every participant can make exactly one bid which the other participants don't see. The bids are then compared with each other. As in "normal" auctions, the participant who made the highest bid gets the good, which in our case is the desired symbol. In contrast to "normal" auctions, however, the winner only pays the second highest bid. To illustrate, consider the following example:

Suppose there are three participants A, B, and C who take part in the auctioning of symbol X. They make the following bids:

- A bids x points
- B bids y points
- C bids z points

Let z be the highest bid and x be the lowest bid: $z > y > x$. C wins the auction and receives symbol X. He or she pays for it the second highest bid which is y points. A and B receive another symbol at random which is not chosen by any other participant. After the auction C's account balance will be 800 minus y points and the account balance of A and B is 800 points each.

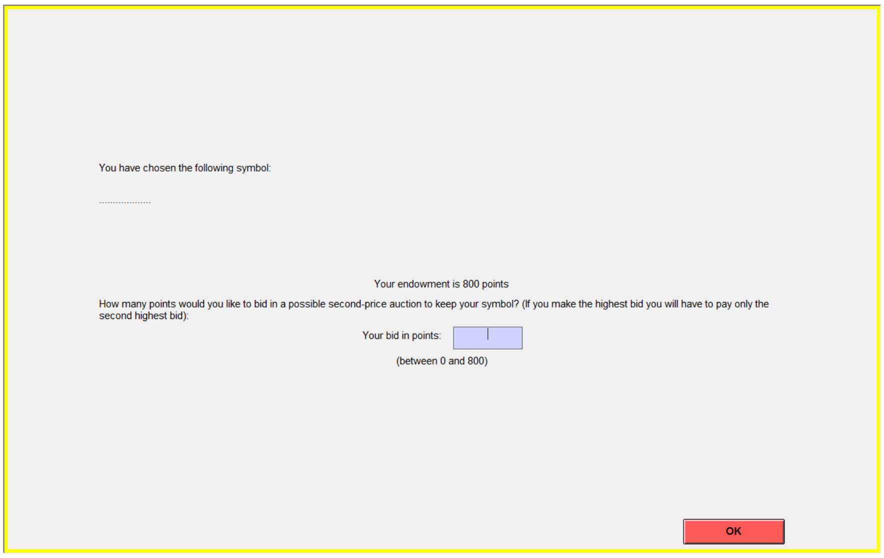
If two or more participants make the same bid, the computer will decide randomly who wins the auction. The winner will then pay exactly his or her bid as the second bid is the same.

The second-price auction has the following interesting property. **It is optimal for every participant to bid exactly as much as the good is worth to him or her.** The following example illustrates why:

> Suppose you are going to the sports store X to buy a skiing equipment set for a few hundred Swiss francs. You know exactly which one you want and you have already compared prices across different sports stores. Thus, you have already taken the decision to buy the skiing equipment set, that is, you will definitely buy it.
>
> Now, you are standing in front of the sports store and you see a crowd of people. A lady offers a voucher for sports store X with a value of 100 Swiss francs. Of course, she sells the voucher by means of a second-price auction. A few people are interested in the auction and the lady is distributing pieces of paper on which the bids should be written. How much should you bet now? Since you know that you are about to spend more than 100 Swiss francs in the sports store X, the 100 Swiss francs voucher is equivalent to 100 Swiss francs in cash. In a second-price auction it is therefore optimal for you to bid exactly 100 Swiss francs for the voucher. The explanation goes as follows:
>
> - If you bid less than 100 Swiss francs, say 90 Swiss francs, then someone who bids, for example, 91 Swiss francs will win the auction. But for 91 Swiss francs you would have been happy to buy the voucher. This reasoning holds for **any** bid below 100 Swiss francs.
> - If you bid more than 100 Swiss francs, say 105 Swiss francs, then you might end up paying too much for the voucher. For instance, if the second highest bid is 103 Swiss francs. This reasoning holds for **any** bid above 100 Swiss francs. It is never worth it to bit more than 100 Swiss francs. Not even if the second highest bid is less than 100 Swiss francs. In this case you could have bid just as well 100 Swiss francs to get the voucher at the same price as if you bid a higher price.

This line of reasoning holds independently of the number of people bidding for the voucher. **Thus, in a second-price auction it is always optimal to bid exactly as much as the good is worth to one.**

Back to the experiment: The assignment of symbols to participants for symbols that are chosen more than once will be settled by a second-price auction. To make sure that everyone is treated equally, **everyone** has to make a bid. The range of possible bids lies between 0 and 800. If you bid 0 you are, de facto, not participating in the auction. If your bid is above 0, you participate in the auction. This happens at a point in time where you don't know yet whether your symbol was chosen multiple times. After everyone has chosen a symbol in the first step, you are shown the following screen:



On the top of the page the symbol that you have chosen will be displayed. Below, you must indicate how much of your endowment of 800 points you bid, **if** a second-price auction takes place. The range of possible bids is between 0 and 800 points.

After inserting your bid, a screen will be displayed that informs you about the results. There are three possible cases:

a) **No one has chosen the same symbol as you.** In this case you will keep your chosen symbol and you will not participate in any auction. You will therefore keep your endowment of 800 points.

b) **Your symbol was chosen multiple times and your bid was the highest among all participants who have chosen the same symbol.** In this case you receive your desired symbol and you pay the second highest bid which was made for your symbol. In case somebody else with the same symbol has made the same bid you will pay exactly the amount that you have bid.

c) **Your symbol was chosen multiple times but your bid was not the highest.** In this case the computer will assign you randomly a symbol which is not yet taken. Since you didn't win the second-price auction, you keep your endowment of 800 points.

Memorize the symbol which you have chosen or the symbol which you were assigned, respectively. At the end of all experiments, there will be a random draw of a symbol to determine whether you win the lottery with your symbol.

Instructions for experiment 2

The new experiment is split into 20 rounds. At the beginning of the experiment, the participants are split into two equally sized groups. The group composition remains unchanged for the whole experiment. It is thus the same in every round.

Each member of a group has to make a choice between Alternative A and Alternative B at the beginning of each round. You can earn points depending on which alternative you choose and depending on the choices that the other participants of your group make. The rounds are identical and independent of each other throughout the experiment. That is, the number of points that you earn in a particular round depends only on the choices in this particular round. After the 20th round, all the points you earned during the experiment will be added up to your **total payoff.**

Computing the payoff in one round

In each round you must choose either Alternative A or Alternative B. If you choose Alternative A, your will participate in a prize competition with the participants of your group who also chose Alternative A. If you choose Alternative B you will not participate in the prize competition.

Your payoff if you choose Alternative A

All participants in a group who chose Alternative A participate in a prize competition where you have the chance to earn between 100 and 215 points.

Whether you win the prize competition is a matter of chance. A random number generator will assign to each participant a number between 0 and 100. For each participant, every number between 0 and 100 is equally likely. For the participants who have chosen Alternative A the random number will determine who wins the prize competition. The participant whom is assigned the *highest number* by the random number generator wins the prize competition. If you are the only participant of your group to choose Alternative A you win the prize competition in any case.

Your payoff in the price competition depends in particular on whether you win the prize competition. If you **don't win** the prize competition you get a payoff of zero points for this round. If you **win** the prize competition, then your payoff depends on how many competitors there are in the prize competition, that is, how many *other* participants also chose Alternative A. The more group members participate in the prize competition, the higher is the winner's payoff. The following Table summarizes the payoff as a function of the number of number of other participants who chose Alternative A:

Number of competitors in the prize competition	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Payoff for the **winner** [in points]	100	130	155	175	190	200	205	209	212	214	215	215	215	215

If, for instance, you and four *other* participants chose A and you are assigned the highest random number, then you receive 190 points and the other 4 participants receive zero points.

Your payoff if you choose Alternative B

If you choose Alternative B you don't participate in the prize competition. You earn 45 points with certainty in the corresponding round.

Detailed procedures of a round

At the beginning of each round, you will see an input screen (Figure 1). On the left hand side of the header you see the number of the round you are currently in. On the right hand side of the header you see how much time you have left to make your choice.

Your task on this screen is to make your choice for Alternative A or Alternative B in this round.

Furthermore, you will be asked what choices you believe the **other members** of your group make. You are asked to indicate how many *other* members of your group you believe to choose Alternative A. Please note that you should indicate the number of **other** participants of your group who chose A, without counting yourself. If your estimate about the number of other group members to choose A turns out to be correct, you will get a bonus of **5 points** for the corresponding round. Apart from that, your indicated estimate is inconsequential for the further procedure of the round.

Once you have completed your inputs, please press the OK-button to confirm.

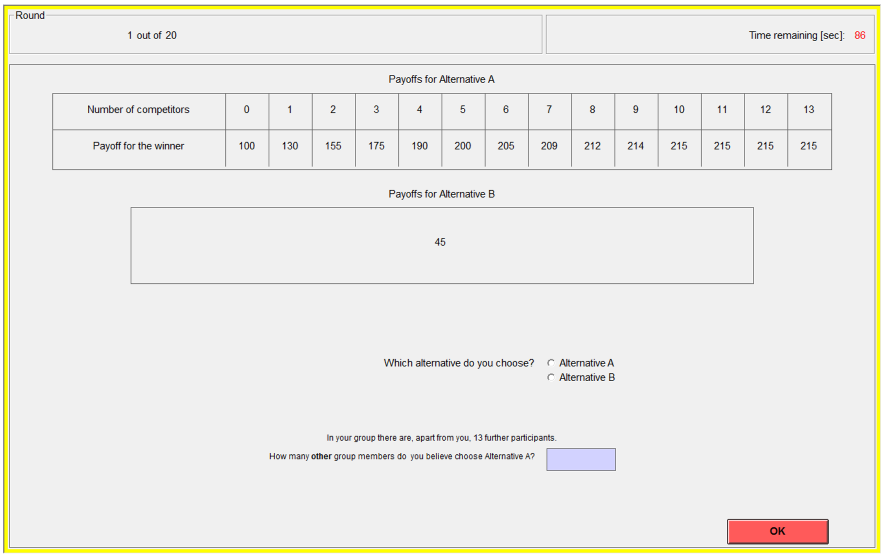


Figure 1: input screen

Subsequently, the computer generates for each participant a random number between 0 and 100. All numbers are equally likely. The participants who chose Alternative B will be assigned a random number as well which, though, is not consequential for any payoff.

For the participants who chose Alternative A, however, the random number determines whether you win the prize competition, that is, whether you earn a payoff between 100 and 215 points. Your random number will be displayed on the screen (Figure 2). Furthermore, it will be displayed how many participants chose Alternative A. Below you will see your payoff in this round and the indication, whether your estimate about the other group members' choices was correct. If this is the case, you receive 5 additional points.

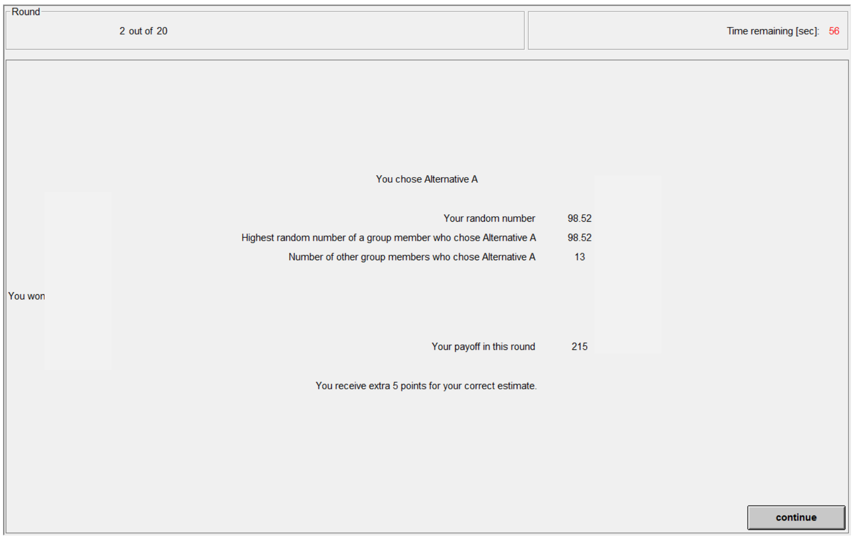


Figure 2: output screen.

Instructions for experiment 3

The third experiment is identical to the previous one in terms of the procedures. Again, there are 20 rounds to play. In each round you have to choose either Alternative A or Alternative B. The group composition is the same as in the previous experiment. The points you earn in this experiment will be added to the points that you earned in the experiments 1 and 3. The same exchange rate applies: 1 point will be converted to 1.5 Swiss cents and paid out at the end of the session.

There is one **important difference** to experiment 2:

The payoff for the participants who choose Alternative A are now determined in a different manner. There is **no** prize competition any more. Instead, everyone who chooses Alternative A receives the same number of points. As there is no prize competition anymore, there won't be random numbers either. But the number of points that you receive under Alternative A still depends on the number of other participants who also choose Alternative A. The payoffs are summarized by the following Table:

Number of other participants who choose A	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Total payoff for **all** participants who choose A	100	130	155	175	190	200	205	209	212	214	215	215	215	215

If, for instance, you and four *other* participants choose Alternative A, then you get 175 points together. This payoff is equally distributed among those who choose Alternative A so that you receive a payoff of 43.75 (=175/4) points.

Everything else remains the same. The payoff possibilities of Alternative B remain at 45. Again, you can earn 5 points by give a correct estimate of the number of other participants to choose Alternative A.

Instructions for experiment 1

Welcome

Welcome to our economic experiment. If you read the instructions carefully, you can, depending on your decisions, earn more or less money. It is therefore important to read the following instructions carefully.

The instructions you received from us are solely for your private information. Communication with other participants is **strictly forbidden** during the whole experiment. Please ask an instructor in case of any questions. If you don't comply with those rules we will have to exclude you from the experiment including from any payments.

During the experiment, your payoff will not be computed in Swiss francs but in **points**. The points that you earn during the experiment will be converted to Swiss francs and paid out in cash. The following exchange rate applies:

1 point = 1.5 Swiss cents.

Instructions for experiment (i)

In this experiment we ask you to make choices between two alternatives, **Alternative A** and **Alternative B**. At the end of each round you receive a payoff (= a certain number of points) which depends on the alternative you choose.

You have to make this choice in a total of 20 consecutive rounds. Those 20 rounds are identical and independent of each other in terms of their procedure. Thus, your payoff in a particular round depends solely on your choice in this particular round. After the 20th round, all points that you received during the 20 rounds will be added to your total payoff.

Computing the payoffs in one round

In each round you must choose either Alternative A or Alternative B. If you choose Alternative A, your payoff will be subject to randomness. If you choose Alternative B you will receive a fixed payoff of 45.

If you choose Alternative A, then it will be determined randomly whether you receive a payoff which is higher than what you get under Alternative B or whether you receive a payoff of zero. The level and the probability of the possible payoffs are variable.

The payoffs are determined based on data of a completed experiment with another group of students. These students have already received their payoffs and they are not participating in today's experiment. To make sure you comprehend the decision situations of the participants of the completed experiment, we ask you to read the **instructions of the already completed experiment.** You find them in the attachment.

++

Please continue with the attachment (colored paper).

++

++

Here we show the instructions for the WTA-condition of the main experiment

++

++

Please return now to the instructions of today's experiment (white paper).

++

++

After reading the attachment, please continue reading here.

++

Today we are interested to know, what choices **you** would make in this situation. You will also face the choice between Alternative A and Alternative B.

In contrast to the completed experiment, however, your payoff for Alternative A will **not** depend on other participants' behavior in this room. Instead, we will use the **data of the completed experiment** to determine the level of your payoffs and their winning probabilities.

For this sake, during the experiment, you will be assigned a data set of a group from the completed experiment. If, in a given round, you choose Alternative A your payoff will be subject to randomness. It will be determined based on the number of participants who chose Alternative A in the corresponding round of the completed experiment. The choices of all participants will be replicated by the computer.

The prize is determined the same way as in the completed experiment:

Number of participants in the assigned group who chose A	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Payoff [in points]	100	130	155	175	190	200	205	209	212	214	215	215	215	215

Detailed procedures of a round

At the beginning of each round an input screen will be displayed (Figure 1). Your task on this screen is to choose either Alternative A or Alternative B for this round. Furthermore, you will be asked what choices you believe the participants of the completed experiment made. More concretely, you are asked to indicate how many of the 14 members of the group which you are assigned to choose Alternative A in the corresponding round. If your estimate regarding the behavior of the participants in the completed experiment is correct, you will get a bonus of 5 points for this round. Apart from that, your indicated estimate is inconsequential for the further procedure of the round.

Once you have completed your inputs, please press the OK-button to confirm.

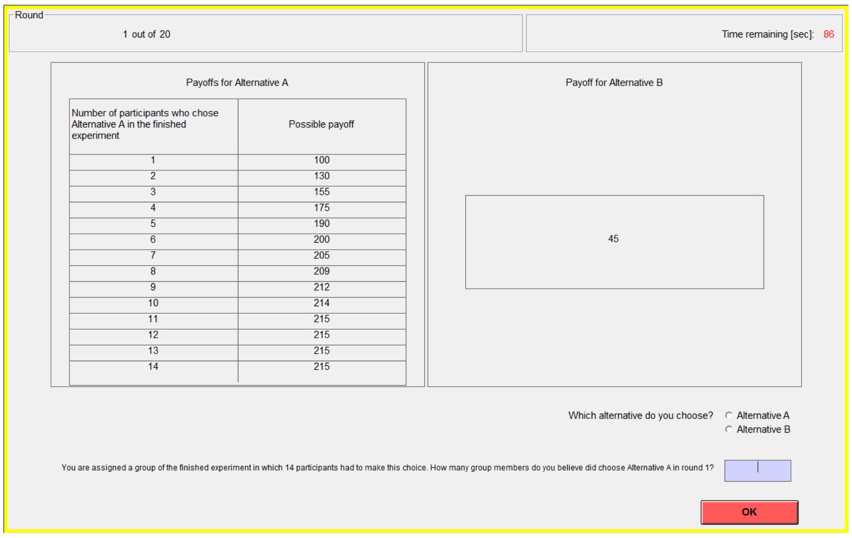


Figure 1: input screen

Subsequently, the computer will generate exactly as many random numbers as there are participants who chose Alternative A in this particular round of the completed experiment. You will be assigned one of those random numbers.

If you have chosen Alternative A, then this random number will determine whether you receive the prize, that is, whether you earn a payoff between 100 and 215 points. Is your random number the highest of all generated random numbers, then you receive a payoff of between 100 and 215 points. Otherwise you receive a payoff of zero.

If, for instance, you choose Alternative A in a given round and five participants of the completed experiment chose Alternative A in this particular round, then five random numbers will be generated. The random number that you are assigned to must be the highest for you to win the prize of 190.

The random numbers are also generated if you chose Alternative B. But they will be inconsequential for your payoff.

The results of these procedures will be displayed on the output screen (Figure 2).

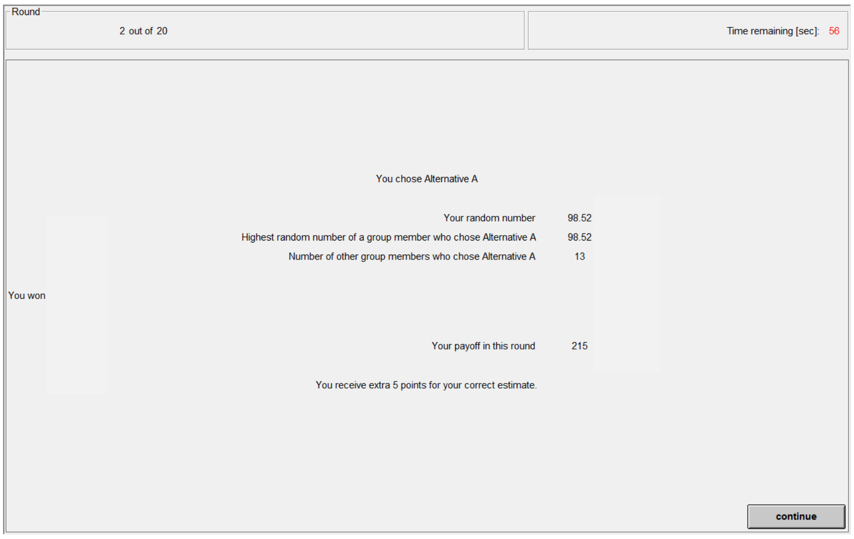


Figure 2: output screen

By clicking the „continue“-button, you get to the next round where you will face the same choice situation again.

Instructions for experiment (ii)

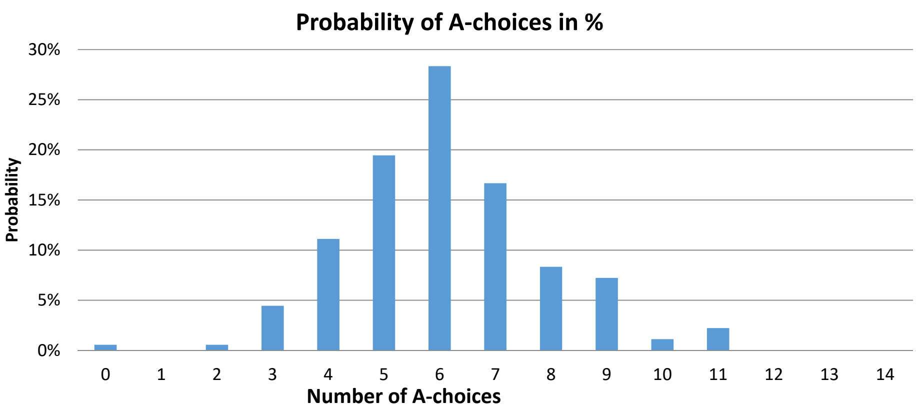
In this experiment you will face again choices between Alternative A and Alternative B. In this stage, however, we change the mechanism to determine your chances to win under Alternative A as well as your level of information.

Computing your payoff in one round

In each round you must choose either Alternative A or Alternative B. If you choose Alternative B, you will still get a payoff of 45. If you choose Alternative A you still get an uncertain payoff in the range of 0 to 215. Your winning chances are still based on the results of the already completed experiment.

In contrast to the previous experiment, you will not be assigned to a particular group for all rounds. Instead, in every round you will be randomly assigned any round of any group. Furthermore, you will learn the full statistical distribution of the A and B choices of the completed experiment.

Overall, 180 rounds were played in the completed experiment: Thereby the following distribution per round resulted:



Number of participants who chose A	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Price to win in points	100	100	130	155	175	190	200	205	209	212	214	215	215	215	215
Absolute number of rounds	1	0	1	8	20	35	51	30	15	13	2	4	0	0	0
Relative frequency in %	**0.6**	**0**	**0.6**	**4.4**	**11.1**	**19.4**	**28.3**	**16.7**	**8.3**	**7.2**	**1.1**	**2.2**	**0**	**0**	**0**

Now, in every round, one of those 180 rounds will randomly be drawn as a basis to determine your payoff profile under Alternative A. If you happen to get a round where no participant chose Alternative A, you will get a payoff of 100 for this round.

Detailed procedures of a round

Again, at the beginning of each round an input screen will be displayed.

Your task is again to indicate for which of the two Alternatives A or B you choose in this round.

Furthermore, we will still ask you to indicate your beliefs about the behavior of the participants of the completed experiment. Now you have to indicate how many of the 14 group members chose Alternative A in the randomly drawn round. If your belief is correct, you get a bonus of 5 points.

Once you have completed your inputs, please press the OK-button to confirm.

Subsequently, the computer will again generate exactly as many random numbers as there are participants who chose Alternative A in this particular round of the completed experiment. You will be assigned one of those random numbers. If you have chosen Alternative A, then this random number will determine your payoff according to the same procedure as described above.

The results of these procedures will be displayed on the output screen. The screens correspond to the ones of the previous experiment.

Instructions CPT elicitation task

Instructions for experiment 3

In this last experiment, you do not interact anymore with the other participants. Your payoff for this experiment will only depend on your choices and luck.

Unlike the first two experiments, we won't use points but ECUs. The new exchange rate for this experiment is:

3 ECU = 1 CHF

You will face 40 situations of similar form and will have to take a decision for each of them.

You will have 40 minutes to take your decisions. During this time, you will be able to modify any decision at any time. Since you do not interact anymore with the other participants, the experiment will end up when you will be done with your decisions. You can move at your own pace. The 40 minutes should be more than enough to take your decisions.

Click on continue to see the next instructions

Continue

Instructions for experiment 3 (continued)

The 40 situations are independent from each other. For each situation, you have to choose between uncertain amounts of ECU (Option A) and certain amounts of ECU (Option B). The uncertain amount of Option A is randomly determined by the computer according to probabilities displayed on the screen.

There are two types of decisions. The amounts can either be positive (gain situation) or negative (loss situation). In loss situations, a specific amount of ECU is given to you (endowment) to compensate for any possible loss. This experiment cannot reduce the amount of money you have earned so far. For each situation, you receive the following information: the uncertain amounts of Option A and the probability attached to each amount, the certain amounts of Option B, and the endowment (only for loss situations).

After you validate your decisions, a situation is randomly drawn by the computer and your payoff will depend on your decision for this situation. Your payoff will be added to your payoff from the two previous experiments. Later, a screen will inform you on your total payoff for the 3 experiments and for each of them.

Click on continue to see an example of gain situation and one of loss situation.

Continue

Instructions for experiment 3 (continued)

Every situation will have a similar structure as the own showed on the right. The table contains 20 rows. Option A is the same for each row whereas Option B is different for each one.

Example of gain situation

In Option A, the gain is of 30.0 ECU with probability 25% and of 10.0 ECU with probability 75%.

In Option B, the certain gains range from 30.0 ECU to 11.00 ECU.

For each row, you have to choose if you prefer the corresponding certain gain (Option B) or if you prefer Option A.

To enter your decision, you will have to click on one of the gray buttons in the middle either in column A or column B. You do not have to enter your choice for each row because if you choose the Option B for a certain amount, then Option B is automatically chosen for all larger certain amounts. In other words, you have to choose from certain amount on you prefer Option B to Option A.

Click on Continue to see an example of loss situation.

n°	Option A	Your Choice				Option B A certain gain of ... ECU
1		A			B	30.0
2		A			B	29.0
3		A			B	28.0
4		A			B	27.0
5	**Gain**	A			B	26.0
6	**of 30.0 ECU with**	A			B	25.0
7	**probability**	A			B	24.0
8	**25%**	A			B	23.0
9		A			B	22.0
10	**and**	A			B	21.0
11		A			B	20.0
12		A			B	19.0
13	**Gain**	A			B	18.0
14	**of 10.0 ECU with**	A			B	17.0
15	**probability**	A			B	16.0
16	**75%**	A			B	15.0
17		A			B	14.0
18		A			B	13.0
19		A			B	12.0
20		A			B	11.0

Continue

Instructions for experiment 3 (continued)

Example of loss situation

In Option A, the loss is of 10.0 ECU with probability 75% and of 30.0 ECU with probability 25%.

In Option B, the certain losses range from 11.0 ECU to 30.00 ECU.

Contrary to a gain situation, you receive an endowment. This endowment ensures you do not lose money if the situation is chosen for the payment. The endowment is displayed at the top of the screen. Here, it is 40.0 ECU.

For each row, you have to choose if you prefer the corresponding certain loss (Option B) or if you prefer Option A.

Click on Continue to see how your payoff is determined.

Endowment: 40.0 ECU

n°	Option A	Your Choice				Option B A certain loss of ... ECU
1		A			B	-11.0
2		A			B	-12.0
3		A			B	-13.0
4		A			B	-14.0
5	**Loss**	A			B	-15.0
6	**of 10.0 ECU with**	A			B	-16.0
7	**probability**	A			B	-17.0
8	**75%**	A			B	-18.0
9		A			B	-19.0
10	**and**	A			B	-20.0
11		A			B	-21.0
12		A			B	-22.0
13	**Loss**	A			B	-23.0
14	**of 30.0 ECU with**	A			B	-24.0
15	**probability**	A			B	-25.0
16	**25%**	A			B	-26.0
17		A			B	-27.0
18		A			B	-28.0
19		A			B	-29.0
20		A			B	-30.0

Continue

Instructions for experiment 3 (continued)

Payoff determination (gain situation)

At the end of the experiment, a situation is drawn at random. Suppose that the situation on the right is drawn and that your choice is as shown. For the line n°10 and all of the above, you prefer Option B to Option A. In other words, for a certain gain larger or equal to 21.0 ECU, you prefer to get the certain amount of the row rather than to choose the uncertain Option A.

A row between 1 and 20 (first column) is then randomly drawn by the computer. In this example, suppose that row n°9 is randomly drawn. Because Option B was chosen for this line, your payoff would be the certain gain for this row (22.0 ECU). The payoff determination is similar for all the rows above row n°9 in this example. The payoff being always the certain amount of the randomly chosen row.

If row n°11 was randomly chosen, your payoff would depend on the outcome of Option A. In this case, you would win 30.0 ECU with probability 25% or 10.0 ECU with probability 75%. Again, the computer would determine which of the two amounts you would win. The payoff determination is the same for all lines below line n°11 in this example.

n°	Option A	Your Choice				Option B A certain gain of ... ECU
1		A		X	B	30.0
2		A		X	B	29.0
3		A		X	B	28.0
4		A		X	B	27.0
5	**Gain**	A		X	B	26.0
6	**of 30.0 ECU with**	A		X	B	25.0
7	**probability**	A		X	B	24.0
8	**25%**	A		X	B	23.0
9		A		X	B	22.0
10		A		X	B	21.0
11	**and**	A	X		B	20.0
12		A	X		B	19.0
13	**Gain**	A	X		B	18.0
14	**of 10.0 ECU with**	A	X		B	17.0
15	**probability**	A	X		B	16.0
16	**75%**	A	X		B	15.0
17		A	X		B	14.0
18		A	X		B	13.0
19		A	X		B	12.0
20		A	X		B	11.0

Continue

Instructions for experiment 3 (continued)

Payoff determination (loss situation)

Instead, suppose that the randomly chosen situation is the one on the right and that your choice is as shown. For the line n°13 and all of the above, you prefer Option B to Option A. In other words, for a certain loss smaller or equal to 23.0 ECU, you prefer to incur the certain loss of the row rather than to choose the uncertain Option A.

Now suppose that row n°12 is randomly drawn for the payment. Because Option B was chosen for this line, your payoff would be the certain loss for this row (22.0 ECU). The payoff determination is similar for all the rows above row n°12 in this example. The payoff being always the certain loss of the randomly chosen row.

If row n°14 was randomly chosen, your payoff would depend on the outcome of Option A. In this case, you would incur a loss of -10.0 ECU with probability 75% or -30.0 ECU with probability 25%. Again, the computer would determine which of the two losses you would incur. The payoff determination is the same for all lines below line n°14 in this example.

Your final payoff would be equivalent to the endowment (40.0 ECU) from which we subtract the loss you incur. Remember, the endowment is always such that you cannot end up with a negative payoff.

Endowment: 40.0 ECU

n°	Option A	Your Choice				Option B A certain loss of ... ECU
1		A		X	B	-11.0
2		A		X	B	-12.0
3		A		X	B	-13.0
4		A		X	B	-14.0
5	**Loss**	A		X	B	-15.0
6	**of 10.0 ECU with**	A		X	B	-16.0
7	**probability**	A		X	B	-17.0
8	**75%**	A		X	B	-18.0
9		A		X	B	-19.0
10	**and**	A		X	B	-20.0
11		A		X	B	-21.0
12		A		X	B	-22.0
13	**Loss**	A		X	B	-23.0
14	**of 30.0 ECU with**	A	X		B	-24.0
15	**probability**	A	X		B	-25.0
16	**25%**	A	X		B	-26.0
17		A	X		B	-27.0
18		A	X		B	-28.0
19		A	X		B	-29.0
20		A	X		B	-30.0

Continue

Instructions for experiment 3 (continued)

Control questions

Please answer to the two following question to make sure you understood the instructions. Suppose your choice is as shown on the right.

Question 1:

What is your payoff if row n°8 is randomly drawn for the payment?

30.0 ECU with probability 25% and 10.0 ECU with probability 75%	23.0 ECU
21.0 ECU	30.0 ECU with probability 75% and 10.0 ECU with probability 25%

n°	Option A	Your Choice				Option B A certain gain of ... ECU
1		A		X	B	30.0
2		A		X	B	29.0
3		A		X	B	28.0
4		A		X	B	27.0
5	**Gain**	A		X	B	26.0
6	**of 30.0 ECU with**	A		X	B	25.0
7	**probability**	A		X	B	24.0
8	**25%**	A		X	B	23.0
9		A		X	B	22.0
10	**and**	A		X	B	21.0
11		A	X		B	20.0
12		A	X		B	19.0
13	**Gain**	A	X		B	18.0
14	**of 10.0 ECU with**	A	X		B	17.0
15	**probability**	A	X		B	16.0
16	**75%**	A	X		B	15.0
17		A	X		B	14.0
18		A	X		B	13.0
19		A	X		B	12.0
20		A	X		B	11.0

Instructions for experiment 3 (continued)

Control questions

Please answer to the two following question to make sure you understood the instructions. Suppose your choice is as shown on the right.

Question 2:

Without considering the endowment of 40.0 ECU, what would be your loss if row n°15 was randomly chosen?

10.0 ECU with probability 75% and 30.0 ECU with probability 25%	25.0 ECU
23.0 ECU	10.0 ECU with probability 25% and 30.0 ECU with probability 75%

Endowment: 40.0 ECU

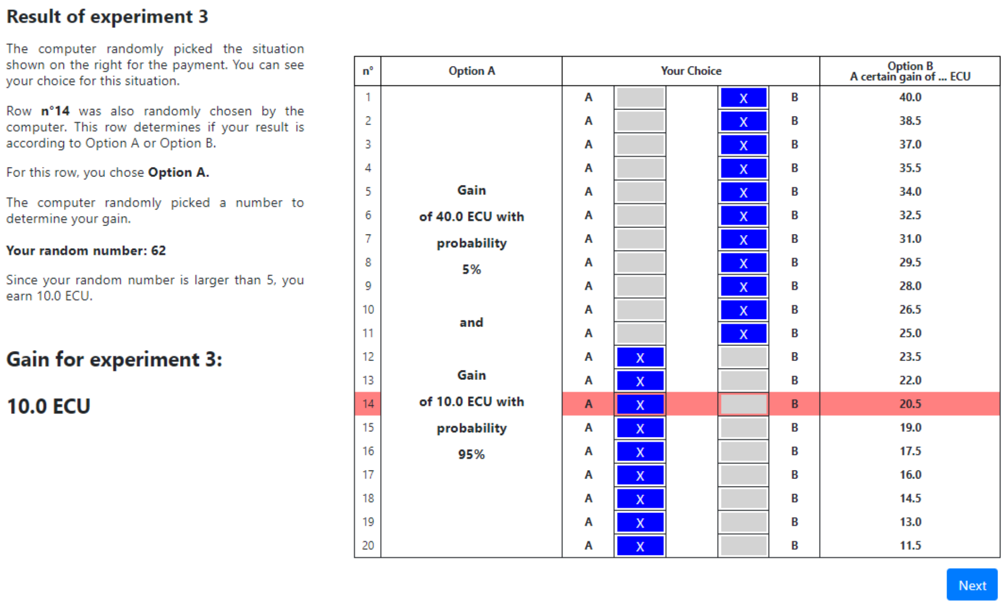
n°	Option A	Your Choice				Option B A certain loss of ... ECU
1		A		X	B	-11.0
2		A		X	B	-12.0
3		A		X	B	-13.0
4		A		X	B	-14.0
5	**Loss**	A		X	B	-15.0
6	**of 10.0 ECU with**	A		X	B	-16.0
7	**probability**	A		X	B	-17.0
8	**75%**	A		X	B	-18.0
9		A		X	B	-19.0
10	**and**	A		X	B	-20.0
11		A		X	B	-21.0
12		A		X	B	-22.0
13	**Loss**	A		X	B	-23.0
14	**of 30.0 ECU with**	A	X		B	-24.0
15	**probability**	A	X		B	-25.0
16	**25%**	A	X		B	-26.0
17		A	X		B	-27.0
18		A	X		B	-28.0
19		A	X		B	-29.0
20		A	X		B	-30.0

Situation 2 / 40

n°	Option A	Your Choice				Option B A certain gain of ... ECU
1		A		X	B	40.0
2		A		X	B	38.5
3		A		X	B	37.0
4	**Gain**	A		X	B	35.5
5	**of 40.0 ECU with**	A		X	B	34.0
6	**probability**	A		X	B	32.5
7		A		X	B	31.0
8	**5%**	A	X		B	29.5
9		A	X		B	28.0
10	**and**	A	X		B	26.5
11		A	X		B	25.0
12		A	X		B	23.5
13	**Gain**	A	X		B	22.0
14	**of 10.0 ECU with**	A	X		B	20.5
15	**une probability de**	A	X		B	19.0
16		A	X		B	17.5
17	**95%**	A	X		B	16.0
18		A	X		B	14.5
19		A	X		B	13.0
20		A	X		B	11.5

Back Next

Figure 1: Decision screen



Result of experiment 3

The computer randomly picked the situation shown on the right for the payment. You can see your choice for this situation.

Row **n°14** was also randomly chosen by the computer. This row determines if your result is according to Option A or Option B.

For this row, you chose **Option A.**

The computer randomly picked a number to determine your gain.

Your random number: 62

Since your random number is larger than 5, you earn 10.0 ECU.

Gain for experiment 3:

10.0 ECU

n°	Option A	Your Choice				Option B A certain gain of ... ECU
1	Gain of 40.0 ECU with probability 5% and Gain of 10.0 ECU with probability 95%	A		X	B	40.0
2		A		X	B	38.5
3		A		X	B	37.0
4		A		X	B	35.5
5		A		X	B	34.0
6		A		X	B	32.5
7		A		X	B	31.0
8		A		X	B	29.5
9		A		X	B	28.0
10		A		X	B	26.5
11		A		X	B	25.0
12		A	X		B	23.5
13		A	X		B	22.0
14		A	X		B	20.5
15		A	X		B	19.0
16		A	X		B	17.5
17		A	X		B	16.0
18		A	X		B	14.5
19		A	X		B	13.0
20		A	X		B	11.5

Next

Figure 2: Result screen

Chapter 2

Multigame Contact: A Double-Edged Sword for Cooperation

VINCENT LAFERRIÈRE†, JOAO MONTEZ, CATHERINE ROUX, AND CHRISTIAN THÖNI

We study the effect of multigame contact on cooperation. In our experimental setup, subjects play a pair of indefinitely repeated prisoner's dilemma games either with the same partner, or with two different partners. In contrast to our theoretical prediction, we find no evidence that multigame contact increases overall cooperation. Nonetheless, we observe that multigame contact systematically affects behavior: subjects link their decisions across games when playing with the same partner. Multigame contact proves to be a double-edged sword, as simultaneous cooperation and simultaneous defection in the two games are both more likely under multigame contact.

2.1 Introduction

Many strategic situations involve players repeatedly interacting across multiple games. Even if each game is payoff-independent—in the sense that the payoffs accruing in each stage game only depend on the actions chosen in that game—repeated games can become strategically connected if players' actions in one game depend on the outcome of another game. For instance, some co-workers are also neighbors, and how loudly they play music at home may influence how they collaborate at work. Likewise, spouses who are business partners must still share household chores, and countries that are trading partners may hold different views on human rights. These links can lead to meaningful differences in strategies and outcomes.

Repeated interactions in a single game have been widely studied from both a theory and an experimental perspective. Most relevant to our work is the recent experimental literature on indefinitely repeated prisoner's dilemmas studying the determinants of cooperation.[1] This literature has identified several factors that help determine the extent to

†This chapter is a joint project with Joao Montez, Catherine Roux, and Christian Thöni. I participated in the design of the experiment, personally ran all experimental sessions, performed a large part of the data analysis, and was heavily involved in writing the article.

[1]For a comprehensive overview, see Dal Bó and Fréchette (2018). Indefinitely repeated games are also known as infinitely repeated games, with the former term being privileged to emphasize that in laboratory

which players cooperate in these games, such as continuation probability (Dal Bó, 2005; Duffy & Ochs, 2009; Normann & Wallace, 2012), experience (Dal Bó & Fréchette, 2011), communication possibilities (Cooper & Kühn, 2014), monitoring structure (Aoyagi et al., 2019; Camera & Casari, 2009), costly punishment (Camera & Casari, 2009; Dreber et al., 2008), timing of play (Bigoni et al., 2015; D. Friedman & Oprea, 2012), and behavioral spillovers (Bednar et al., 2012).[2] However, we do not know if and how simultaneously playing multiple indefinitely repeated prisoner's dilemmas with the same partner (multigame contact) or different partners influences cooperation. We start to fill this gap.

The theoretical understanding of our subject has been a staple in industrial economics, where it is well established that multimarket contact can increase collusive behavior among firms in pricing games with a prisoner's dilemma structure. The mechanism behind this result, first formalized by Bernheim and Whinston (1990), is that firms can pool their incentive constraints across markets, i.e., use the slack in the collusive incentives in one market to compensate for the lack of collusive incentives in another market. This reduces the critical discount factor that equalizes the long-term gains from continuing collusion with the short-term gains from deviation followed by perpetual punishment. Spagnolo (1999a) further shows that concavity in the players' utility functions reinforces this result: the critical discount factor for which collusion in all markets can be sustained becomes even lower relative to the linear utility case. Some industry studies have found evidence consistent with this hypothesis, but the endogeneity problem remains challenging to address in empirical work.[3] This further motivates an experimental approach.

In this article, we set up a laboratory experiment where each subject simultaneously plays a pair of indefinitely repeated prisoner's dilemmas. Our main treatment comprises the presence or absence of multigame contact. Multigame contact is present when a subject interacts with the same partner in both indefinitely repeated prisoner's dilemmas, and it is absent when the subject faces a distinct partner in each of the two games. We pair a *hard* with an *easy game* such that the incentives to deviate from the most cooperative path are higher in the former than the latter, and therefore the critical discount factor at which cooperation is sustainable is higher in the *hard* than the *easy game*.

In this framework, theory predicts that multigame contact weakly increases the possibilities for cooperation. We consider three discount factor conditions such that, in theory: i) with the lowest discount factor, cooperation is possible in neither of the games with and without multigame contact, ii) with the intermediate discount factor, cooperation is possible in both games with multigame contact, but only possible in the *easy game* without multigame contact, and iii) with the highest discount factor, cooperation is possible in both games with and without multigame contact.

In contrast to the theoretical predictions, we find no evidence of multigame contact facilitating overall cooperation in our experiment. Nonetheless, we find strong evidence that subjects' behavior in one game is influenced by what happens in the other game in the presence of multigame contact, which is what the theory predicts. When playing with the same partner, we observe that i) subjects tend to revert to uncooperative behavior in all

experiments such games are implemented using some random stopping procedure. The literature tends to use the two terms interchangeably (Dal Bó & Fréchette, 2018; Fréchette & Yuksel, 2017).

[2]A related body of literature studies the behavioral effects of playing multiple repeated games with a finite time horizon (Falk et al., 2013; Savikhin & Sheremeta, 2013).

[3]These industries include cement (Ghemawat & Thomas, 2008; Jans & Rosenbaum, 1997), telecommunications (Busse, 2000; Parker & Röller, 1997), radio (Waldfogel & Wulf, 2006), hotels (Fernandez & Marin, 1998), airlines (Ciliberto & Williams, 2014; Evans & Kessides, 1994; Miller, 2010; Singal, 1996), hospitals (Schmitt, 2018), and banking (Coccorese & Pellecchia, 2009; Heggestad & Rhoades, 1978).

games in reaction to a deviation from cooperative behavior in a single game, and ii) cooperation in the *easy game* is more strongly linked with cooperative outcomes in the *hard game*. To the extent that people resort to uncooperative behavior at times, this implies that punishment occurs more often than theory would predict, and—in this experiment—the effect of multigame contact averages out.[4] In summary, multigame contact acts as a double-edged sword: it increases not only cooperation but also defection in both games (and thus, cooperation in only one of the games becomes less likely).

The notion that linking games can have adverse effects has been previously explored in a handful of theoretical and experimental papers. In a theoretical framework, Spagnolo (1999b) studies the case where, with some probability, a player does not behave cooperatively in one of the two games, in which case linking the games may have an adverse effect, not only for the players who link but also for the players who do not link. In a setting with imperfect monitoring, Thomas and Willig (2006) show that strategically linking multiple games may be disadvantageous because a mistaken deviation from cooperation in one game triggers punishments with uncooperative behavior in all games. The losses due to this contagion outweigh the gains from strategic linkage when the level of accuracy in monitoring the actions in one game is very low. In this case, players may want to avoid linkages if they have the possibility to do so. Experimentally, Vespa and Wilson (2019) observe contagion in a dynamic version of an infinitely repeated prisoner's dilemma with a high and low payoff state. Despite initial cooperation in the low state, subsequent defection in the high state can reduce cooperation rates in the low state.

To further explore the mechanism, we set up a second experiment with a sequential variant of the game, in which the cooperation-enhancing effect of multigame contact is in theory strongly increased. We again find compelling evidence for linkage but no effect on overall cooperation, and thus we conclude that multigame contact is indeed a double-edged sword, being a benefit for some and a curse for others.

Our study contributes to experimental literature focusing on multimarket contact and cooperation, which finds mixed results (Feinberg & Sherman, 1985, 1988; Freitag et al., 2021; Güth et al., 2016; Modak, 2021; Phillips & Mason, 1992, 1996; Yang et al., 2016). We improve on the existing literature in multiple ways. First, we build on recent methodological contributions by implementing indefinite repetition with a random stopping rule instead of the commonly-used finite horizon (Dal Bó, 2005). Second, in our setup, the subjects always interact in two games, with the variation being whether they interact with the same or a different partner in each game. This stands in contrast to earlier papers that compare a single-market to a multimarket environment with the same partner in every market. Third, we vary the continuation probability. This is an important step since the theoretical predictions on multimarket contact change with this parameter, which could explain the mixed results found in the literature so far (as none of the papers has taken this into account). Finally, we introduce a novel design that provides a stronger test of the theory in a second experiment: the sequential variant of the game ensures—in theory—that cooperation is possible with multigame contact for almost any discount factor.

The remainder of this article is organized as follows. Section 2.2 derives the theoretical predictions. Section 2.3 describes the experimental design. Section 2.4 discusses our results. Finally, Section 2.5 briefly concludes.

[4] This result is reminiscent of Dreber et al. (2008), which finds—in a single prisoner's dilemma setting—that adding a costly punishment strategy increases the frequency of cooperation but does not significantly change the average group payoff (since benefits are dissipated by the usage of the costly punishment option).

2.2 Theory

In this section, we explain the theory underlying our experiment. In Section 2.2.1, we investigate the effect of multigame contact in indefinitely repeated prisoner's dilemmas when players either choose their actions for both games at the same time or play the games sequentially. In Section 2.2.2, we derive the theoretical predictions for the game parameters used in our experiments.

2.2.1 Cooperation with multigame contact

Consider the stage game in Figure 2.1 where C stands for cooperation, and D for defection. The payoff matrix comprises four elements: the reward payoff from joint cooperation (R), the temptation payoff earned from defection when the other player cooperates (T), the sucker's payoff from cooperation when the other player defects (S), and the punishment payoff from mutual defection (P). Under the restriction $T > R > P > S$, the stage game is a prisoner's dilemma, and (D, D) its unique Nash equilibrium.

		Player 2	
		C	D
Player 1	C	R, R	S, T
	D	T, S	P, P

Figure 2.1: Payoff matrix of a single stage game

This stage game is repeated infinitely, and players discount the future with a common discount factor $\delta \in (0, 1)$. If $2R > T + S$, then a dynamic cooperative path with (C, C) in every period dominates a path with alternating strategies (C, D) and (D, C) across periods for every δ. In our experiment, we interpret δ as the probability with which the game will continue into the next round t. In laboratory experiments, such games are known as *indefinitely* repeated games, since players know the game stops after any period with probability $(1 - \delta)$ and they cannot infer for certain how long the game will last.

For sufficiently high discount factors, mutual cooperation (C, C) in every period can be sustained as a subgame-perfect equilibrium. The lowest δ at which cooperation is subgame perfect is achieved with the following grim trigger strategies: play C in every t, and play D forever after any deviation from (C, C). This critical threshold for δ is obtained by solving the incentive compatibility constraint below:

$$\frac{R}{1-\delta} \geq T + \frac{\delta P}{1-\delta} \Leftrightarrow \delta \geq \frac{T-R}{T-P} \tag{2.1}$$

The left-hand side of the first inequality denotes the present discounted payoff from cooperation in every period, whereas the right-hand side denotes the present discounted payoff from deviation. This critical threshold is decreasing in R, and increasing in T and P.

Given that we are interested in the effect of multigame contact on cooperation, we turn now our attention to a situation in which a player simultaneously engages in two indefinitely repeated prisoner's dilemmas and immediately learns the outcome of each stage game she plays. Simultaneously playing the stage games of two identical indefinitely repeated prisoner's dilemmas (with either the same or different partners) does not affect the critical discount factor at which cooperation is sustainable (Bernheim & Whinston, 1990).[5] We thus consider stage games with asymmetric payoffs, where we add a factor $z > 0$ to the temptation payoff T in one game and subtract z from T in the other one (see Figure 2.2). Because the gain from deviating from (C, C) in the game in which we added z is higher than in the game in which we subtracted it, we call the former the *hard game* and the latter the *easy game*. To keep a similar incentive structure to the stage game above, we assume that $T - z > R > P > S$, and $2R > (T + z) + S$.

Hard game

	C	D
C	R, R	$S, T+z$
D	$T+z, S$	P, P

Easy game

	c	d
c	R, R	$S, T-z$
d	$T-z, S$	P, P

Figure 2.2: Payoff matrices of the two stage games

In the absence of multigame contact, i.e., when facing a different partner in each game, a player's strategy in one game cannot affect the action of her partner in the other game. Therefore, each game can be treated independently, and cooperation in each game is respectively sustainable if:

$$\delta \geq \frac{(T+z)-R}{(T+z)-P} = \delta_{hard}, \quad \delta \geq \frac{(T-z)-R}{(T-z)-P} = \delta_{easy} \tag{2.2}$$

Note that cooperation is easier to sustain in the *easy* than the *hard game*, i.e., $\delta_{easy} < \delta_{hard}$.

Consider now the situation in which two players interact with each other in both the *easy* and *hard game*, i.e., a situation with multigame contact. The two players may still play these games as if they faced a different partner in each game, i.e., as if the games were independent, and the critical discount factors would still be the ones presented above. However, the two players may achieve cooperation in both games more easily if they link the strategies, as discussed next. When facing the same partner, each player can use the threat of punishment in both games following any deviation. This threat will pool the two incentive constraints, which induces cooperation in both games if the following single incentive constraint is satisfied:

$$\frac{2R}{1-\delta} \geq (T+z) + (T-z) + \frac{\delta 2P}{1-\delta} \Leftrightarrow \delta \geq \frac{T-R}{T-P} = \delta_{pool} \tag{2.3}$$

The payoff from perpetual cooperation in both games is given by the left-hand side of the first inequality. The payoff from defection followed by perpetual punishment is given by the right-hand side: as punishment is expected to occur in both games regardless of the form of deviation, a player who defects would optimally do so in both games simultaneously. Because $\delta_{pool} < \delta_{hard}$, cooperation in both games is indeed easier to sustain if strategies are linked.

The most cooperative outcomes are achieved—using grim trigger strategies—as follows: for $\delta \geq \delta_{hard}$, players should cooperate in each game separately, and thus linkage becomes superfluous; for $\delta_{pool} \leq \delta < \delta_{hard}$ players should link the strategies in the two games; for $\delta_{easy} \leq \delta < \delta_{pool}$ players should not link the strategies to be able to cooperate

[5] Spagnolo (1999a) shows that this neutrality result hinges on the linearity assumption, as the critical discount factor is reduced with multigame contact if utilities are concave.

in at least the *easy game*; and for $\delta < \delta_{easy}$ players should not cooperate in either game and again linkage becomes superfluous.

Note that the discount factors that enable cooperation in *every period*—as characterized above—remain unchanged if the indefinitely repeated game is preceded by a phase in which the stage game is repeated with certainty a finite number of times (as we do in our experiment, where subjects—at the outset—play two additional guaranteed rounds).

Next, we explain why linkage may not be optimal. Notice that for $\delta_{easy} \leq \delta < \delta_{pool}$, if a player expects her partner to use a grim trigger strategy with linkage following a deviation in one game, her optimal response is to deviate immediately in both games since the incentive constraint (2.3) is violated. In this case, cooperation will not be achieved in either game. However, cooperation in the *easy game* alone is still achievable if players do not link the grim trigger strategies, and instead play each game independently. So linkage helps sustaining full cooperation for intermediate discount factors, but it may destroy partial cooperation when the discount factor is sufficiently low. That is, in theory linkage should be made conditional on the discount factor.[6]

Turning to our main treatment of the presence or absence of multigame contact, and focusing on the most cooperative outcome, our theoretical predictions are summarized in Figure 2.3 below: if $\delta < \delta_{easy}$ then in both cases—with and without multigame contact—cooperation is not sustainable in either game; if $\delta_{easy} \leq \delta < \delta_{pool}$, then cooperation is only sustainable in the *easy game* in both cases; if $\delta_{pool} \leq \delta < \delta_{hard}$, then cooperation is sustainable in both games in the presence of multigame contact whereas cooperation is only sustainable in the *easy game* in the absence of it; and if $\delta_{hard} \leq \delta$, cooperation is sustainable in both games in the presence and the absence of multigame contact.

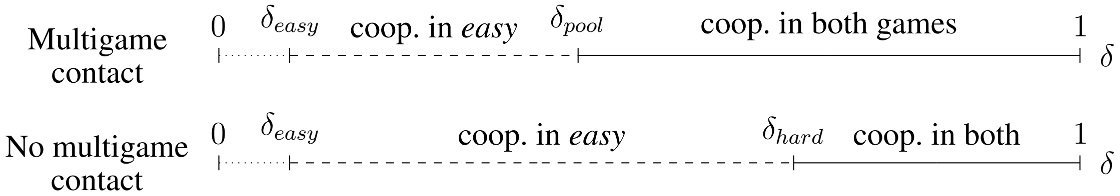


Figure 2.3: Most cooperative outcomes

In light of the results of our initial experiment where the two games are played simultaneously, we decided to study a second setting in which the effect of multigame contact on cooperation is theoretically reinforced by playing the two games sequentially. Each period now comprises two stages: players first interact in the *hard game* and then—knowing the outcome of that stage game—play the *easy game* (we focus on this sequence, since in theory it has a stronger effect on the critical discount factor than the reverse order). In the absence of multigame contact, cooperation is independent of playing the stage games sequentially or simultaneously, and thus there is no effect on the critical discount factors at which cooperation is sustainable in the *hard* and *easy game*. However, in the presence of multigame contact, the critical discount factor for joint cooperation in the *hard* and *easy game* is lower when played sequentially rather than simultaneously, and it surprisingly may be even lower than for the *easy game* alone.

The reason is twofold. First, when games are played sequentially, a partner who defects receives the temptation payoff in at most one game, and thus she gains a lower benefit

[6]As discussed in the introduction, Spagnolo (1999b) and Thomas and Willig (2006) reach similar conclusions in settings with imperfect information.

from deviation than in simultaneous play. Second, the punishment is weakly harsher: it is harsher if she deviates in the *hard game* (as then she will see her punishment start immediately within the same time period, i.e., in the second stage *easy game*), and it is the same if she deviates in the *easy game* (as—like in the simultaneous case—the punishment in both games then starts in the next time period). These considerations alter the incentive constraint and make cooperation in both games sustainable for

$$\delta \geq \begin{cases} \frac{T+z+P-2R}{T+z-P}, & \text{if} \quad 2z+P-R \geq 0, \\ \\ \frac{T-z-R}{T-z+R-2P}, & \text{otherwise.} \end{cases} \tag{2.4}$$

The top case refers to the optimal deviation taking place in the *hard game* and the bottom case in the *easy game*. If the games are not too asymmetric (i.e., z is not too large), this critical discount factor for cooperation in both games is lower than the critical discount factor for cooperation in the *easy game* alone.

Thus, with multigame contact, the consequences of linkage can be fundamentally different with simultaneous and sequential play. With simultaneous play, linkage may be beneficial for δ sufficiently high but it may need to be avoided for δ sufficiently low. On the other hand, with sequential play, if games are not too asymmetric, there is no region where linkage should be avoided.

2.2.2 Experimental predictions

To test the effect of multigame contact, we use the concrete parameterization of the stage game payoffs seen in Figure 2.4. Numbers represent monetary payoffs in ECU (experimental currency unit). We derive our predictions under the assumption that utilities are a linear transformation of a player's own monetary payoffs.

In our first experiment, we consider the simultaneous variant. If we compute the different critical discount factors discussed in Section 2.2.1, we obtain the following values when the games are played simultaneously: $\delta_{easy} = 0.11$, $\delta_{pool} = 0.38$, and $\delta_{hard} = 0.52$. We conduct sessions with three different continuation probabilities: $\delta \in \{0.1, 0.5, 0.9\}$. For $\delta = 0.1$, the theory predicts no cooperation, both with and without multigame contact; for $\delta = 0.9$ cooperation in both games can be part of an equilibrium both with and without multigame contact. For $\delta = 0.5$, the beneficial effects of multigame contact should crop up: whereas cooperation in the *easy game* is sustainable in both treatments, cooperation in the *hard game* is only possible with multigame contact.

Hard game

	C	D
C	$135, 135$	$45, 216$
D	$216, 45$	$60, 60$

Easy game

	c	d
c	$135, 135$	$45, 144$
d	$144, 45$	$60, 60$

Figure 2.4: Payoff matrices of the experimental stage games

In our second experiment, we consider the sequential variant. As presented in Section 2.2.1, with our particular payoffs, the optimal defection with multigame contact takes place in the second stage, i.e., in the *easy game* (as $135+144 > 216+60$). This lowers the critical discount factor for cooperation in both games from $\delta_{pool} = 0.38$ to $\delta_{pool} = 0.06$.[7]. All other critical discount factors are unaffected by moving from simultaneous to sequential play. This time, we conduct sessions with a single continuation probability: $\delta = 0.5$. The theoretical predictions based on subgame perfectness are therefore the same as before: whereas cooperation in the *easy game* is sustainable in both treatments, cooperation in the *hard game* is only possible with multigame contact. However, in single repeated prisoner's dilemma experiments, not only subgame perfectness matters for cooperation, but also the distance from this critical discount factor to the game's discount factor seems important (Dal Bó & Fréchette, 2018). We expect similar effects to be at work in a multigame contact setting, and thus to find a more significant effect of multigame contact on full cooperation in the sequential relative to the simultaneous case.

The critical discount factors presented above were derived under the assumption that other subjects are following cooperative grim trigger strategies, possibly with linkage. However, subjects cannot be certain that the other players will use such strategies, and the experimental literature based on single repeated games has shown that such strategic risk matters. Following Blonski and Spagnolo (2015), the literature has adapted the concept of risk dominance of Harsanyi and Selten (1988) by controlling for the critical discount factor at which grim becomes a best response when the other player randomizes with equal probabilities between the strategies of grim and always defect. Experimentally, this more stringent requirement seems to matter more for cooperation than subgame perfectness, and the distance from this critical discount factor to the game's discount factor remains important (Dal Bó & Fréchette, 2018).

Our choice of parameters was such that for each of the three different continuation probabilities $\delta \in \{0.1, 0.5, 0.9\}$, the qualitative predictions remain the same if we consider this notion of strategic risk rather than subgame perfection. Indeed, if we consider the *hard* and *easy game* alone, we obtain 0.56 and 0.24 as these critical discount factors, respectively. With multigame contact, this critical discount factor for full cooperation is 0.44 when the games are played simultaneously, and it is reduced to 0.12 when the games are played sequentially (since strategic uncertainty is fully resolved after the very first stage, i.e., the *hard game*).[8]

2.3 Experimental procedures

Subjects play a sequence of indefinitely repeated prisoner's dilemmas. In every round, each subject plays two prisoner's dilemmas in parallel. Henceforth, we refer to the experimental implementation of the *hard game* as *hard* and the *easy game* as *easy*.[9] We call the combination of both indefinitely repeated prisoner's dilemmas a supergame. The first three rounds of a supergame are played for certain and, at the end of the third round,

[7]Playing the *hard game* first leads to a much lower critical discount factor than playing the *easy game* first under multigame contact (0.06 vs. 0.35). Despite the theoretical argument in favor of our sequence, we cannot rule out the notion that the opposite sequence could have a larger impact on cooperation due to gradualism (see, e.g., Kartal et al., 2021)

[8]The idea that sequential play can reduce strategic uncertainty and thus foster cooperation has been experimentally tested in Ghidoni and Suetens (2020), although in their case the subgame-perfect critical discount factor remains unaffected.

[9]For half of the subjects, *hard* is always displayed on the left of the screen and *easy* on the right. The order is reversed for the other half. We use neutral labels (A, B and X, Y) for the actions in the games. We randomize by subject whether they see A, B or X, Y as labels for *hard* or *easy*. See Appendix 2.A.3 (p. 76) for the experimental instructions and screenshots.

a computerized stopping rule is introduced. From round three onward, the supergame either proceeds to the next round with continuation probability δ, or it stops and subjects move to a new supergame. After the termination of a supergame, subjects are randomly rematched for the subsequent supergame. All of this information is common knowledge.

The main treatment variation manipulates multigame contact: within a supergame, subjects either interact in both games with one partner (multigame contact, henceforth *1Partner*) or they play *hard* with one partner and *easy* with another partner (no multigame contact, henceforth *2Partner*).[10] The matching in a supergame is fixed. The second treatment variation is the expected length of the supergames. We implement three different continuation probabilities: $\delta \in \{0.1, 0.5, 0.9\}$.

The initial guaranteed rounds of play enable us to observe how subjects deviate and react to deviations, which is especially helpful for low and intermediate continuation probabilities where longer supergames are rare. These guaranteed rounds do not affect the theoretical predictions in the parameter space considered in the experiment. While they could affect subjects' perceptions of the true continuation probability, this should equally affect the treatments with one and two partners.[11]

We run our two-by-three factorial design as a between-subjects design, i.e., subjects only play one of the six treatments. All subjects in a session play the same treatment. Subjects in a session are randomly allocated to matching groups, which remain fixed throughout the session. At the beginning of each supergame, the computer randomly matches subjects with one or two partners from their matching group depending on the treatment.[12]

In the second experiment, subjects play the two games in a sequence instead of simultaneously. The order of events within a round is as follows: subjects take their decision simultaneously in *hard*, they are informed about the outcome in *hard*, they take their decision in *easy*, and are finally informed about the outcome in *easy*. For the sequential variant, we run only $\delta = 0.5$, the most interesting case in the first experiment. Whereas supergame durations were generated on the spot in the first experiment, we use the realizations of the six matching groups of the first experiment at $\delta = 0.5$ in *1Partner* for all treatments of the second experiment to maximize comparability between the two experiments. The rest of the experimental procedures is identical to the first experiment.[13]

Participants were paid out the sum of the payoffs of all rounds. During the experiment, we measured earnings in ECUs and the exchange rate was 1,000 ECUs = CHF 1 ($\approx$ USD 1.10). In addition, participants received a show-up fee of CHF 10 ($\approx$ USD 11). Sessions were run in the laboratory of the University of Lausanne (LABEX) with undergraduate students from the University of Lausanne and the EPFL recruited with ORSEE

[10]Labels do not allow subjects to identify with whom they interact. The other subject in the game is always labeled as "Your partner" in the treatments with one partner, and "Your partner 1" and "Your partner 2" in the treatments with two partners. At the beginning of each supergame, we inform subjects that a new partner or new partners 1 and 2 are randomly drawn.

[11]Several alternatives have been proposed in the literature to deal with these issues, including the block method (Fréchette & Yuksel, 2017). Our methodological choice was driven by the desire to keep the experimental instructions as simple as possible.

[12]See Appendix 2.A.2 (p. 81) for more details on the formation of matching groups and the stopping procedure. Table 2.A.5 in the appendix (p. 82) provides detailed information on matching groups' size and supergames' duration.

[13]Changes to the computer screens were kept minimal. They have the same structure as in the first experiment, but at the beginning of the round the part of the screen for *easy* is shaded and inactive. It becomes active when subjects have to take their decision in *easy* and the other part of the screen keeps displaying the results in *hard*.

(Greiner, 2015). The experiment was programmed in oTree (Chen et al., 2016). We ran a pilot in May 2020; the sessions of the first experiment took place in September and October 2020, while the sessions of the second experiment took place in May 2021.[14]

A total of 436 and 128 subjects participated in the first and second experiments, respectively. We end up with six matching groups per treatment, except at the continuation probability $\delta = 0.9$, for which we have data from five matching groups in *1Partner* and five matching groups in *2Partner*. Table 2.A.2 in the appendix (p. 76) provides detailed information about the observations per treatment. The average payment per participant was CHF 31 ($\approx$ USD 34) and sessions lasted between 70 and 113 minutes. The designs of both experiments and the hypotheses were pre-registered prior to data collection in two OSF registries.[15]

2.4 Results

This section is organized in four parts. In Section 2.4.1, we investigate the effect of multigame contact on cooperation under simultaneous play in our first experiment. In Sections 2.4.2 to 2.4.3, we perform exploratory analyses on first-round behavior and strategic linkage. Finally in Section 2.4.4, we present the results from our second experiment in which the effect of multigame contact on cooperation should be particularly strong according to theory.

2.4.1 Cooperation across treatments

The left panel of Figure 2.5 shows the mean cooperation rates with 95 percent confidence intervals in *hard* and *easy* for each treatment across all rounds. The cooperation rate in a game refers to the proportion of participants who choose to cooperate.

In line with the literature (Dal Bó & Fréchette, 2018), we observe an increase in cooperation rates for both games and for both partner treatments as the continuation probability increases. Contrary to the theoretical predictions at $\delta = 0.5$, we do not observe a higher cooperation rate in *1Partner* than *2Partner* for *hard* ($p = .394$).[16]

Although the theory does not predict differences in any of the other comparisons, it is clear that we may not want to take these predictions too literally. After all, we observe quite some cooperation at $\delta = 0.1$, as well as defection in $\delta = 0.9$. Given this variation, it seems reasonable to expect that the general mechanism of multigame contact should also affect behavior in the remaining comparisons.[17] For all other comparisons between *1Partner* and *2Partner*, we find no consistent difference in the average cooperation rates ($p > 0.589$). This also holds for overall cooperation (pooling the actions of *hard* and *easy*, $p > .588$ at any δ). We do not find interesting dynamic effects regarding the effect

[14] Data from the pilot were excluded from the analysis.

[15] First experiment: https://osf.io/u7hwe, second experiment (after analysis of the first experiment but prior to data collection of the second one): https://osf.io/6qcjt

[16] Unless specified, we always report the exact $p-$values from Wilcoxon rank-sum tests on matching group averages. The number of independent observations for each test comparing *1Partner* and *2Partner* is twelve for each δ, except $\delta = 0.9$, where we observe ten matching groups.

[17] Bruttel (2009) and Dal Bó and Fréchette (2018) argue that—rather than a stepwise increase—the distance between the implemented discount factor and the critical discount factor (δ^*) is a continuous predictor of cooperation. For discount factors below δ^*, the cooperation rates are typically at a low level, whereas above δ^* cooperation gradually increases in δ.

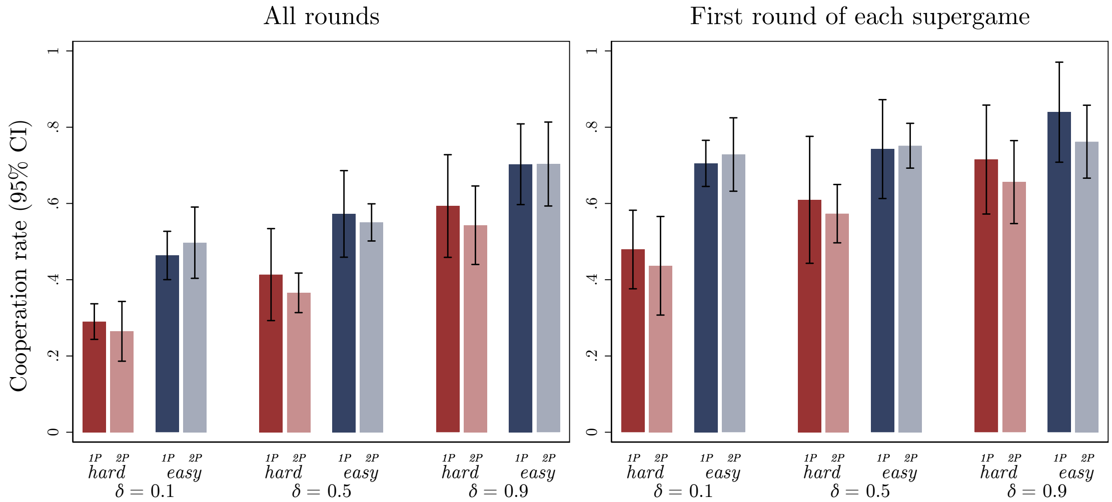


Figure 2.5: Cooperation rates by treatment. Mean cooperation rates and 95% confidence intervals are computed using the matching group averages. The left panel shows the results for all rounds; the right panel shows the first round in each supergame.

of multigame contact in any of the treatments. Figure 2.A.1 in the appendix (p. 77) shows the cooperation rates over time. This leads to our first result:

Result 1: Multigame contact does not increase average cooperation.
Cooperation rates in both hard and easy are statistically indistinguishable between 1Partner and 2Partner, and the same holds for overall cooperation.

However, average cooperation rates in *hard* and *easy* are only one way to investigate potential differences between treatments. A first indication for a treatment difference may manifest itself—at least at $\delta = 0.5$—in the confidence intervals shown in Figure 2.5, which are substantially larger in *1Partner*.

In theory, multigame contact should foster cooperation in *hard*. More specifically, simultaneous cooperation in *hard* and *easy* should be more prevalent in *1Partner* than *2Partner*. Although cooperation rates in *hard* do not differ between *1Partner* and *2Partner*, simultaneous cooperation in both games may still be more prevalent under multigame contact if simultaneous defection in both games is more prevalent as well. After all, defecting in both games is the reaction we expect after any deviation under multigame contact.

To check whether multigame contact leads more often to situations of either simultaneous cooperation or defection in both games, we look at the outcome of the stage games. We are interested in whether or not subjects reach a cooperative outcome in the stage games. A cooperative outcome in *hard* (*easy*) refers to a situation in which both the subject and her partner in *hard* (*easy*) cooperate. The game outcomes are considered from the perspective of each subject. Whereas game outcomes are identical from the perspective of a subject and her partner in *1Partner*, this is not necessarily the case in *2Partner*. Figure 2.6 shows the outcome of the stage games by treatment. For each continuation probability and for both *1Partner* and *2Partner*, we report the relative frequencies of four outcomes: full defection occurs when a subject defects in both games and also face defection in both games; partial cooperation occurs when a subject reaches a cooperative

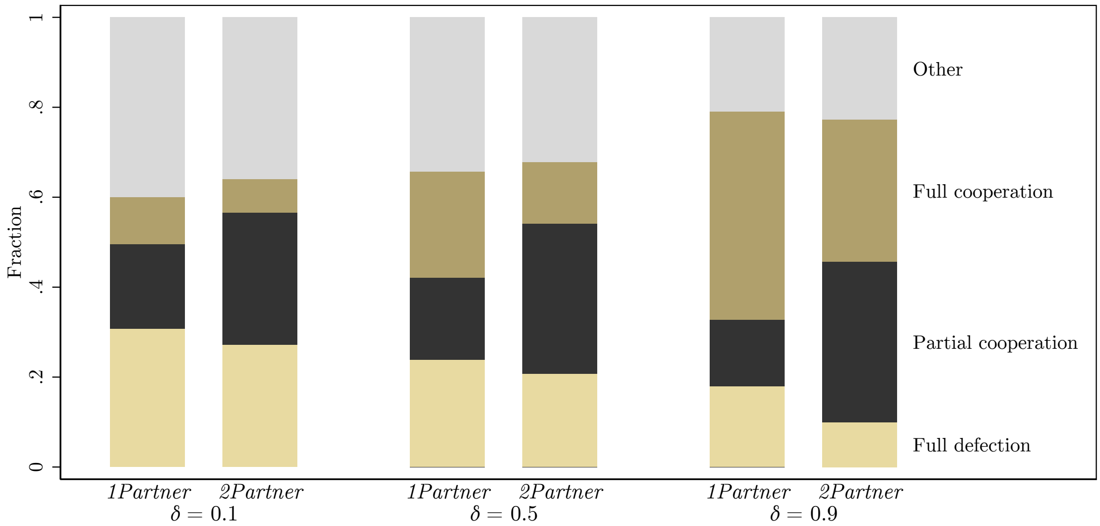


Figure 2.6: Outcome of the stage games by treatment for all rounds. Relative frequencies of four categories of outcomes. Full defection: a subject defects in both games and faces defection in both. Partial cooperation: cooperation of a subject and her partner for the game only occurs in one of the two games (*hard* or *easy*). Full cooperation: a subject cooperates in both games and faces cooperation in both. Other contains the remaining cases.

outcome in only one of the two games; full cooperation occurs when a subject reaches a cooperative outcome in both games; the remaining category (other) contains cases in which no cooperative outcome was reached despite some cooperative actions. This last category contains only uncoordinated decisions and presumably reflects transitory states towards cooperation or defection.[18]

At any continuation probability, the two extreme outcomes (full defection and full cooperation) are more frequent in *1Partner* than *2Partner*. Conversely, partial cooperation occurs approximately twice as often in *2Partner* than *1Partner*. To compare the distribution of outcomes between *1Partner* and *2Partner*, we use Rao-Scott χ^2-tests, which correct for dependence within matching group (henceforth RS-test, Rao & Scott, 1984). Leaving aside the other category, which occurs with similar frequencies across partner treatments, we find statistically significant differences between *1Partner* and *2Partner* at $\delta = 0.5$ and $\delta = 0.9$ ($p < 0.041$) but no at $\delta = 0.1$ ($p = 0.213$).

The higher frequencies of extreme outcomes, at the cost of cooperation in a single game, suggest that multigame contact is a double-edged sword; namely, a blessing for some and a curse for others. Because we did not pre-register the hypothesis that multigame contact leads to more extreme outcomes, we will treat this as a tentative result at this point. Our second experiment (Section 2.4.4) will replicate these results.

To investigate the effect of multigame contact in more detail, we now take a closer

[18]This would not be true if we also considered cooperative strategies in which players do not simultaneously cooperate in the same game. For example, subjects could take turns getting the temptation payoff in *hard* (216) and/or in *easy* (144). Such strategies lead to a lower average payoff than the kind of strategies we consider in Section 2.2, but they can still form a subgame-perfect Nash equilibrium for sufficiently high continuation probabilities (see Stahl, 1991, for a characterization of subgame-perfect strategies in infinitely repeated prisoner's dilemma). Note that alternating is more attractive in *hard* than *easy* (average payoffs per round of 130.5 and 94.5, respectively) and results in an average payoff per round close to simultaneous cooperation (130.5 vs. 135). We checked for evidence of such strategies in our dataset and—similarly to Dal Bó (2005) and Fréchette and Yuksel (2017)—we find little evidence of subjects consistently alternating.

look at individual decisions. We separate our analysis into two parts: in Section 2.4.2, we analyze decisions in the first round of each supergame and in Section 2.4.3 we use behavior from round two onwards to investigate whether subjects link the two games when playing with the same partner.

2.4.2 First-round behavior

Looking at the first round of each supergame is especially interesting since it allows us to observe subjects' behavior before it has been influenced by the partner's or partners' decisions.[19] Cooperation rates in the first round of each supergame are often used as a measure of subjects' intention to cooperate (Dal Bó & Fréchette, 2018, p.90). The result in Section 2.4.1 that multigame contact does not affect cooperation rates in both *easy* and *hard* also holds when restricting to the first round in each supergame (see right panel of Figure 2.5). Indeed, all comparisons between *1Partner* and *2Partner* are neither large nor statistically significant ($p > .420$ at any δ).

How do subjects enter into a new supergame? As there are two stage games in which subjects can either cooperate or defect, there are four possible decision pairs in each round. We denote a decision pair by a capital letter for the decision in *hard* and a lowercase letter for the decision in *easy*. For example, Dc means that a subject defects in *hard* and cooperates in *easy*. In what follows, we pool Cd and Dc as it is difficult to rationalize wanting to cooperate in *hard* but not in *easy* in the first round of a supergame, especially for *1Partner*. According to our theoretical predictions, we expect cooperation in none of the games (Dd) for $\delta = 0.1$ and full cooperation (Cc) for $\delta = 0.9$ in both *1Partner* and *2Partner*. At $\delta = 0.5$, cooperation is only sustainable in *easy* (Dc) in *2Partner*, whereas full cooperation (Cc) is sustainable in *1Partner*.

Table 2.1: Subjects' decision pair, first round of each supergame

	$\delta = 0.1$		$\delta = 0.5$		$\delta = 0.9$	
	1Part.	*2Part.*	*1Part.*	*2Part.*	*1Part.*	*2Part.*
Cc	0.45	0.42	0.60	0.54	0.70	0.64
Cd/Dc	0.29	0.33	0.16	0.25	0.15	0.14
Dd	0.26	0.25	0.24	0.21	0.15	0.22
N	1,970	3,770	1,488	2,812	330	500

Notes: Relative frequencies of subjects' decision pair in the first round of each supergame. The first letter (capital) of each pair refers to the decision in *hard* and the second (lowercase) to the decision in *easy*. RS-tests for the difference between *1Partner* and *2Partner* yield $p = .766$ at $\delta = 0.1$, $p = .457$ at $\delta = 0.5$, and $p = .504$ at $\delta = 0.9$.

Table 2.1 presents the relative frequencies of the decision pairs in the first round of each supergame. While none of the comparisons in first-round decisions reaches significance (see notes of Table 2.1), a few observations are worth mentioning. Cc is the modal

[19] Admittedly, this is strictly only true in the first round of the first supergame. Dal Bó and Fréchette (2018) report that cooperation rates in the first round of later supergames are influenced by both the realized duration of the previous supergames and the choices of the past subjects with whom one interacted. However, looking only at the very first round played does not allow us to observe experienced subjects. Dal Bó and Fréchette (2018) also find that letting subjects gain experience with the environment is often required for a treatment effect to appear. Since our main interest is the difference between *1Partner* and *2Partner*, we will consider the first round of all supergames.

decision pair in every treatment and always more frequent in *1Partner*. This suggests that many subjects try to instigate full cooperation, especially in *1Partner*. At $\delta = 0.1$ and $\delta = 0.5$, Cd/Dc occurs more often in *2Partner* than in *1Partner*. This is in line with our theoretical prediction for $\delta = 0.5$. Finally, with the exception of Dd at $\delta = 0.9$, extreme decision pairs (Cc and Dd) are always more frequent in *1Partner* than *2Partner*.

To conclude, Result 1 also holds when only considering the first round of each supergame. On the other hand, the two extreme decision pairs (Cc and Dd) are more frequent in *1Partner* than *2Partner* in almost all cases, which is consistent with the double-edged sword we observed in Section 2.4.1. Assuming that a subject links her decisions in the two games and/or believes that her partner does so under multigame contact, we would draw the conclusion that the two possible paths forward would be either full cooperation or full defection. In the first round of a supergame, this can only work through anticipation, which may be one reason why the treatment differences remain statistically weak. In order to investigate whether subjects indeed link the two games under multigame contact, we now study subjects' transitions between cooperation and defection.

2.4.3 Strategic linkage

The prediction that multigame contact helps to sustain simultaneous cooperation in both games relies on the assumption that players link the two games when matched with a single partner. If this link holds, then we should observe different reactions to the partner's or partners' previous decisions between *1Partner* and *2Partner*. For the following analysis we restrict our attention to situations where theory predicts different reactions. Suppose that a subject cooperates in both games in round $t-1$ but reaches a cooperative outcome only in one of the two games. For *2Partner* theory predicts that the subject will continue cooperating in the one game and defect in the other. In contrast, the multigame environment (*1Partner*) calls for deviation in both games (linkage).

Figure 2.7 shows heatplots with the decisions of a subject's partner(s) in round $t-1$ of a supergame on the horizontal axis and the subjects' reaction in t on the vertical axis. We restrict our attention to the cases in which the subject in question cooperated in both games in round $t-1$. The figure shows the results for the treatments with $\delta = 0.5$, for the results with the other two continuation probabilities see Figure 2.A.2 in the appendix (p. 78). We consider three situations depending on whether the partner(s) cooperated in *hard* (capital letter) and *easy* (lowercase letter). In the left panel (*1Partner*), these decisions refer to the subject's partner, whereas in the right panel (*2Partner*) the first letter is the decision of the partner in *hard* and the second is the decision of the partner in *easy*. Each cell reports the mean proportion of the corresponding reaction (in percent) conditional on the decisions of the partner(s). Cells within a column add up to 100 percent. The relative frequencies of the partner's or partners' decisions are given in parentheses below the labels.

The left and right columns in each panel document how subjects react to either full cooperation (Cc) or full defection (Dd). Observing the partner(s) choosing Cc or Dd are both more likely in *1Partner* than *2Partner* (63 vs. 47 percent and 21 vs. 13 percent, $p = .009$, RS-test). After Cc and Dd the modal reaction is to respond in kind, independent of the treatment.

To investigate whether subjects in multigame contact link the two games the middle column of each panel is most informative. This column shows how a subject having played Cc reacts to Cd or Dc, namely cooperation in one game but not in the other.[20]

[20]For ease of exposition we pool Cd and Dc. The qualitative results do not change if we treat these two

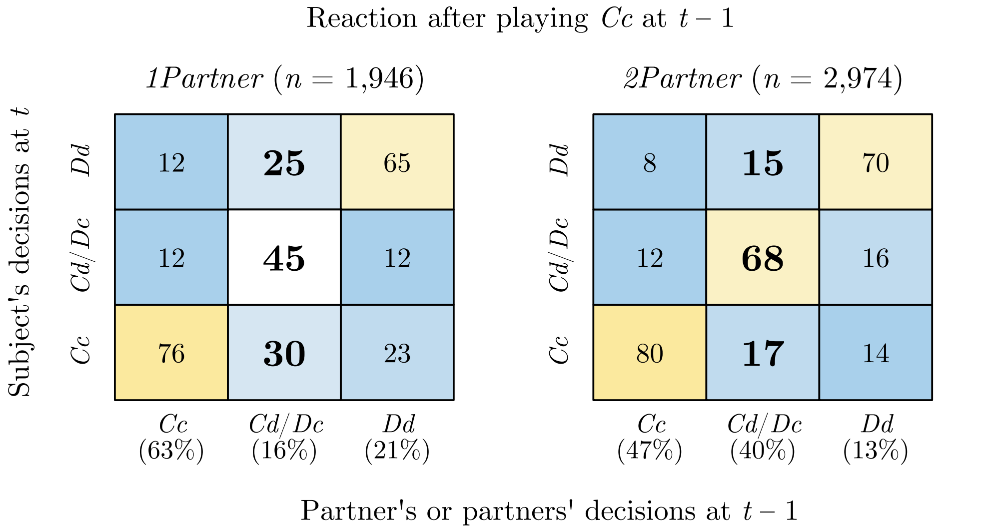


Figure 2.7: Reactions to observed previous-round decisions ($\delta = 0.5$ only). Numbers in the cells show the subject's decision at t in percentage within each column; bold numbers indicate significant differences for a given column between *1Partner* and *2Partner* (RS-tests); numbers below the labels show the frequencies of the partner's or partners' decisions. The coloring indicates frequencies, going from blue for values close to 0 up to gold for values close to 100 percent.

Observing uncoordinated decisions from the partner(s) is much less likely in *1Partner* than *2Partner* (16 vs. 40 percent). Subjects' reactions to Cd/Dc are significantly different between *1Partner* and *2Partner* ($p = .002$, RS-test, indicated by the bold numbers in Figure 2.7). Although the modal reaction is Cd/Dc in both treatments, it happens less often in *1Partner* than *2Partner* (45 vs. 68 percent). Conversely, reacting with either Cc or Dd are both almost twice as likely in *1Partner* than *2Partner* (25 vs. 15 percent for Cc and 30 vs. 17 percent for Dd).

The treatments with $\delta = 0.1$ and $\delta = 0.9$ qualitatively confirm these patterns (see Figure 2.A.2 in the appendix, p. 78). Compared to *2Partner*, subjects in *1Partner* are more likely to deviate with full defection if they face Cc, and subjects are more likely to fully defect when facing partial defection.

Linkage does not only imply reacting with full defection following any deviation by the partner(s). By backward induction, linkage is also the optimal deviation under multigame contact and what a rational player would choose if she expects a deviation by the partner in any game in the following round. To have a comprehensive measure for linkage we should therefore consider all transitions out of Cc. In particular, we analyze how likely it is for a subject to move from Cc to Dd (as opposed to Cd or Dc) and compare these frequencies across *1Partner* and *2Partner* to identify the strength of linkage. Table 2.2 shows the transition matrices for the decisions from round $t-1$ to t at $\delta = 0.5$ for *1Partner* and *2Partner* (see Tables 2.A.3 and 2.A.4 in the appendix, p. 79, for the other continuation probabilities). For example, subjects who played Cc in the previous round play Dd in the current round in 25 percent of the cases and Cd/Dc in 17 percent of the cases under multigame contact. Conversely, in *2Partner* subjects move to Dd in 18 percent and Cd/Dc in 35 percent of the cases.

cases separately.

Our measure for the strength of linkage compares these relative frequencies across treatments. In particular, we ask how much more likely it is that a subject who stops cooperating chooses full defection (Dd). We can calculate this measure for linkage by the ratio of the relative frequencies in the two treatments. For $\delta = 0.5$ we get $r = (25/17)/(18/35) = 2.86$, 95% CI: $[1.47, 5.03]$.[21] This means it is close to three times more likely that a subject who stops cooperating moves to full defection (rather than partial defection) under multigame contact compared to interactions in single games. Also for the other two continuation probabilities we find very consistent and highly significant results (for the transition matrices see Tables 2.A.3 and 2.A.4 in the appendix, p. 79). For $\delta = 0.1$ we get $r = 2.85$, 95% CI $[1.64, 5.40]$; for $\delta = 0.9$ we get $r = 3.56$, 95% CI $[1.81, 6.06]$.

Table 2.2: Transition matrices $t-1$ to t, $\delta = 0.5$

		1Partner			*2Partner*		
	t	Cc	Cd/Dc	Dd	Cc	Cd/Dc	Dd
	Cc	0.58	0.17	0.25	0.47	0.35	0.18
$t-1$	Cd/Dc	0.18	0.46	0.36	0.13	0.55	0.32
	Dd	0.08	0.15	0.77	0.06	0.15	0.79

Notes: Transition matrices of the subjects' decisions from round $t-1$ to t at $\delta = 0.5$. The first letter (capital) of each pair refers to the decision in *hard* and the second (lowercase) to the decision in *easy*. Each cell shows the relative frequency of the column decision pair at t given the row decision pair at $t-1$. Each row within a matrix adds up to one.

To summarize, we find no support for our hypothesis that multigame contact increases cooperation in *hard*, but we find strong evidence of linkage in *1Partner* at all three continuation probabilities. In line with Bernheim and Whinston (1990), this often enables subjects to reach fully cooperative outcomes. However, linkage also frequently leads subjects towards fully defective situations, which prevents multigame contact from producing overall benefits. One reason why multigame contact did not increase overall cooperation in our experiment may be that the shift in the critical discount factor was not strong enough in our design. In order to give the theory the best shot we introduce a small change in the sequencing of events, which has large effects on the critical discount factor under multigame contact. We examine this new environment in our second experiment.

2.4.4 Powering multigame contact through sequential play

In this second experiment we explore a new design: instead of playing the two games simultaneously, subjects play the two games in a sequence. First, both subjects take their decision in *hard*, after which they learn the outcome of *hard* and proceed to *easy*. Recall from the theory outlined in Section 2.2 that playing the two games sequentially rather than simultaneously within a round has no impact on the sustainability of cooperation in *2Partner*. In *1Partner*, the sequential variant of the game lowers the critical discount factor necessary to sustain cooperation in both games substantially from 0.38 to 0.06.

The left panel of Figure 2.8 shows mean cooperation rates with 95% confidence intervals for all the rounds. Contrary to the simultaneous case, the cooperation rate in *hard*

[21] We use a bootstrap with clustering on matching group (1000 repetitions) to estimate the 95% confidence interval of $\ln(r)$. The confidence interval in the main text shows the exponentiated form.

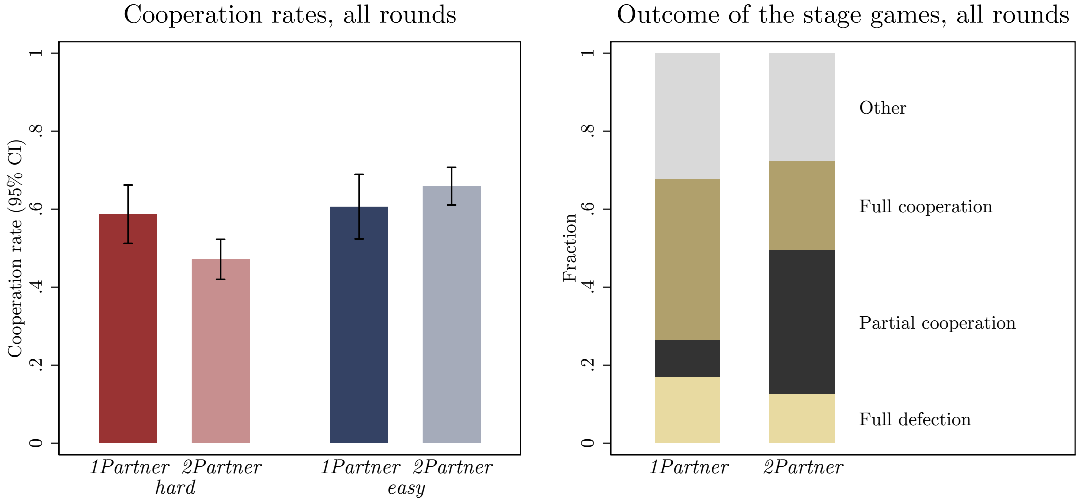


Figure 2.8: Left panel: cooperation rates by treatment; mean cooperation rates and 95% confidence intervals for all the rounds. Right panel: outcome of the stage games by treatment for all the rounds; relative frequencies of four categories of outcomes. See the captions of Figures 2.5 and 2.6.

is now significantly higher in *1Partner* than *2Partner* (0.59 vs. 0.47, $p = .026$) and it is almost indistinguishable from the cooperation rate in *easy* (*1Partner*, 0.59 and 0.61, $p = .937$). On the other hand, the cooperation rate in *easy* is now slightly lower in *1Partner* than *2Partner* (0.61 vs. 0.67, $p = .310$). In all situations, cooperation rates tend to increase when moving from simultaneous to sequential games.[22] From these results, we conclude that the largest effect of moving from simultaneous to sequential games is—as expected—the increase in cooperation in *hard* for *1Partner*. Nevertheless, because cooperation rates have also increased in *2Partner*, we do still not find significant improvements in overall cooperation with multigame contact (0.60 vs. 0.57, $p = .589$).

The lack of clear beneficial effect from multigame contact is surprising as the implemented continuation probability is significantly higher than the critical discount factor required to sustain cooperation in *hard* for *1Partner*.

The results of the second experiment clearly confirm the double-edged sword. The right panel of Figure 2.8 demonstrates the move towards extreme outcomes in *1Partner*. In this case, full cooperation occurs twice as often in *1Partner* than *2Partner* and full defection is almost 50 percent more likely in the former than the latter. Conversely, a cooperative outcome in only one of the two games is four times more frequent in *2Partner* than *1Partner*. The difference in the distribution of outcomes is highly significant (RS-test, $p = .000$) and leads us to our second result:

Result 2: Multigame contact is a double-edged sword.
Subjects in 1Partner are significantly more likely to realize cooperative outcomes in both or neither game, whereas subjects in 2Partner are more likely to realize partially cooperative outcomes. On average, multigame contact fails to increase cooperation.

In Section 2.4.3, we evidenced linkage by showing that subjects in *1Partner* react more often with a coordinated decision in *hard* and *easy* following defection in one game

[22] *1Partner*: from 0.42 to 0.59 in *hard* ($p = .041$) and from 0.57 to 0.61 in *easy* ($p = .589$). *2Partner*: from 0.37 to 0.47 in *hard* ($p = .026$) and from 0.55 to 0.67 in *easy* ($p = .026$).

by the partner than subjects in *2Partner*. We can perform a similar analysis on the data of the second experiment. Again, we expect that a deviation by the partner in one game will trigger defection in both games under multigame contact. When games are played sequentially, a deviation in the first game (*hard*) should trigger in reaction in the same round.

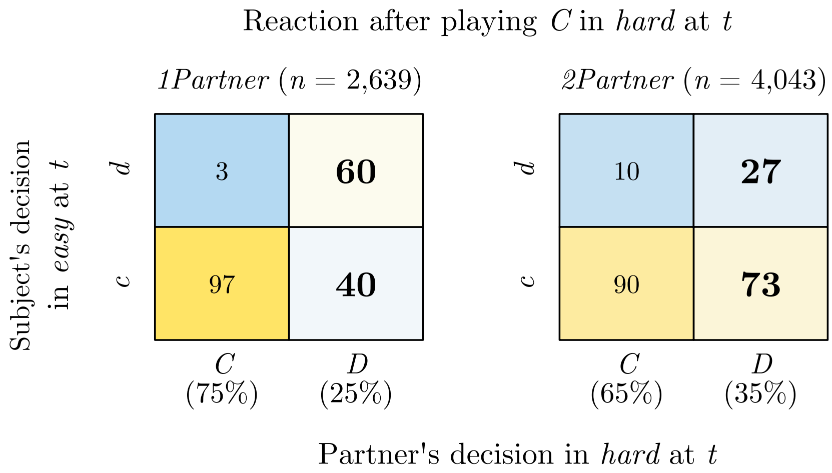


Figure 2.9: Reactions to observed decision in *hard* at t. Numbers in the cells show the subject's decision at t in percentage within each column; bold numbers indicate significant differences for a given column between *1Partner* and *2Partner* (RS-tests); numbers below the labels show the frequencies of the partner's or partners' decisions. The coloring indicates frequencies, going from blue for values close to 0 up to gold for values close to 100 percent.

Figure 2.9 shows how a subject having cooperated in *hard* (C) reacts in *easy* (c or d) conditional on the partner's decision in *hard* for a given round t. Cells within a column add up to 100 percent and the relative frequencies of the partner's decision are given in parentheses below the labels. In most cases subjects cooperate in *easy* when observing cooperation in *hard* (left columns). Although the point estimate is considerably larger in *1Partner* (97 percent vs. 90 percent), there is no statistically significant difference between the two treatments. However, we observe a large and statistically significant difference between *1Partner* and *2Partner* in the reaction after facing defection in *hard* (indicated by bold numbers, RS-test, $p = .000$). Whereas the modal reaction is to defect in *1Partner* (60 percent), the modal reaction in *2Partner* is to cooperate (73 percent).[23]

In a next step, we will provide a more general test for linkage. In our previous analysis, we did not control for possible time effects within and across supergames. Not controlling for such effects may explain why we observe different reactions in the right panel of Figure 2.9 when we do not expect them if subjects in *2Partner* treat the two games independently.

In a broad sense, linkage between the two games can be understood as mutual dependence of the decisions between the two games. We expect to observe linkage in *1Partner*, meaning that the decisions between *hard* and *easy* should be linked. Whereas cooperating

[23]Instead of focusing on the reaction to decisions in *hard* in the same round we could also look at the reaction in *easy* to decisions in *hard* in the previous round. Figure 2.A.3 in the appendix (p. 80) shows that the results are very similar, with subjects in *1Partner* being significantly more likely to defect in the subsequent game.

only in *easy* is possible, cooperating only in *hard* under multigame contact is difficult to rationalize. It follows that for a given round, a subject that reaches a cooperative outcome in *hard* should then cooperate in *easy*. In *2Partner*, there is no obvious reason to link the two games and we cannot explain the decision in one game by the decision in the other game. Most common strategies for repeated games condition the decision in a given round on what happened in the game in the previous round. Therefore, the outcome of the game in the previous round is presumably the best predictor for the decision to cooperate without multigame contact. Indeed, cooperating is only attractive if the partner is also willing to cooperate. Consequently, we should be able to explain the decision in *easy* by the outcome in *hard* for *1Partner* whereas the previous *outcome* in *easy* should be the best predictor of cooperation in *easy* for *2Partner*. We put this intuitive conjecture to a test in a regression analysis.

Table 2.3: Linkage in the sequential games

	Dep. var.: cooperation in *easy* (c_t)		
	(1)	(2)	(3)
2Partner	0.052	0.073^{**}	0.010
	(0.044)	(0.018)	(0.037)
$(C,C)_t$ [cooperative outcome in *hard*]		0.205^{**}	0.519^{**}
		(0.046)	(0.046)
$(C,C)_t \times$ *2Partner*			-0.408^{**}
			(0.046)
$(c,c)_{t-1}$ [cooperative outcome in *easy*]		0.492^{**}	0.186^{**}
		(0.048)	(0.054)
$(c,c)_{t-1} \times$ *2Partner*			0.383^{**}
			(0.066)
Constant	0.487^{**}	0.140^{**}	0.196^{**}
	(0.057)	(0.023)	(0.028)
Time controls	Yes	Yes	Yes
χ^2-test	383.1	1574.0	2951.2
p	0.000	0.000	0.000
R^2	0.081	0.433	0.463
N	$13{,}076$	$9{,}796$	$9{,}796$

Notes: Random effects estimates. Dependent variable is cooperation in *easy*. Independent variables are a dummy for the treatments with two partners (with one partner as baseline case); $(C,C)_t$ indicates a cooperative outcome in *hard*; $(c,c)_{t-1}$ indicates a cooperative outcome in *easy* in the previous round of the supergame. Time controls are dummies for the first and second round of the supergame and the supergame round, as well as the overall round in the experiment. Robust standard errors, clustered on matching group, in parentheses. $^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$.

Table 2.3 shows the results of linear probability models. In all models, we report standard errors clustered on matching groups.[24] In Model (1), we regress cooperation in *easy* (c_t) on the exogenous treatment variable *2Partner* and the coefficient estimate confirms the lack of differences in the cooperation rate in *easy* between the partner treatments. In

[24] We treat the matching group as the level of independent observation. This approach does not account for static session effects (Fréchette, 2012). However, our inferential results are almost identical if we cluster at the session level.

all estimates, we control for time effects with dummies for the first and second round of the supergame, the supergame round and the overall round. In Model (2), we add two explanatory variables: $(C, C)_t$ indicates a cooperative outcome in *hard* in the same round and $(c, c)_{t-1}$ indicate a cooperative outcome in *easy* in the previous round. Both variables predict a significantly higher likelihood to cooperate in *easy*, with the effect of the outcome in *easy* being stronger than the link between *hard* and *easy*. To address the core question of this analysis we need to check whether these coefficients react to multigame contact. In Model (3) we interact $(C, C)_t$ and $(c, c)_{t-1}$ with *2Partner*, respectively. In *1Partner*, the reaction to a cooperative outcome in *hard* is positive and large (0.519). The highly significant and negative interaction term indicates that the link between *hard* and *easy* is smaller for *2Partner* ($0.519 - 0.408 = 0.111$, $p = .000$). Having reached a cooperative outcome in *easy* in the previous round has a positive effect on cooperation in *easy* in the following round, and this effect is much stronger in *2Partner*.

Our second experiment strongly confirms that subjects link the two games in the presence of multigame contact. Although multigame contact seems to help to establish cooperation in *hard*, overall the results do not suggest important improvements in terms of cooperation. These findings lead us to our final result:

Result 3: Under multigame contact, subjects strategically link the two games.
In the presence of multigame contact, defection of other subjects is more likely to provoke full defection in response. The link between cooperation in the hard game and cooperation in the easy game is stronger in the presence of multigame contact.

2.5 Conclusion

According to a well-established theoretical argument, interacting on multiple fronts should enable people to establish and maintain cooperation on all fronts more often. The idea is to link the games and use any existing slack in the incentive constraint of some games to enforce cooperation in the other.

Our experimental results suggest that things are more complicated in practice than theory would predict. In our main treatment, even when the discount factor is such that subjects should be able to cooperate in both games in the presence of multigame contact but only cooperate in the *easy game* in the absence of multigame contact, we find that on average subjects fail to reach any significant benefits from multigame contact. Our experimental findings led us to conclude that multigame contact is a double-edged sword.

Although our results shed some light on the beneficial and detrimental effects of linking the actions in one game to the outcome of the other, it remains a conundrum why the negative effects of linkage cannot be avoided. After all, if linking the two games leads to unfavorable outcomes, rational players should be able to unlink the two situations. In other words, everything that is possible with single-game contact should also be possible with multigame contact (and more).

The insight that subjects link even when they should not carries potential important implications in a multitude of settings. Besides obvious applications in industrial organization and labor economics, the issue of undesirable linkage may have ramifications for fields such as international relations and diplomacy. For instance, it has been shown that linking independent policy games should not harm international cooperation (Spagnolo, 2001). In reality, we see that policy issues are often kept separate and in particular the WTO even promotes the idea that countries should use sanctions against other countries in the same sector in which the violation took place. For example, the response to a violation in the area of patents should also relate to patents. Scholars have reacted to this principle by highlighting that there are strategic reasons for linking trade and environmental policies in multilateral negotiations and that global cooperation may be easier to sustain when pursued through linked negotiations (see, for example, Barrett, 1994, and Blackhurst and Subramanian, 1992). Our findings challenge this idea that linking different separable policy issues into an overarching international agreement is advantageous for cooperation. The relevance of all of these issues points to the need to create a body of work that leads to a better understanding not only of the benefits but also the potential costs of linkage.

2.A Appendix

2.A.1 Additional tables and figures

Table 2.A.1: Cooperation as Subgame-Perfect Equilibrium (SPE) and Risk-Dominant (RD) strategy

	Simultaneous			Sequential
Continuation prob.	$\delta = 0.1$	$\delta = 0.5$	$\delta = 0.9$	$\delta = 0.5$
- *1Partner*				
- *easy* game	-	SPE & RD	SPE & RD	SPE & RD
- *hard* game	-	SPE & RD	SPE & RD	SPE & RD
- *2Partner*				
- *easy* game	-	SPE & RD	SPE & RD	SPE & RD
- *hard* game	-	-	SPE & RD	-

Simultaneous: $\delta^{SPE}_{easy} = 0.11$, $\delta^{RD}_{easy} = 0.24$, $\delta^{SPE}_{pool} = 0.38$, $\delta^{RD}_{pool} = 0.44$, $\delta^{SPE}_{hard} = 0.52$, $\delta^{RD}_{hard} = 0.56$
Sequential: $\delta^{SPE}_{easy} = 0.11$, $\delta^{RD}_{easy} = 0.24$, $\delta^{SPE}_{pool} = 0.06$, $\delta^{RD}_{pool} = 0.12$, $\delta^{SPE}_{hard} = 0.52$, $\delta^{RD}_{hard} = 0.56$

Table 2.A.2: Summary of the sessions

	$\delta = 0.1$, Sim.		$\delta = 0.5$, Sim.		$\delta = 0.9$, Sim.		$\delta = 0.5$, Seq.	
	1Part.	*2Part.*	*1Part.*	*2Part.*	*1Part.*	*2Part.*	*1Part.*	*2Part.*
Sessions	3	6	3	6	2	3	3	6
Matching gr.	6	6	6	6	5	5	6	6
Subjects	60	116	58	114	34	54	44	84
Decisions	6,030	11,754	5,924	11,574	3,752	5,800	4,496	8,580

Notes: Number of sessions, matching groups, subjects, and decisions by treatment. We count one decision each time a subject has to take a decision in both games, i.e. this is equivalent to the total number of rounds played by all the subjects.

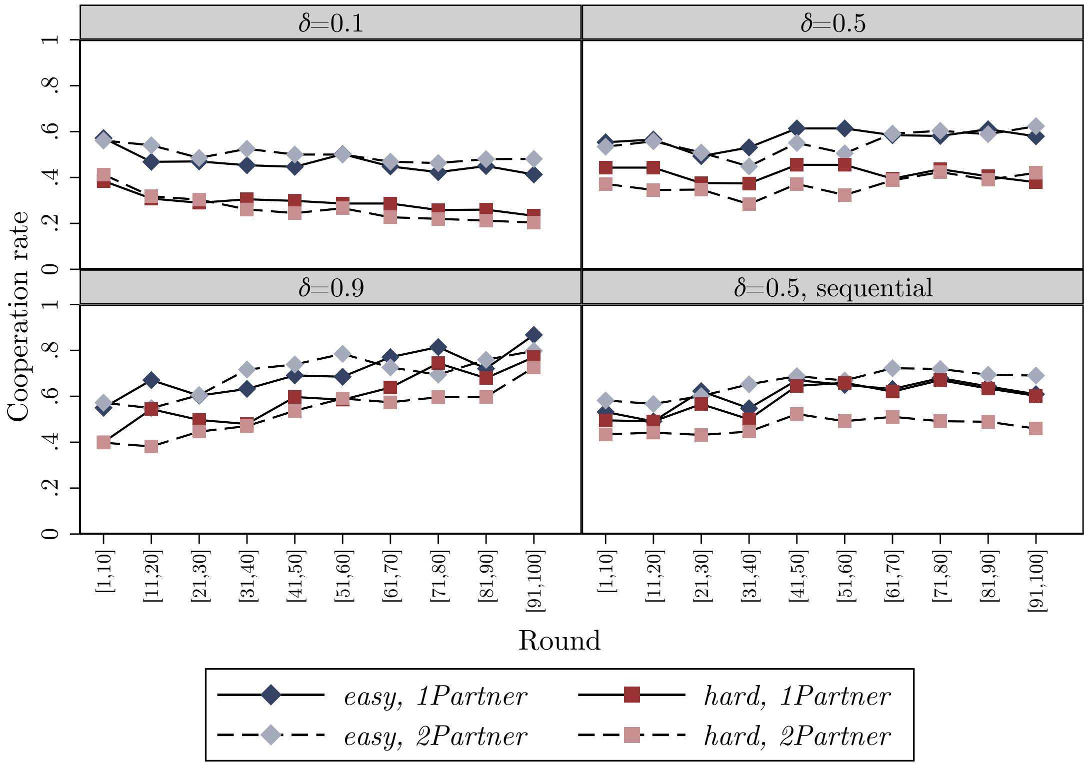


Figure 2.A.1: Cooperation rates in *hard* and *easy* over time by treatment.

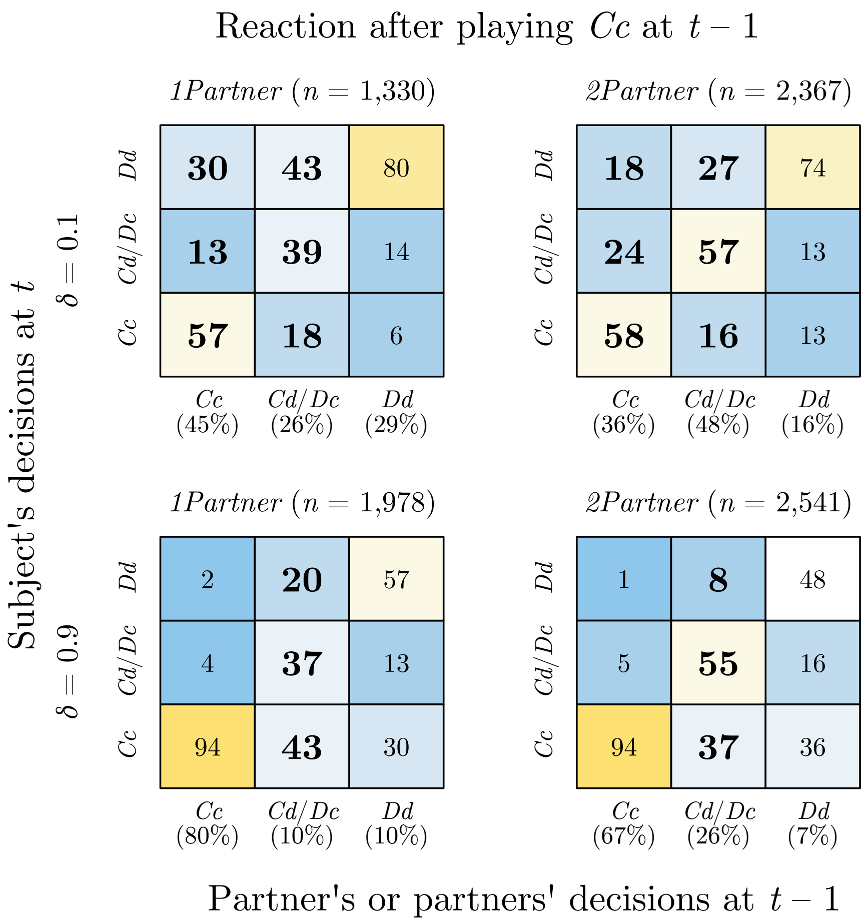


Figure 2.A.2: Reactions to observed previous round decisions for $\delta = 0.1$ (upper half), and $\delta = 0.9$ (lower half). Numbers show the percentage in each column; bold numbers indicate significant differences for a given column between *1Partner* and *2Partner* (RS-tests); numbers below the labels show the frequencies of the partner's or partners' decisions. The coloring indicates frequencies, going from blue for values close to 0 up to gold for values close to 100 percent.

Table 2.A.3: Transition matrices $t-1$ to t, $\delta = 0.1$

		1Partner			*2Partner*		
	t	Cc	Cd/Dc	Dd	Cc	Cd/Dc	Dd
	Cc	0.32	0.20	0.48	0.30	0.38	0.32
$t-1$	Cd/Dc	0.11	0.38	0.51	0.09	0.39	0.52
	Dd	0.05	0.15	0.80	0.05	0.12	0.83

Notes: Transition matrices of the subjects' decisions from round $t-1$ to t at $\delta = 0.1$. The first letter (capital) of each pair refers to the decision in *hard* and the second (lowercase) to the decision in *easy*. Each cell shows the relative frequency of the column decision pair at t given the row decision pair at $t-1$. Each row within a matrix adds up to one.

Table 2.A.4: Transition matrices $t-1$ to t, $\delta = 0.9$

		1Partner			*2Partner*		
	t	Cc	Cd/Dc	Dd	Cc	Cd/Dc	Dd
	Cc	0.83	0.08	0.09	0.75	0.19	0.06
$t-1$	Cd/Dc	0.31	0.42	0.27	0.25	0.60	0.15
	Dd	0.13	0.15	0.72	0.12	0.20	0.68

Notes: Transition matrices of the subjects' decisions from round $t-1$ to t at $\delta = 0.9$. The first letter (capital) of each pair refers to the decision in *hard* and the second (lowercase) to the decision in *easy*. Each cell shows the relative frequency of the column decision pair at t given the row decision pair at $t-1$. Each row within a matrix adds up to one.

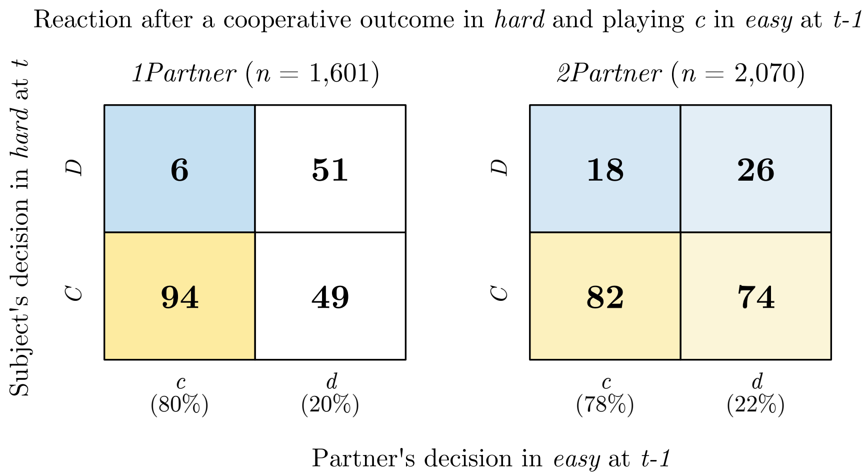


Figure 2.A.3: Reactions to observed decision in *easy* at $t-1$. Numbers in the cells show the subject's decision at t in percentage within each column; bold numbers indicate significant differences for a given column between *1Partner* and *2Partner* (RS-tests); numbers below the labels show the frequencies of the partner's or partners' decisions. The coloring indicates frequencies, going from blue for values close to 0 up to gold for values close to 100 percent.

2.A.2 Experimental procedures

2.A.2.1 Matching procedure

Subjects in a session are randomly allocated to matching groups and only interact with the other subjects in their matching group. All matching groups (and therefore all subjects) in a session play the same treatment. We vary matching group sizes (6 to 20 subjects) across treatments to keep the number of times a subject interacts with another subject in her matching group comparable. Because the lower the continuation probability, the more supergames a subject plays, matching group size gets smaller as the continuation probability increases. Similarly, subjects in *2Partner* always interact with two different partners, whereas subjects in *1Partner* only interact with one partner at a time. Therefore, *2Partner* requires larger matching groups than *1Partner*. See columns 4 and 5 of Table 2.A.5 for information about the matching groups.

2.A.2.2 Stopping procedure and supergames' duration

The first three rounds of a supergame are played for certain and at the end of the third round, a computerized stopping rule is introduced. From round three onward, the supergame either proceeds to the next round (with continuation probability δ), or it stops and subjects move to a new supergame. Rather than randomly stopping after each round greater or equal than 3, the computer generates a sequence of supergames at the beginning of the session. The duration of each supergame is drawn from a geometric distribution and a new supergame is added to the sequence up until the total number of rounds exceeds 100. We wanted to let matching groups within a session go through independent sequences of supergame durations. Subjects in a matching group go through the same sequence of supergames. At the end of each round, subjects have to wait until all the subjects in their session, irrespective of their matching group, have taken their decisions and observed the results before moving to the next round. This spreads out waiting out times more evenly throughout the session.

Since the total number of rounds played by a matching group is random, matching groups within a session do not necessarily finish at the same time. To avoid having subjects in matching groups, which are not last to finish, wait on others and possibly infer with whom they interacted, we make those matching groups play a last supergame of finite duration. The duration of the matching group's finite supergame is chosen to ensure that all matching groups within a session finish at the same time. To illustrate this, consider matching groups 71 and 72 in Table 2.A.5, which were the only two in session 7. The computer drew 33 and 34 supergames summing up to 101 and 102 rounds for matching groups 71 and 72, respectively. Since matching group 71 would have played one round less than matching group 72, we add a supergame of one round to matching group 71. Subjects who play a finite supergame are informed about the finite character of the game and its duration. Data from these finite supergames are not part of the analysis. To avoid any effects on the main part of the experiment we do not mention the possibility of playing a supergame of finite length in the instructions.

For the second experiment, we did not generate the supergame durations on the spot, but used the realizations of the six matching groups of the first experiment at $\delta = 0.5$ in *1Partner* for all treatments. The goal was to maximize comparability between the two experiments.

Table 2.A.5: Supergames' duration by matching group

Treatment			Match. gr.		Supergames duration						Statistics supergames					
δ	Sim.	$2P$	ID	Subj.	3	4	5	6–8	9–11	12+	N	Mean	Min	Max	Total	Finite
0.1	1	0	21	10	29	2	1	0	0	0	32	3.1	3	5	100	0
			22	10	32	1	0	0	0	0	33	3.0	3	4	100	0
			71	10	31	2	0	0	0	0	33	3.1	3	4	101	1
			72	10	34	0	0	0	0	0	34	3.0	3	3	102	0
			221	10	32	1	0	0	0	0	33	3.0	3	4	100	0
			222	10	28	4	0	0	0	0	32	3.1	3	4	100	0
		1	11	20	31	2	0	0	0	0	33	3.1	3	4	101	0
			81	20	26	6	0	0	0	0	32	3.2	3	4	102	0
			141	20	31	2	0	0	0	0	33	3.1	3	4	101	0
			191	18	27	5	0	0	0	0	32	3.2	3	4	101	0
			211	18	30	3	0	0	0	0	33	3.1	3	4	102	0
			231	20	28	3	1	0	0	0	32	3.2	3	5	101	0
0.5	1	0	61	10	10	8	1	2	1	1	23	4.4	3	12	102	1
			62	8	13	7	3	3	0	0	26	4.0	3	8	103	0
			151	10	13	7	3	3	0	0	26	3.9	3	7	102	0
			152	10	15	7	3	2	0	0	27	3.8	3	8	102	0
			171	10	12	7	4	3	0	0	26	4.0	3	7	103	0
			172	10	15	6	1	4	0	0	26	3.9	3	8	101	2
		1	51	20	13	5	5	3	0	0	26	4.0	3	7	104	0
			91	20	10	7	1	4	1	0	23	4.4	3	11	101	0
			101	18	12	9	2	3	0	0	26	3.9	3	8	102	0
			121	18	11	5	3	5	0	0	24	4.2	3	8	101	0
			181	20	11	7	4	3	0	0	25	4.0	3	8	101	0
			201	18	14	3	1	5	1	0	24	4.2	3	9	100	0
	0	0	251	8	10	8	1	2	1	1	23	4.4	3	12	102	1
			252	6	13	7	3	3	0	0	26	4.0	3	8	103	0
			281	8	13	7	3	3	0	0	26	3.9	3	7	102	0
			282	8	15	7	3	2	0	0	27	3.8	3	8	102	0
			301	8	12	7	4	3	0	0	26	4.0	3	7	103	0
			302	6	15	6	1	4	0	0	26	3.9	3	8	101	2
		1	241	16	10	8	1	2	1	1	23	4.4	3	12	102	0
			261	16	13	7	3	3	0	0	26	4.0	3	8	103	0
			271	12	13	7	3	3	0	0	26	3.9	3	7	102	0
			291	16	15	7	3	2	0	0	27	3.8	3	8	102	0
			311	10	12	7	4	3	0	0	26	4.0	3	7	103	0
			321	14	15	6	1	4	0	0	26	3.9	3	8	101	0
0.9	1	0	31	6	0	1	1	1	0	4	7	15.4	4	48	108	11
			32	6	2	0	0	2	0	4	8	12.6	3	29	101	18
			33	6	0	0	1	6	2	3	12	9.9	5	21	119	0
			131	8	0	1	2	2	0	4	9	13.2	4	30	119	0
			132	8	1	1	5	0	3	2	12	8.7	3	24	104	15
		1	41	10	1	2	1	1	1	4	10	11.1	3	29	111	0
			42	10	0	2	0	2	2	3	9	11.9	4	28	107	4
			111	14	2	1	0	1	1	5	10	10.5	3	19	105	0
			161	10	0	1	1	1	3	4	10	10.9	4	21	109	0
			162	10	1	2	0	2	0	2	7	15.1	3	65	106	3

Notes. Summary of the supergames duration for each matching group. The first three columns show the treatment played. *2P* is a binary variable taking the value 1 for *2Partner* and 0 for *1Partner*. Sim. is a binary variable taking the value 1 if the *easy* and *hard* games were played simultaneously and 0 if they were played sequentially. The last digit of a matching group ID is its number in the session and the other digits form the session ID (e.g., 21 refers to matching group 1 in session 2 and 222 refers to matching group 2 in session 22). There are up to 3 matching groups within a session. Subj. is the number of subjects in the matching group. Columns 6 to 11 show frequencies of supergames duration for six selected intervals. Columns 12 to 16 show summary statistics for supergame duration.

2.A.3 Experimental instructions

The instructions were originally written in French. Depending on the treatment, minimum changes were made to the instructions. Below you can see the translated instructions for the treatment *2Partner* at $\delta = 0.5$ when both games are played simultaneously.

Instructions

General information

You are going to participate to a study financed by the Swiss National Science Foundation (SNSF). Depending on your decisions, you will have the opportunity to earn a substantial amount of money. Please read the following instructions carefully.

These instructions are exclusively reserved for your usage. You are not allowed to communicate with the other participants. If you violate this rule, you will be banned from the experiment and receive no payment.

Throughout the study, we will not speak in CHF but in points. At the end of the study, your gains will be converted to CHF. The exchange rate between CHF and points is CHF 1 = 1000 points. Once the study is finished, you will receive your gains in cash plus a show-up fee of CHF 10.

The study is divided into matches. For each match, you are paired with two other randomly drawn participants in the room. These participants are called your partners. You will interact with these same two partners for several rounds. We will see later what determines the length of a match. Once a match is over, two new partners are randomly drawn. The figure below shows the difference between matches and rounds:

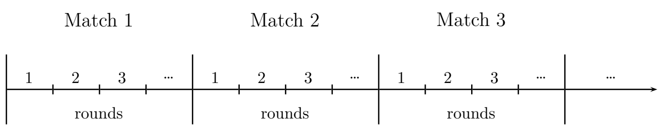


Figure 1: Matches and rounds

Your identity will never be revealed and you will never receive information about your partners.

You will play several matches and your partners will change between each match. You do not know how many matches you will play.

Rules of the game

Below, you can see the decisions screen. The header shows the current match number as well as the round number in the current match. Here you have the example of round 1 in match 2.

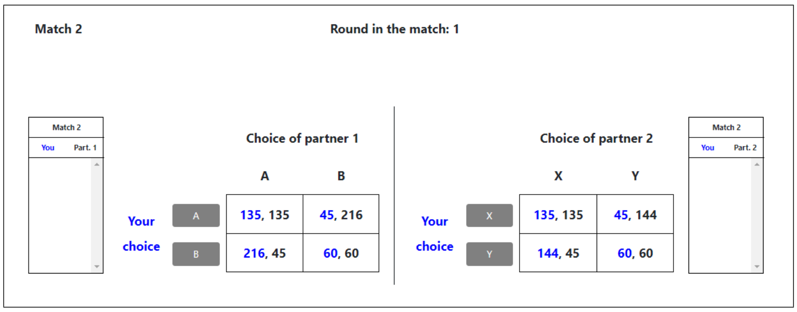


Figure 2: Decisions screen

The body of the screen is divided into two parts by a vertical line. At each round, you have two decisions to take. Specifically, you have to take one decision for the left part of the screen and one for the right part. The two tables in the middle show the possible gains for you and your partners. The decision for each table consists of choosing between the first line and the second line. In the left table, click on the gray button **A** or **B** to choose the first or second line. The decision is similar for the right part clicking the gray button **X** or **Y**. Your partner 1 does the same either choosing column **A** or column **B** on the left part and your partner 2 either chooses column **X** or column **Y** on the right part.

Each table contains four cells. The first number in blue of each cell is your gain for the round if this cell is the result of your decision and the one of your partner. The second number in black of each cell is your partner's gain. The following lines shows the four possible cases for the table on the left.

- You - **A** / Partner 1 - **A** → You - 135 Points / Partner 1 - 135 Points
- You - **A** / Partner 1 - **B** → You - 45 Points / Partner 1 - 216 Points
- You - **B** / Partner 1 - **A** → You - 216 Points / Partner 1 - 45 Points
- You - **B** / Partner 1 - **B** → You - 60 Points / Partner 1 - 60 Points

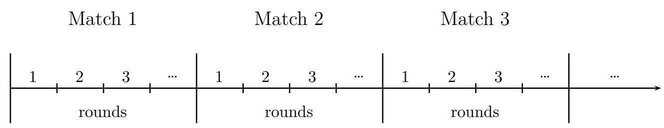
The reasoning is similar for the table on the right.

The tables on the very left and right parts of the screen remind you of your decisions and those of your respective partner for each half of the screen. Only the decisions of the current match are shown. Since the example is for round 1, the summary tables are still empty.

To take your decisions, you have to click on the gray buttons for each of the two tables in the middle. By clicking on a button, it becomes blue. Once you have taken your two decisions, a green button "Validate" appears in the lower right corner. By clicking this button, you move to the results screen. This screen will inform you about the choice of each of your partners. The results will be highlighted and your gain for each part will be displayed. If the match continues, you move to the next decisions screen and play the same game with the same two partners. If the match ends, a new screen will appear and inform you that two new partners will be randomly drawn.

Length of a match

Finally, we are going to look what determines the length of a match.



A match lasts at least 3 rounds. That means you are going to interact at least three times in a row with the same two partners.

From round 3 on, the match will stop randomly. More precisely, the match can stop at the end of round 3 with a probability of 1 chance out of 2. If the game does not stop, you move to a round 4 and there is again 1 chance out of 2 the match will stop at the end of round 4. The reasoning is identical for rounds 5 on, the match stopping at the end of each round with 1 chance out of 2. The computer randomly determines the stopping of a match.

To summarize, a match lasts at least 3 rounds. Starting from the end of round 3, the match stops at the end of each round with a probability of 1 chance out of 2.

The results screen will inform you whether the match continues or stops. Once a match is over, you move to the next one. As a reminder, two new partners are randomly drawn for the next match.

Make sure you understand the instructions. If something is not clear, please raise your hand and the organizer will come to help.

Chapter 3

Multigame Contact: There Can Be More to Lose than to Win

VINCENT LAFERRIÈRE[†]

We report on an experiment examining the effect of multigame contact on cooperation. Subjects play two indefinitely repeated games, a prisoner's dilemma and a stag hunt game, either with the same partner, or with two different partners. The second treatment dimension is the order of play within a round: either the two games are played simultaneously, or the prisoner's dilemma is played before the stag hunt game. Contrary to the theoretical predictions, multigame contact does not improve cooperation in the prisoner's dilemma. When games are played simultaneously, multigame contact even leads to less efficient outcomes. Non-credible threats can explain why multigame contact does not help, or is detrimental, in our experiment.

3.1 Introduction

Because of its importance for human interactions, cooperation is a prominent topic in experimental economics. The most widely used game to study cooperation between two individuals is the prisoner's dilemma, especially its indefinitely repeated version. In the prisoner's dilemma, two players simultaneously decide whether to cooperate or not. This simple binary choice allows us to study how subjects balance the short-term incentives to defect and the long-term incentives to cooperate. A large body of literature focused on cooperation in various situations and whether or not we can create environments that foster cooperation.[1]

In this chapter, we build on Laferrière, Montez, et al. (2022) and provide further experimental evidence on a factor that—in theory—can foster cooperation between two in-

[†]Chapter 3 is a follow-up project based on Chapter 2 (henceforth, Laferrière, Montez, et al., 2022). The concept of this experiment was developed jointly with Joao Montez, Catherine Roux, and Christian Thöni. I worked on my own for the rest of the project; I designed the specifics of the experiment, ran all experimental sessions, collected and cleaned the data, performed the data analysis, and wrote this chapter.

[1]Many factors affect cooperation in the prisoner's dilemma, such as continuation probability (Dal Bó, 2005; Duffy & Ochs, 2009; Fréchette & Yuksel, 2017; Normann & Wallace, 2012), experience (Dal Bó & Fréchette, 2011), communication possibilities (Cooper & Kühn, 2014), monitoring structure (Aoyagi et al., 2019; Camera & Casari, 2009), costly punishment (Camera & Casari, 2009; Dreber et al., 2008), timing of play (Bigoni et al., 2015; D. Friedman & Oprea, 2012), and behavioral spillovers (Bednar et al., 2012). See Dal Bó and Fréchette (2018) for a comprehensive overview of the literature.

dividuals: multigame contact. We consider a situation with multigame contact when two individuals interact across multiple games. The theoretical foundations describing the effects of multigame contact on cooperation come from industrial economics. Bernheim and Whinston (1990) show that two firms interacting in more than one market can sustain collusive behaviors that are impossible to sustain when interacting with different firms across markets. In a multimarket situation, behaviors in one market can now be rewarded or punished in both markets; the two games become strategically connected. Collusive behaviors emerge because a threat of future punishment disciplines short-term behaviors. By improving the ability to punish, multimarket contact reduces the incentives to deviate, which fosters collusive behavior. This intuition is directly applicable to the prisoner's dilemma as players face the same dilemma between the short- and long-term benefits. To better understand how multigame contact can impact cooperation, we pair a prisoner's dilemma with a stag hunt game. With two obvious Nash equilibria in pure strategies, one leading to a higher payoff for both players than the other, the stag hunt game offers a simple way to punish or reward behaviors. The stag hunt game is also a staple in experimental economics (see Dal Bó et al., 2021, for a comprehensive overview) but, to our knowledge, has never been paired with an indefinitely repeated prisoner's dilemma to study multigame contact.[2]

We set up a laboratory experiment where subjects play two indefinitely repeated games in parallel: a prisoner's dilemma and a stag hunt game. Our primary treatment is the presence or absence of multigame contact. Multigame contact occurs when a subject plays the two indefinitely repeated games with the same partner. There is no multigame contact when a subject plays each game with a different partner. We add a second treatment dimension: the order of play within a round. In the simultaneous treatment, subjects take their decisions for the prisoner's dilemma and the stag hunt game at the same time. In the sequential treatment, they first take their decision in the prisoner's dilemma, learn the outcome of the game, and then take their decision in the stag hunt game.

The payoffs in each game and the critical discount factor we implement are such that—in theory—cooperation in the prisoner's dilemma is possible with multigame contact but not without; this is true whether the games are played simultaneously or sequentially. In every treatment, we expect frequent coordination on the payoff-dominant equilibrium in the stag hunt game.

In our experiment, we observe little cooperation in the prisoner's dilemma but frequent coordination on the payoff-dominant equilibrium in the stag hunt game without multigame contact. Counter to the theoretical predictions, multigame contact does not help to increase efficiency, measured by the frequencies of cooperation in the prisoner's dilemma and coordination on the payoff-dominant equilibrium in the stag hunt game. When games are played sequentially, we replicate the "double-edged sword" that Laferrière, Montez, et al. (2022) find: multigame contact leads more often to situations where the efficient outcome is either reached in both games at the same time or in none of them. When the two games are played simultaneously within a round, multigame contact is even detrimental. It does not increase cooperation in the prisoner's dilemma, while it decreases coordination on the payoff-dominant equilibrium in the stag hunt game. This negative effect goes counter to the theoretical predictions that multigame contact can help to achieve

[2]Duffy and Fehr (2018) look at equilibrium selection when subjects play a sequence of indefinitely repeated prisoner's dilemmas followed by a similar sequence of stag hunt games (or vice versa). They find that a precedent for efficient or inefficient play in one sequence does not carry over to the other. They conclude that equilibrium selection depends more on strategic considerations than historical precedents.

cooperation but never hurts.

A possible explanation for the absence of beneficial effects from multigame contact is that, for it to be effective, subjects have to punish across games and this punishment has to be credible. Since coordination on the payoff-dominant equilibrium in the stag hunt game is very frequent without multigame contact, subjects in a situation of multigame contact must be ready to sacrifice future payoffs in the stag hunt game to induce cooperation in the prisoner's dilemma. Since we observe that subjects do not always use the stag hunt game to punish deviations in the prisoner's dilemma, such strong punishments become less credible and thus lose some disciplinary power. Therefore, there may be a trade-off between a punishment's credibility and its strength.

This chapter closely follows the novel experimental design in Laferrière, Montez, et al. (2022), which improved on the previous literature on multimarket contact (Feinberg & Sherman, 1985, 1988; Freitag et al., 2021; Güth et al., 2016; Modak, 2021; Phillips & Mason, 1992, 1996; Yang et al., 2016). So far, the literature found mixed effects of multigame contact on cooperation. By pairing a prisoner's dilemma and a stag hunt game, we take a step towards understanding why the beneficial effects of multigame contact often fail to materialize. This chapter also builds on the literature focusing on the effect of punishment on cooperation (e.g., Carpenter, 2007; Egas & Riedl, 2008; Fehr & Gächter, 2002; Nikiforakis & Normann, 2008; Roux & Thöni, 2015). The fact that being able to punish does not lead to more efficient outcomes is reminiscent of Dreber et al. (2008). They make subjects play an indefinitely repeated prisoner's dilemma but add a third strategy to the stage game that allows subjects to punish their partner. In their experiment, this new strategy increases the frequency of cooperation but does not increase average payoffs. In all these experiments, the extended ability to punish either comes from a change in the available strategies or by adding a costly punishment phase after the game outcomes are revealed. By contrast, multigame contact allows us to study the effect of increasing the ability to punish without modifying the games.

In what follows, Section 3.2 derives the theoretical predictions, Section 3.3 presents the experiment, Section 3.4 discusses our results, and Section 3.5 briefly concludes.

3.2 Theory

This section presents the theory underlying our experiment. In Section 3.2.1, we analyse the two games of interest without multigame contact. Section 3.2.2 provides an analysis for situations with multigame contact and we derive the theoretical predictions for our experiment in Section 3.2.3.

3.2.1 Without multigame contact

Figure 3.1 presents the two stage games of interest. In each period, players have to take a decision in both games and we consider an infinite horizon. At any point in time, players discount the next period payoffs with a common discount factor $\delta \in (0, 1)$. Another way to look at such a game is to consider δ as the probability to reach the next period, which is usually how infinitely repeated games are implemented in laboratory experiments (Murnighan & Roth, 1983; Roth & Murnighan, 1978). At any point in time, the game continues—at least for another period— with probability δ or stops with probability $(1-\delta)$. In this case, δ is called the continuation probability and it becomes an indefinitely repeated game. The theoretical predictions are identical whether we interpret δ as a discount factor or a continuation probability. As a consequence, we first interpret δ as a discount factor (infinitely repeated game), but when we derive the testable predictions in Section 3.2.3 we interpret it as a continuation probability (indefinitely repeated game).

Prisoner's dilemma

	C	D
C	R, R	S, T
D	T, S	P, P

Stag hunt game

	$stag$	$hare$
$stag$	r, r	s, t
$hare$	t, s	p, p

Figure 3.1: Payoff matrices of the two stage games

Without multigame contact, players are matched with a different player for each game. Therefore, we can treat the two games independently; what happens in one game should bear no influence on the other game. Note that this is only true if we assume that players' utility function depends linearly on payoffs and/or utility is additive in the payoffs of each game. Assuming, for example, strict concavity (convexity) in the utility function and that payoffs in both games are perfect substitutes, a higher payoff in one game decreases (increases) marginal utility in the other game, which then affects a player's incentive in the other game. For simplicity, we assume throughout the paper that utility is a linear transformation of payoffs.[3]

In the prisoner's dilemma, each player can either cooperate (C) or defect (D). R is the reward payoff from joint cooperation, whereas P, the punishment payoff, results from joint defection. In the asymmetric outcomes (C, D) and (D, C), the defecting player receives the temptation payoff T, whereas the player who cooperates receives the sucker's payoff S. By imposing the following conditions on the payoffs, the stage game has a prisoner's dilemma structure with (D, D) being the only stage-game Nash equilibrium: $T > R > P > S$. The following condition ensures that joint cooperation (C, C) is more efficient than takings turns in getting the temptation and sucker's payoffs when the game is repeated: $2R > T + S$. Although the stage game has a single Nash equilibrium (D, D), cooperative equilibria emerge with infinite repetition if the discount factor is high enough (J. W. Friedman, 1971). We want to find the minimum critical discount factor above which reaching the most efficient outcome (C, C) in every round is part of a subgame-perfect Nash equilibrium. It occurs when both players follow a *grim trigger* strategy: starts with C in period 1, plays C if (C, C) was the outcome in the previous period, or plays D forever otherwise. In other words, players revert to the stage-game Nash equilibrium forever after any deviation from (C, C), which is the strongest punishment possible. Therefore, joint cooperation forms a subgame-perfect Nash equilibrium if the following incentive compatibility constraint holds:

$$\frac{R}{1-\delta} \geqslant T + \frac{\delta P}{1-\delta} \Leftrightarrow \delta \geqslant \frac{T-R}{T-P} = \delta_{PD}^{SPE}.$$

The left-hand side of the first inequality is the sum of discounted payoffs on the equilibrium path. The right-hand side is the sum of discounted payoffs when deviating; that is, a player gets the payoff T in the period she deviates and the payoff P in all following

[3]See Spagnolo (1999a) for a discussion on the effect of concave utility functions in situations with multigame contact.

periods. Intuitively, the critical discount factor is increasing in the payoff a player gets when deviating (T) and the payoff on the punishment path (P), and decreasing in the payoff on the equilibrium path (R).

In the stag hunt game, each player has two pure strategies: $stag$ and $hare$. By imposing the following conditions on the payoffs, the stage game has a stag hunt structure: $r > t \geqslant p > s$. There are two Nash equilibria in pure strategies $(stag, stag)$ and $(hare, hare)$ with the former one being payoff dominant (Harsanyi & Selten, 1988) because $r > p$. There is also a Nash equilibrium in mixed strategies where each player chooses $stag$ with probability $(p - s)/(r - t + p - s)$ and $hare$ with probability $(r - t)/(r - t + p - s)$. In contrast to the prisoner's dilemma, infinite repetition of the stage game does not help to reach more efficient Nash equilibria. The most efficient outcome $(stag, stag)$ is already a Nash equilibrium of the stage game. Therefore, both players choosing $stag$ in every period forms the most efficient Nash equilibrium of the repeated game and it is subgame perfect. The stag hunt game is, in essence, a coordination problem. Although $(stag, stag)$ payoff dominates $(hare, hare)$, playing $stag$ comes at the risk of getting the lowest payoff s. Lowering s does not reduce the set of Nash equilibria in the stage game but makes $stag$ less attractive, especially if the payoff difference between the two Nash equilibria in pure strategies is low $(r - p)$. This consideration led Harsanyi and Selten (1988) to propose risk dominance as a refinement concept to take into account how a Nash equilibrium is robust to strategic uncertainty.[4] In a symmetric two-player game with two pure strategies, like the stag hunt game, risk dominance has an intuitive explanation. We say that $stag$ $(hare)$ is risk dominant if $stag$ $(hare)$ is the best response to the other player mixing between $stag$ and $hare$ with probability one half. It follows that $stag$ $(hare)$ is risk dominant if: $r + s \geqslant (\leqslant)\ t + p$.[5]

Although risk dominance was developed as a refinement concept for static games with multiple Nash equilibria, it is possible to extend the underlying idea to infinitely repeated games. Before applying it to the prisoner's dilemma as proposed by Blonski et al. (2011) and Blonski and Spagnolo (2015), let us focus first on an infinitely repeated stag hunt game. Thanks to the repetition, a player can take more risks in the first period because she will be able to react to the other player's move. Suppose the following trigger strategy: start with $stag$ in period 1, play $stag$ if $(stag, stag)$ was the outcome in the previous period, or play $hare$ otherwise. With this strategy, a player tries to coordinate on the payoff dominant equilibrium but reverts to the "safe" equilibrium forever when faced with $hare$. Both players following this strategy trivially forms a subgame-perfect Nash equilibrium because every move is part of a stage-game Nash equilibrium. This strategy can be compared with the "safest" strategy, which is to play $hare$ in every period. It is the safest strategy as it is the only one that guarantees never to get the lowest payoff (s). By restricting to these two strategies, we are back in a situation where players have two

[4]When payoff dominance and risk dominance contradict, the authors first considered that payoff dominance should prevail. After the publication of Aumann (1990), Harsanyi (1995) revised his stance and agreed that risk dominance should prevail over payoff dominance in such situations.

[5]In order to get a continuous measure on how close to risk dominance a strategy is, more recent studies (Dal Bó et al., 2021) have focused on the analog concept of the basin of attraction. To identify whether $stag$ $(hare)$ is risk dominant, we look for the maximum probability of the other playing $hare$ $(stag)$—i.e. the basin of attraction $stag$ $(hare)$—that still makes $stag$ $(hare)$ a best response. Because a player is indifferent between the two pure strategies in the Nash equilibrium in mixed strategies, the basin of attraction of $stag$ $(hare)$ is the same as the probability of the other player choosing $hare$ $(stag)$ in the Nash equilibrium in mixed strategies. If the basin of attraction of $stag$ $(hare)$ is larger than one half, then $stag$ $(hare)$ is risk dominant.

strategies available and symmetric payoffs. Therefore, the proposed trigger strategy is risk dominant if:

$$\frac{1}{2}\left(\frac{r}{1-\delta}\right)+\frac{1}{2}\left(s+\frac{\delta p}{1-\delta}\right)\geqslant\frac{1}{2}\left(t+\frac{\delta p}{1-\delta}\right)+\frac{1}{2}\left(\frac{p}{1-\delta}\right)$$
$$\Leftrightarrow\delta\geqslant\frac{t-r+p-s}{t-s}=1-\frac{r-p}{t-s}=\delta_{SH}^{RD}.$$

Unsurprisingly, when $stag$ is risk dominant in the stage game, then the numerator, and thus δ_{SH}^{RD}, is negative, meaning that the proposed trigger strategy is always risk dominant. More interestingly, there is now a region of parameters for which the trigger strategy is risk dominant, although $stag$ is not risk dominant in the stage game. Intuitively, the less attractive the payoff dominant Nash equilibrium is ($r-p$ is small), or the riskier $stag$ is ($t-s$ is large), the higher must be the critical discount factor.

In the prisoner's dilemma, cooperation can emerge as subgame-perfect Nash equilibrium through the use of *trigger* strategies for sufficiently high discount factors. By restricting the set of strategies to the *grim trigger* strategy discussed above and the *always defect* strategy (plays D in every round), both strategies form subgame-perfect Nash equilibria. Each strategy is the best response to the other playing the same strategy, and we end up with a game similar to the stag hunt game: *(grim trigger, grim trigger)* payoff dominates *(always defect, always defect)* with the former being riskier as a player takes the risk of getting the lowest payoff S if the other player deviates. As Skyrms (2001, p. 34) explains: "The Shadow of the Future has not solved the problem of cooperation in the Prisoner's Dilemma; it has transformed it into the problem of cooperation in the Stag Hunt." Although restricting to these two strategies seems arbitrary, Blonski et al. (2011) and Blonski and Spagnolo (2015) provide intuitions and theoretical arguments justifying these choices. *always defect* is the unique "safe" equilibrium as it is the only one that guarantees never getting the lowest payoff (S). The *grim trigger* strategy is shown to be the least "risky" strategy; that is, if any cooperative strategy risk dominates *always defect*, then *grim trigger* also risk dominates *always defect*.[6] The *grim trigger* strategy risk dominates *always defect* if it is the best response to the other player choosing each strategy with probability one half:

$$\frac{1}{2}\left(\frac{R}{1-\delta}\right)+\frac{1}{2}\left(S+\frac{\delta P}{1-\delta}\right)\geqslant\frac{1}{2}\left(T+\frac{\delta P}{1-\delta}\right)+\frac{1}{2}\left(\frac{P}{1-\delta}\right)$$
$$\Leftrightarrow\delta\geqslant\frac{T-R+P-S}{T-S}=1-\frac{R-P}{T-S}=\delta_{PD}^{RD}.$$

Again, the critical discount factor δ_{PD}^{RD} is increasing in T and P and decreasing in R. Contrary to the critical discount factor for subgame perfection (δ_{PD}^{SPE}), δ^{RD} depends on the sucker payoff (S) and the relationship is negative. Intuitively, the higher this payoff is, the less risky it is to try to establish cooperation. Note that risk dominance is always more stringent than subgame perfection in the prisoner's dilemma $(\delta_{PD}^{RD}>\delta_{PD}^{SPE})$. Since the

[6]Note that we only consider cooperative strategies in which players coordinate on the most efficient outcome (C,C). Other cooperative strategies could be considered, such as alternating between the outcomes (C,D) and (D,C). Such cooperative paths are less efficient than the (C,C) path but can be sustained at lower discount factors and can also be less risky. However, it is hard to imagine that subjects in a laboratory experiment could coordinate on such complicated equilibria, especially without communication. Dal Bó (2005), Fréchette and Yuksel (2017), and Laferrière, Montez, et al. (2022) look for such strategies in the data and find very little evidence.

interest of this paper lies in the effect of multigame contact and not on the determinants of cooperation, we are going to choose a parameter space in which these two criteria agree.

3.2.2 With multigame contact

For games with a prisoner's dilemma structure, Bernheim and Whinston (1990) show that multigame contact can allow achieving simultaneous cooperation in both games at a lower critical discount factor than without multigame contact and it never restricts players' ability to cooperate.[7] Their key insight is that players can link the two games, meaning that a deviation from cooperation in one game can lead to punishment in both games. Enhancing players' ability to punish makes deviations less attractive, which reduces the critical discount factors required to achieve simultaneous cooperation in both games. Pairing a prisoner's dilemma and a stag hunt game provides an intuitive way to think about the effect of multigame contact on cooperation. The multiplicity of equilibria in the stag hunt game gives a simple way to reward cooperation in the prisoner's dilemma by playing the efficient Nash equilibrium $(stag, stag)$ or to punish defection with the payoff dominated Nash equilibrium $(hare, hare)$.[8]

Before deriving the critical discount factors for cooperation in the prisoner's dilemma, we have to make a distinction between two situations relevant to our experiment. Players may take their decision simultaneously for both games and learn the outcome of each game before moving to the next period, or players could play the two games sequentially within a period; that is, taking their decision in one game and learning the outcome of this game before playing the second one. Although this distinction makes no difference without multigame contact—the two games being independent—this has implications for the ability to punish with multigame contact.

When the two games are played simultaneously, we can construct a strategy similar to the previous *grim trigger* to achieve cooperation in a subgame-perfect Nash equilibrium: plays C and $stag$ in period 1, plays C and $stag$ if the outcomes in the previous period were (C, C) and $(stag, stag)$, or play D and $hare$ otherwise. Both players following this strategy forms a subgame-perfect Nash equilibrium if the following incentive compatibility constraint holds:

$$\frac{R+r}{1-\delta} \geqslant T+r+\frac{\delta(P+p)}{1-\delta} \Leftrightarrow \delta \geqslant \frac{T-R}{T-P+r-p} = \delta_{Sim}^{SPE}.$$

The left-hand side of the first inequality is the sum of discounted payoffs on the equilibrium path and the right-hand side is the sum of discounted payoffs when deviating. In this case, the optimal deviation is to deviate only in the prisoner's dilemma because $stag$ is still the best response to the other player choosing $stag$ in the stag-hunt stage game. The last term on the right-hand side is the continuation payoff on the punishment path. The

[7] Multigame contact does not reduce the critical discount factor for simultaneous cooperation in both games if the payoffs in one game are a linear transformation of the other game or if there is too much asymmetry in the incentive to cooperate between the two games.

[8] This is reminiscent of Krishna et al. (1985) who look at finite repetition of a stage game that has multiple Nash equilibria. In such situations, it may be possible to construct subgame-perfect Nash equilibria of the repeated game in which players coordinate on a cooperative outcome, which is not itself a Nash equilibrium of the stage game. However, the increase in cooperation is only limited in time as the game has to converge back to a stage game Nash equilibrium by the end of the game. Such new equilibria can emerge when pairing a prisoner's dilemma and a stag hunt game, but we will focus our attention on the infinitely repeated variant of the game.

critical discount factor (δ^{SPE}_{Sim}) is still increasing in T and P and decreasing in R. However, it is now decreasing in r and increasing in p, with $r - p$ being the loss per period in the stag hunt game after defection. Since this loss is positive—i.e., punishment has increased—the critical discount factor to sustain cooperation in the prisoner's dilemma has decreased compared with the situation without multigame contact ($\delta^{SPE}_{Sim} < \delta^{SPE}_{PD}$). The beneficial effects of multigame contact on cooperation materialize in this lower critical discount factor.

We can again apply risk dominance to this situation by restricting our attention to the *grim trigger* strategy proposed above and to the safe strategy consisting of always playing D in the prisoner's dilemma and $hare$ in the stag hunt game, which guarantees never getting the lowest payoffs (S and s). This new *grim trigger* strategy is risk dominant if it is the best response to the other player choosing each strategy with probability one half:

$$\frac{1}{2}\left(\frac{R+r}{1-\delta}\right)+\frac{1}{2}\left(S+s+\frac{\delta P+p}{1-\delta}\right)\geqslant\frac{1}{2}\left(T+t+\frac{\delta P+p}{1-\delta}\right)+\frac{1}{2}\left(\frac{P+p}{1-\delta}\right)$$
$$\Leftrightarrow \delta \geqslant \frac{T-R+P-S+t-s-(r-p)}{T-S+t-s}=1-\frac{R-P+r-p}{T-s+t-s}=\delta^{RD}_{Sim}.$$

We can compare the critical discount factor required for *grim trigger* to be the risk-dominant strategy between the single prisoner's dilemma case (without multigame contact) to this new situation. Multigame contact leads to a lower critical discount factor ($\delta^{RD}_{Sim} < \delta^{RD}_{PD}$) if:

$$\frac{R-P}{T-S}\leqslant\frac{r-p}{t-s}$$

Since $T > R > P > S$, the left-hand side is always strictly below one and the right-hand side is larger than one if $stag$ is risk dominant in the stag-hunt ($r + s > t + p$). Therefore, if $stag$ is risk dominant, then the *grim trigger* strategy for multigame contact we proposed leads to a lower critical discount factor than without multigame contact. This inequality can even hold when $stag$ is not risk dominant if the parameters of the prisoner's dilemma are such that the left-hand side is low enough. Therefore, it is even possible to reduce the critical discount factor that makes the *grim trigger* strategy risk dominant ($\delta^{RD}_{Sim} < \delta^{RD}_{PD}$) by pairing the prisoner's dilemma with a stag hunt game in which $stag$ is not a risk-dominant strategy of the stage game.

Moving to the sequential variant of the game, we only consider the case of playing the prisoner's dilemma before the stag hunt game within a period. This is the order leading to the lowest critical discount factor and the one we run in our experiment. We will briefly mention what could be expected in the reverse order. Since players observe the outcome of the prisoner's dilemma before taking their decision in the stag hunt game, they are able to react to a deviation immediately. This further increases their ability to punish. If a player considers deviating, she would nevertheless do it in the prisoner's dilemma. Indeed, no deviation is possible in the stag hunt game, because a stage-game Nash equilibrium is played on and off the equilibrium path. The new incentive compatibility constraint for subgame perfection is:

$$\frac{R+r}{1-\delta}\geqslant T+p+\frac{\delta(P+p)}{1-\delta}\Leftrightarrow\delta\geqslant\frac{T-R-(r-p)}{T-P}=\delta^{SPE}_{Seq}.$$

This leads to an even lower critical discount factor ($\delta^{SPE}_{Seq} < \delta^{SPE}_{Sim}$), because sequential choices have enhanced a player's ability to punish without affecting the achievable payoffs.

Since players play one game after the other, a player's strategy is revealed after the first game. For *grim trigger* to be risk dominant, the following condition has to hold:

$$\frac{1}{2}\left(\frac{R+r}{1-\delta}\right)+\frac{1}{2}\left(S+p+\frac{\delta P+p}{1-\delta}\right)\geqslant\frac{1}{2}\left(T+p+\frac{\delta P+p}{1-\delta}\right)+\frac{1}{2}\left(\frac{P+p}{1-\delta}\right)$$
$$\Leftrightarrow \delta\geqslant\frac{T-R+P-S-(r-p)}{T-S}=1-\frac{R-P+r-p}{T-S}=\delta^{RD}_{Seq}.$$

The critical discount factor for risk dominance is again lower ($\delta^{RD}_{Seq}<\delta^{RD}_{Sim}$) because a player who tries to establish cooperation, but faces deviation, can react in the same period.

We have shown that whether we use subgame perfection or risk dominance as a criterion for cooperation, multigame contact always leads to a lower critical discount factor for sustaining cooperation in the prisoner's dilemma and coordination on the payoff dominant Nash equilibrium in the stag hunt game at the same time. Moreover, this beneficial effect is even more substantial when players observe the outcome of the prisoner's dilemma before playing the stag hunt game within a period. In theory, multigame contact cannot hurt since players could always treat the two games independently if linking the two games was not optimal. Anything achievable without multigame contact is achievable with multigame contact and possibly more. We chose not to look into the reverse order—to play the stag hunt game before the prisoner's dilemma—as the effect on the critical discount factor is weaker in the reverse order. Indeed, the critical discount factor for subgame perfection, assuming players follow a *grim trigger* strategy, for the reverse order of play is identical to the one for the simultaneous case. Intuitively, a player who wants to deviate would do it in the prisoner's dilemma (second game) and the punishment would start in the next period, which is the same situation as the simultaneous case with multigame contact.

3.2.3 Experimental predictions

Prisoner's dilemma

	C	D
C	$135, 135$	$40, 228$
D	$228, 40$	$60, 60$

Stag hunt game

	$stag$	$hare$
$stag$	$135, 135$	$50, 132$
$hare$	$132, 50$	$60, 60$

Figure 3.2: Payoff matrices of the experimental stage games

Figure 3.2 presents the game parameters used in our experiment. Numbers represent monetary payoffs in ECUs (experimental currency unit). The games share the same payoffs when players coordinate: 135 for coordinating on C or $stag$ and 60 for D or $hare$. The payoffs of the asymmetric outcomes differentiate the two games. In the prisoner's dilemma, the stage game payoff when deviating from cooperation is 228 for the player who deviates and 40 for the other player. In the stag hunt game, choosing $stag$ when the other player chooses $hare$ leads to a payoff of 50 and 132 for choosing $hare$ against $stag$.

Without multigame contact, we obtain the following critical discount factors for the infinitely repeated prisoner's dilemma: $\delta^{SPE}_{PD}=0.55$ and $\delta^{RD}_{PD}=0.60$. In the stag hunt stage game, $stag$ is not risk dominant because $135 + 50 < 132 + 60$, but the critical discount factor for the trigger strategy to be risk dominant is $\delta_{SH}^{RD} = 0.09$. All these critical discount factors apply to both simultaneous and sequential cases.

With multigame contact and simultaneous games, the critical discount factors for joint cooperation in the prisoner's dilemma and coordination on $stag$ in the stag hunt game decrease to $\delta_{Sim}^{SPE} = 0.38$ and $\delta_{Sim}^{RD} = 0.44$. These critical discount factors are even lower when the games are played sequentially: $\delta_{Seq}^{SPE} = 0.11$ and $\delta_{Seq}^{RD} = 0.20$.

In our experiment, we implement all treatments with the continuation probability $\delta = 0.5$. Without multigame contact, joint cooperation in the prisoner's dilemma cannot be part of a subgame-perfect Nash equilibrium and the *grim trigger* strategy is not risk dominant ($\delta_{PD}^{SPE}, \delta_{PD}^{RD} > 0.5$). Although $stag$ is not risk dominant in the stage game, the proposed trigger strategy of playing $stag$ until $hare$ is played is risk dominant ($\delta_{SH}^{RD} < 0.5$).

With multigame contact, joint cooperation can be sustained in a subgame-perfect Nash equilibrium and the *grim trigger* is risk dominant for both simultaneous ($\delta_{Sim}^{SPE}, \delta_{Sim}^{RD} < 0.5$) and sequential cases ($\delta_{Seq}^{SPE}, \delta_{Seq}^{RD} < 0.5$).

If we assume that subjects coordinate on the most efficient equilibrium, we expect no cooperation in the prisoner's dilemma without multigame contact but cooperation with multigame contact. Coordination on $stag$ in the stag hunt game is expected in both cases.

Of course, it is optimistic to expect that subjects behave in such a binary way. Previous experiments on the indefinitely repeated prisoner's dilemma (Bruttel, 2009; Dal Bó & Fréchette, 2018) tend to show that the distance to the critical discount factor is a better predictor. Below the critical discount factor, low cooperation rates should be expected and cooperation rates tend to increase linearly as the distance above the critical discount factor increases. Therefore, we expect low cooperation rates in the prisoner's dilemma without multigame contact and no difference between the simultaneous and sequential cases. With multigame contact, cooperation rates in the prisoner's dilemma should increase and should be the highest when games are played sequentially. As $stag$ can always be sustained in a risk-dominant strategy, it's prevalence should be high in all cases.

3.3 Experimental procedures

Subjects play a sequence of indefinitely repeated games that we call supergames. A supergame consists of a random number of rounds and, in each round, a subject has to take a decision for two games: a prisoner's dilemma and a stag hunt game. Figure 3.2 in Section 3.2.3 presents the two stage games. The first treatment dimension is the presence or absence of multigame contact. If a subject plays both games with the same partner, there is multigame contact (henceforth *1Partner*). If she is matched with two different partners for a supergame, there is no multigame contact (henceforth *2Partner*). The second treatment dimension is the order of play within a round. Subjects either take their two decisions at the same time (henceforth *Simultaneous*) or they first take their decision in the prisoner's dilemma, observe the outcome, and then take their decision in the stag hunt game (henceforth *Sequential*). We end up with a two-by-two factorial designs for a total of four treatments. We use a between-subjects design; that is, subjects within a session only play one of the four treatments.

Within a session, subjects are allocated to matching groups and only interact with others in their matching group.[9] At the beginning of each supergame, a subject is randomly matched with one or two partners from her matching group (stranger matching). For each treatment, we have data from six matching groups. Subjects can never infer with whom they interact. We use neutral labels to denote the partner or partners and the strategies. Both games appear on the same screen and we randomize across subjects, whether the prisoner's dilemma or the stag hunt game appears on the left or the right of the screen.[10]

The first three rounds of a supergame are played with certainty and, at the end of the third round, a random stopping procedure is introduced. The supergame continues with probability $\delta = 0.5$ or stops and this procedure is repeated for any subsequent round. Each matching group of a treatment goes through an independent sequence of supergames duration. Rather than drawing supergames duration on the spot, we generated six sequences of supergames duration from a geometric distribution and we match the six sequences to the six matching groups of each treatment. The goal is to maximize comparability across treatments.[11] Subjects play between 20 and 23 supergames for a total of at least 80 rounds.[12] Table 3.A.2 in the appendix (p. 108) shows detailed information on the size of the matching groups and the supergames' duration.

The final payment of a participant is the total amount of ECUs earned for all the rounds and a show-up fee of CHF 15 ($\approx$ USD 17). 1,000 ECUs were equivalent to CHF 1 ($\approx$ USD 1.1). Sessions were run in the laboratory of the University of Lausanne (LABEX) with undergraduate students from the University of Lausanne and the EPFL recruited with ORSEE (Greiner, 2015). The experiment was programmed in oTree (Chen et al., 2016).

We ran 18 sessions for a total of 340 subjects. Table 3.A.1 in the appendix (p. 107) shows a summary of the sessions. A session lasted on average 85 minutes and the average payment per participant was CHF 33 ($\approx$ USD 36). The design of the experiment and the hypotheses were pre-registered prior to data collection in an OSF registry.[13]

3.4 Results

This section is organized in three parts. First, we investigate the effect of multigame contact on the frequency of cooperation in the prisoner's dilemma and the frequency of playing $stag$ in the stag hunt game. Next, we analyze subjects' behavior in the first round of each supergame. In the final part, we examine the outcomes of the stage games and the effects on the stag hunt game when cooperation in the prisoner's dilemma breaks down.

[9]To keep constant the expected number of interactions with a given other subject, the size of a matching group in *2Partner* (16 to 22 subjects) is approximately twice the size as in *1Partner* (8 to 10 subjects).

[10]See Appendix 3.A.2 (p. 109) for screenshots of the experiment in the instructions.

[11]Mengel et al. (2022) show that the realized length of early supergames has substantial impact on cooperation rates in subsequent supergames. This could have strong implications for treatment comparison. Using the same sequences for all treatments helps to mitigate this issue.

[12]In each sequence, we keep drawing matching groups until a total of 80 rounds is reached. Since matching groups within a session go through independent sequences, they may not play the same total number of rounds. To ensure that all matching groups within a session finish at the same time, we add a last supergame of finite duration to those who are not first to finish. Data from the finite supergames are removed from our analysis and subjects were not informed in advance that they could play a last finite supergame. Therefore, they do not bear any influence on the rest of the data.

[13]https://osf.io/p5t2v

3.4.1 Testing the predictions

The left panel of Figure 3.3 shows the frequency of cooperation (C) in the prisoner's dilemma (henceforth *PD*) and the frequency of $stag$ in the stag hunt game (henceforth *SH*) when looking at all the rounds. Contrary to the theoretical predictions, *1Partner* does not lead to more cooperation in *PD* than *2Partner*. In *Simultaneous*, it even points towards *1Partner* reducing the frequency of C (0.33 vs. 0.40, $p = .200$) but the difference is not statistically significant.[14] In *Sequential*, C is slightly more frequent in *1Partner* (0.44 vs. 0.40, $p = .748$) but the difference is small and far from statistical significance. Moving from *Simultaneous* to *Sequential* in *1Partner* increases cooperation in *PD* (0.33 vs. 0.44, $p = .262$) but the difference is again not statistically significant. As expected, there is no difference between *Simultaneous* and *Sequential* in *2Partner* (0.40 vs. 0.40, $p = 1.000$). Although we predicted no cooperation in *2Partner*, it would be unrealistic to assume that subjects would never try to cooperate. Subjects trying to cooperate 40% of the time is almost the same estimates that Laferrière, Montez, et al. (2022) find in their *hard* prisoner's dilemma, which has almost identical theoretical critical discount factors as the prisoner's dilemma used in this experiment. The surprising finding is that playing the two games with the same partner does not help to increase cooperation in the prisoner's dilemma and it even seems to hurt in *Simultaneous*.

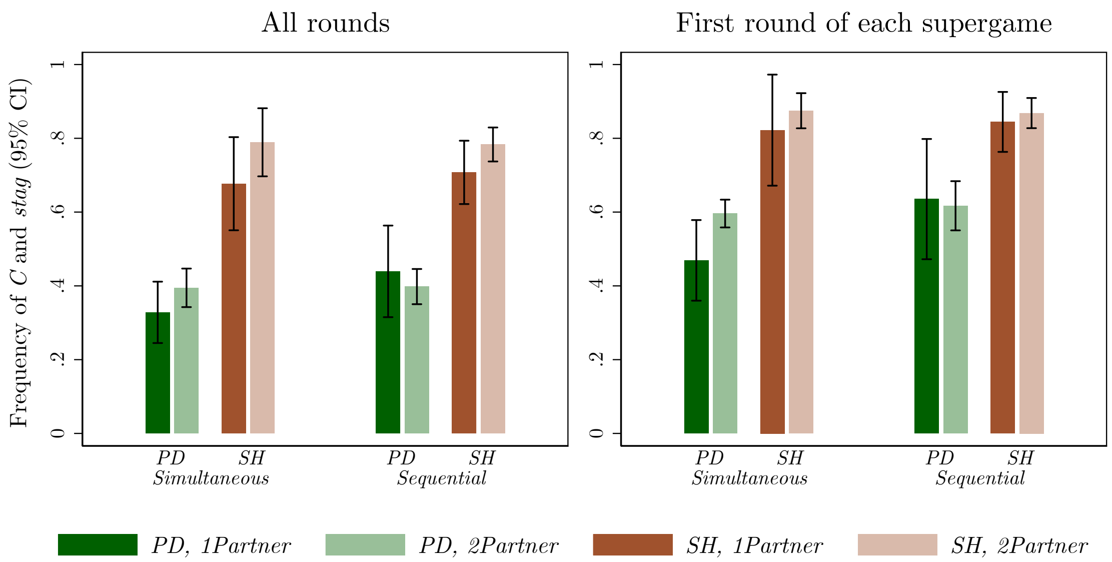


Figure 3.3: Frequency of cooperation (C) in the prisoner's dilemma and $stag$ in the stag hunt game by treatment. Mean cooperation rates and 95% confidence intervals are computed using the matching group averages. The left panel contains all the rounds; the right panel shows the first round of each supergame.

Looking at the frequency of $stag$, we find lower frequencies in *1Partner* than *2Partner* for both *Simultaneous* (0.68 vs. 0.79, $p = .200$) and *Sequential* (0.71 vs. 0.78, $p = .262$) but the differences are not statistically significant. Unsurprisingly, $stag$ is by large the

[14]Unless specified, we always report $p-$values from Wilcoxon rank-sum tests on matching group averages. There are six matching groups per treatment and, therefore, 12 independent observations for each test. Using the *hard* game in Laferrière, Montez, et al. (2022) as prior, we estimate a minimum detectable effect of 10 percentage points difference in cooperation rates between *1Partner* and *2Partner* (80% power, 5% significance level).

modal decision for all treatments meaning that subjects usually try to coordinate on the efficient outcome in *SH*. Since repetition makes the $(stag, stag)$ Nash equilibrium very attractive, it is not surprising to observe high frequencies of $stag$ in every treatment.

When looking at the evolution of cooperation in *PD* and $stag$ in *SH* as subjects gain experience, we find no interesting dynamics over time. Figure 3.A.1 in the appendix (p. 107) shows the mean frequencies of C and $stag$ for each supergame by treatment. With the exception of the frequency of $stag$ in *2Partner* for *Sequential*, which slightly increases over time, all estimates oscillate around the mean over time.

Overall, the supposed beneficial effects of multigame contact do not materialize in *PD*. This is surprising because multigame contact allows—in theory—to sustain cooperation at a much lower critical discount factor, especially in *Sequential*, which should translate into a higher frequency of cooperation. This leads to the following result:

Result 1: Multigame contact does not increase cooperation in the prisoner's dilemma
Playing the prisoner's dilemma and the stag hunt game with the same partner rather than with two different partners does not make cooperation in the prisoner's dilemma more frequent. It is true whether the two games are played simultaneously or if the prisoner's dilemma is played before the stag hunt game within a round.

Although multigame contact has little effect on average frequencies of C and $stag$, there are reasons to expect it affects how subjects play the games. In Laferrière, Montez, et al. (2022), the authors also find no average effect on the frequency of C but substantial effects on subjects' behavior and the distribution of game outcomes. To better understand subjects' behavior, we first restrict our attention to subjects' decisions in the first round of each supergame. Since subjects change partner or partners at the start of every supergame, it allows to observe their decisions before being affected by their partner or partners' decisions.[15] Moreover, most strategies for the prisoner's dilemma discussed in the literature condition a player's decision on the previous round outcome. For such strategies, the outcome in any period can be predicted from the outcome in the first round.

3.4.2 First-round behavior

The right panel of Figure 3.3 shows the frequency of cooperation (C) in *PD* and the frequency of $stag$ in *SH* when restricting to the first round of each supergame. The previous observations based on all the rounds also hold when restricting to the first round of each supergame. In *Simultaneous*, C is somewhat lower in *1Partner* than *2Partner* (0.47 vs. 0.60, $p = .200$) and the same holds for $stag$ (0.82 vs. 0.88, $p = .872$) but none of the comparisons are statistically significant. In *Sequential*, there are almost no difference between *1Partner* and *2Partner* for both C (0.64 vs. 0.62, $p = .872$) and $stag$ (0.85 vs. 0.87, $p = .872$).

In our experiment, subjects always have to take two decisions in a round, either simultaneously or one after the other. To better understand subjects' behavior, it is not sufficient to look at the average frequencies. Table 3.1 presents the relative frequencies of subjects' decision pair in the first round of each supergame. A decision pair is the joint decisions in *PD* and *SH* with the first letter (capital) referring to the decision in *PD* and the rest (lowercase) to the decision in *SH*. Since there are two stage games with two possible choices in each, there are four possible decision pairs. To compare the distribution of decisions

[15] This is only true if we assume that a subject's decisions in one supergame is not affected by what happened in the previous supergames.

Table 3.1: Subjects' decision pair, first round of each supergame

	Simultaneous		*Sequential*	
	1Partner	*2Partner*	*1Partner*	*2Partner*
$Cstag$	0.45	0.58	0.57	0.58
$Chare$	0.02	0.02	0.07	0.03
$Dstag$	0.37	0.30	0.28	0.28
$Dhare$	0.16	0.10	0.08	0.11
N	1,102	2,468	1,228	2,348

Notes: Relative frequencies of subjects' decision pair in the first round of each supergame. The first letter (capital) of each pair refers to the decision in the prisoner's dilemma and the rest (lowercase) to the decision in the stag hunt game. RS-tests for the difference between *1Partner* and *2Partner* yield $p = .250$ when games are played simultaneously and $p = .514$ when sequentially.

between *1Partner* and *2Partner*, we use Rao-Scott χ^2-tests, which correct for dependence within matching group (henceforth RS-test, Rao & Scott, 1984). Although the comparison between *1Partner* and *2Partner* does not reach statistical significance neither in *Simultaneous* (RS-test, $p = .250$) nor in *Sequential* (RS-test, $p = .514$), there are a few observations worth mentioning. In *Simultaneous*, $Cstag$ is surprisingly less frequent in *1Partner* than in *2Partner* and is the modal decision pair in both. $Chare$ is almost never observed in both treatments. It is reassuring as it would have been difficult to rationalize wanting to cooperate in the prisoner's dilemma while not wanting to coordinate on the efficient equilibrium in the stag hunt game, the former being significantly riskier than the latter. Finally, both $Dstag$ and $Dhare$ are more likely in *1Partner* than *2Partner*.

Moving from *Simultaneous* to *Sequential*, we predicted no difference in *2Partner* as we assume subjects treat the two games independently. The fact that the relative frequencies are almost identical between those two treatments is consistent with our hypothesis. It is for *1Partner* that we predicted an increase in cooperation when moving from *Simultaneous* to *Sequential*. The *grim trigger* strategy we assumed to derive the critical discount factors in Section 3.2.2 makes use of the stag hunt game as an additional mechanism to punish deviations in the prisoner's dilemma. Whereas playing $Chare$ is difficult to rationalize in *Simultaneous*, it can be expected when games are played sequentially in *1Partner*. Indeed, the most notable difference between *1Partner* and *2Partner* in *Sequential* is that $Chare$ occurs more than twice as often in *1Partner* than *2Partner*. However, the relative frequencies in Table 3.1 do not allow us to observe whether subjects in *1Partner* use the stag hunt game as a punishment mechanism to punish deviations in the prisoner's dilemma.

Table 3.2 shows subjects' decision in *Sequential* restricting to the first round of each supergame. Column 1 shows the frequency of C, which is the first decision subjects have to take in the round. The frequencies are almost identical in *1Partner* and *2Partner*. Columns 2 to 5 show the frequency of $stag$ following all possible outcomes in the prisoner's dilemma. For example, column 3 $(stag \mid (C, D))$ shows the frequency of $stag$ following a prisoner's dilemma in which the subject chose C and her partner chose D. The game outcomes are always written from the perspective of the subject. In *1Partner*, $stag$ is almost always played following the cooperative outcome (C, C) and the frequency of $stag$ is similar following the three other outcomes $(0.74$ to $0.78)$. These facts are consistent with some use of trigger strategies that punish any deviation in the prisoner's dilemma

Table 3.2: Subjects' decisions in *Sequential*, first round of each supergame

	PD	*SH*			
	C	$stag \mid (C, C)$	$stag \mid (C, D)$	$stag \mid (D, C)$	$stag \mid (D, D)$
1Partner	0.64	0.96	0.74	0.78	0.75
2Partner	0.62	0.98	0.88	0.76	0.72

Notes: Frequency of cooperation (C) in the *PD* and frequency of $stag$ in *SH* conditional on the outcome of *PD* in the first round of each supergame in *Sequential*. Frequencies computed using the matching groups averages. RS-tests for the difference between *1Partner* and *2Partner* yield $p = .838$ for C, $p = .351$ for $stag \mid (C, C)$, $p = .014$ for $stag \mid (C, D)$, $p = .825$ for $stag \mid (D, C)$, and $p = .546$ for $stag \mid (D, D)$.

by playing $hare$ in the stag hunt game. It does not seem to matter whether the subject and/or her partner defected in the prisoner's dilemma, which suggests that some subjects anticipate the fact that defection in the prisoner's dilemma can lead to playing the payoff-dominated equilibrium in the stag hunt game.

In *2Partner*, we expect subjects to treat the two games independently. However, we observe that $stag$ is more frequent when a subject chose C in the prisoner's dilemma (0.98 and 0.88) rather than D (0.76 and 0.72). To assume some heterogeneity between subjects could explain these differences. As mentioned before, a subject willing to cooperate in the prisoner's dilemma should probably be willing to coordinate on the efficient Nash equilibrium in the stag. This could also explain why $stag$ is more frequent following C than D but it cannot explain why $stag$ is less likely following (C, D) than (C, C). However, this difference is much larger in *1Partner* than *2Partner*, which are our main treatments of interest. In fact, the difference in the frequency of $stag$ following (C, D) is the largest and only statistically significant difference between *1Partner* and *2Partner* (RS-test, $p = .014$).

In the end, we observe that subjects in *1Partner* are not more likely to start with cooperation in the prisoner's dilemma than *2Partner*; they may even be less likely in *Simultaneous*. In *Sequential*, there is a notable difference between *1Partner* and *2Partner* in the frequency of $stag$ after a subject tried to establish cooperation in the prisoner's dilemma but her partner did not. At least some subjects in *1Partner* use the stag hunt game—as predicted in theory—to punish defection in the prisoner's dilemma. None of these observations suggest a beneficial effect of multigame contact on cooperation in the prisoner's dilemma and coordination on the efficient Nash equilibrium in the stag hunt game. There are even some indications that multigame contact could be counterproductive. In the next section, we extend the analysis to all the rounds by looking at the outcomes of the stage games across treatments.

3.4.3 Outcomes of the stage games

To simplify the analysis, we restrict our attention to four possible outcomes of the stage games. We consider that a subject has reached the efficient outcome in *PD* (*SH*) if both the subject and her partner chose C ($stag$) in a given round; that is, the outcome is (C, C) in *PD* ($(stag, stag)$ in *SH*). Combining the two stage games results in four possible outcomes. Either a subject reached the efficient outcome in *PD* and *SH*, only in *PD*, only in *SH*, or in neither. Of course, the subject or her partner choosing C ($stag$) and the other D ($hare$) is more efficient than both choosing D ($hare$). However, it is hard to imagine that such asymmetric outcomes can be sustained for more than one round. Since we are interested in the ability of subjects to coordinate on efficient outcomes, we pool these outcomes with the most inefficient ones (D, D) and $(hare, hare)$.

In our experiment, the continuation probability is introduced at the end of round 3. Therefore, rounds 1 and 2 differ from the others in the sense that subjects know they will play at least another round after those two. It appears as a natural threshold for our analysis. In what follows, we perform the same analysis on rounds 1 and 2 together and on rounds 3 and further. The results are similar if we pool all the rounds together.

Table 3.3 shows the relative frequencies of the outcome of the stage games in rounds 1 and 2 of each supergame. Again, the differences between *1Partner* and *2Partner* are not statistically significant, but the following is in particular worth noticing. In *Simultaneous*, the relative frequencies point towards *1Partner* being detrimental to reach efficient outcomes compared with *2Partner*. All relative frequencies for efficient or partially efficient outcomes (rows 1 to 3) are lower in *1Partner* than *2Partner*. In *Sequential*, reaching the efficient outcome in the two stage games in a round is more likely in *1Partner* than *2Partner*, which supports some beneficial effect of multigame contact on cooperation in the prisoner's dilemma. However, this comes at the cost of reaching the efficient outcome in only one of the two stage games less frequently and reaching the efficient outcome in neither also more frequently. This is similar to what Laferrière, Montez, et al. (2022) observe when pairing two prisoner's dilemmas with different incentives to cooperate. They find that multigame contact both leads to more simultaneous cooperation in the two games and simultaneous defection in the two games; they conclude that multigame contact is a "double-edged sword" for cooperation.

Table 3.3: Outcome of the stage games in rounds 1 and 2 of each supergame

Rounds 1 and 2	*Simultaneous*		*Sequential*	
	1Partner	*2Partner*	*1Partner*	*2Partner*
Efficient outcome ...				
... in *PD* and *SH*	0.19	0.28	0.38	0.29
... only in *PD*	0.02	0.05	0.02	0.06
... only in *SH*	0.45	0.47	0.32	0.45
... in neither	0.34	0.21	0.28	0.21
N	2,204	4,936	2,456	4,696

Notes: Relative frequencies of the outcome of the stage games in rounds 1 and 2 of each supergame. A cooperative outcome in *PD* (*SH*) refers to a situation where both the subject and her partner in *PD* (*SH*) cooperate. RS-tests for the difference between *1Partner* and *2Partner* yield $p = .166$ when games are played simultaneously and $p = .270$ when sequentially.

Table 3.4 shows the relative frequencies of the outcome of the stage games in rounds 3 and further of each supergame. In *Simultaneous*, reaching the efficient outcome in both stage games at the same time is now rare and almost as likely in both treatments, meaning that many deviations occur once the continuation probability is introduced. However, the relative frequency of reaching the efficient outcome in neither of the two stage games is much larger in *1Partner* than *2Partner*. Whereas this is the modal outcome in *1Partner* (0.57), the modal outcome in *2Partner* is to reach the efficient outcome only in *SH* (0.54). The difference between *1Partner* and *2Partner* is now statistically significant (RS-test, $p = .047$). It leads to the following result:

Result 2: When the prisoner's dilemma and the stag hunt game are played simultaneously within a round, multigame contact leads to less efficient outcomes.
Comparing 1Partner with 2Partner in Simultaneous, we find that the efficient outcome in the prisoner's dilemma is not reached more frequently in the former than the latter and coordination on the payoff dominant Nash equilibrium in the stag hunt game is less likely in 1Partner.

In *Sequential*, the results are a bit different with *1Partner* resulting more often than *2Partner* in an efficient outcome in both *PD* and *SH* (0.14 vs. 0.08) or in neither (0.52 vs. 0.34). The difference between the two treatments is statistically significant (RS-test, $p = .003$). The fact that—with multigame contact—*Sequential* results more often than *Simultaneous* in the efficient outcome in *PD*, especially in the early rounds, can be explained by the lower critical discount factor required for reaching the efficient outcome in the former than the latter. In this case, this somewhat counterbalances the reduction in coordination on the payoff dominant Nash equilibrium in *SH*.

Result 3: When the prisoner's dilemma and the stag hunt game are played sequentially within a round, multigame contact is a double-edged sword.
Comparing 1Partner with 2Partner in Sequential, we find that reaching the efficient outcome in the prisoner's dilemma and the stag hunt game at the same time or in neither of those are both more likely in the former than the latter.

Table 3.4: Outcome of the stage games in rounds 3 and further of each supergame

Rounds 3 and further	*Simultaneous*		*Sequential*	
	1Partner	*2Partner*	*1Partner*	*2Partner*
Efficient outcome ...				
... in *PD* and *SH*	0.10	0.09	0.14	0.08
... only in *PD*	0.01	0.03	0.01	0.03
... only in *SH*	0.31	0.54	0.33	0.55
... in neither	0.57	0.35	0.52	0.34
N	2,086	4,746	2,322	4,486

Notes: Relative frequencies of the outcome of the stage games in rounds 3 and further of each supergame. A cooperative outcome in *PD* (*SH*) refers to a situation where both the subject and her partner in *PD* (*SH*) cooperate. RS-tests for the difference between *1Partner* and *2Partner* yield $p = .047$ when games are played simultaneously and $p = .003$ when sequentially.

To better understand the mechanisms leading to Results 2 and 3, we make use of Table 3.5 and investigate the effects of cooperation in the prisoner's dilemma breaking down. We condition the analysis on having reached a cooperative outcome in *PD* (C, C) and in *SH* $(stag, stag)$ in the first round of a supergame (first row).[16] As already observed, reaching a cooperative outcome in both games in the first round is less frequent in *1Partner* than *2Partner* in *Simultaneous* (RS-test, $p = .097$). In *Sequential*, coordination on the two efficient outcomes in the first round happens slightly more often in *1Partner* than *2Partner* but the difference is far from statistical significance (RS-test, $p = .570$).

[16]For this whole analysis, we decided not look at supergames in which cooperation in *PD* emerge later than the first round or at supergames starting with cooperation in *PD* but no coordination on $(stag, stag)$ in *SH*. Indeed, such occurrences are very rare encompassing only one to eight percent of the supergames depending on the treatment and they would unnecessarily complicate the analysis.

The second row shows the shares of supergames in which cooperation in *PD* breaks down in later rounds when we condition on having reached (C, C) and $(stag, stag)$ in the first round. In *Simultaneous*, we see that cooperation in *PD* is significantly less likely to break down in *1Partner* than *2Partner* (RS-test, $p = .009$). This suggests that multigame contact may increase the stability of cooperation in *PD*. The point estimate for *1Partner* is also lower in *Sequential* (0.74 vs. 0.80) but the difference is much lower and far from statistical significance (RS-test, $p = .274$).

Table 3.5: Effects of cooperation in the prisoner's dilemma breaking down

	Simultaneous		*Sequential*	
	1Partner	*2Partner*	*1Partner*	*2Partner*
(C, C) and $(stag, stag)$ in round 1...	0.19^{+}	0.31^{+}	0.40	0.33
...coop. in *PD* stops in round $t > 1$	0.54^{**}	0.80^{**}	0.74	0.81
Coop. in *PD* stops in round t ...				
... $(stag, stag)$ in SH at t	0.80	0.86	0.36^{**}	0.80^{**}
... $(stag, stag)$ in SH after t	0.56	0.75	0.40^{**}	0.76^{**}
... (C, C) in PD after t	0.05	0.04	0.04	0.06

Notes: The first row shows the relative frequencies of a cooperative outcome in *PD* and in *Sh* in the first round 1 of a supergame. A cooperative outcome in *PD* (*SH*) refers to a situation where both the subject and her partner in *PD* (*SH*) cooperate. The second row shows the fraction of supergame starting with (C, C) and $(stag, stag)$ in which cooperation stops in *PD*. The last three rows show the relative frequencies of $(stag, stag)$ and (C, C) after a breakdown of cooperation in *PD*. RS-tests for the difference between *1Partner* and *2Partner*: $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Under multigame contact, we assumed that cooperation breaking down in *PD* should lead to coordination on the payoff dominated equilibrium $(hare, hare)$ in *SH* as a punishment. In *Simultaneous*, the punishment should start in the following period and should continue until the end of the supergame. In *Sequential*, the punishment should already start in the same period. The third row of Table 3.5 shows the relative frequencies of $(stag, stag)$ in round t where t is the round when cooperation in *PD* stops. As expected, we see high relative frequencies for both *1Partner* and *2Partner* in *Simultaneous*. Indeed, we expect a subject who wants to deviate in *PD* at t to still choose $stag$ in the round. In *Sequential*, we see that cooperation breaking down in *PD* leads directly to less coordination on the payoff dominant equilibrium $(stag, stag)$ (0.36 vs. 0.80; RS-test, $p = .004$). Looking at the rounds after t (fourth row), we see a similar effect in *Simultaneous* although not statistically significant (RS-test, $p = .136$). Coordination on $(stag, stag)$ stabilizes around 40 percent (*1Partner*) and 76 percent (*2Partner*) in *Sequential* after t (RS-test, $p = .004$). The last row of Table 3.5 shows that once cooperation breaks down in *PD*, it is very rare that subjects manage to reestablish cooperation.

The differences in the relative frequencies of $(stag, stag)$ between *1Partner* and *2Partner* demonstrate that at least some subjects do link the two games by using *SH* to punish deviations in *PD*. However, this threat may not be strong or credible enough to deter deviations. Using *2Partner* as a benchmark, we see that the additional probability of departing from $(stag, stag)$ after cooperation breaking down in *PD* is approximately 25 percent in *Simultaneous* and 50 percent in *Sequential* for *1Partner*.[17] Result 2 suggests that the threat of departing from $(stag, stag)$ in *Simultaneous* is not credible enough to

[17] *Simultaneous*: $0.25 \approx 1 - \frac{56}{75}$ after t. *Sequential*: $0.55 \approx 1 - \frac{36}{80}$ at t and $0.47 \approx 1 - \frac{40}{76}$ after t

discipline behavior in *PD* and may even have adverse effects in terms of efficiency. In *Sequential*, we observed a double-edged sword, which suggests that the threat of departing from $(stag, stag)$ has some ability to discipline behavior in *PD* but not enough to be beneficial in terms of efficiency.

3.5 Conclusion

From a theoretical point of view, interacting in multiple games with the same partner—what we call multigame contact—can expand the set of possible cooperative equilibria compared with situations where players face a different partner in each game. The kind of strategies usually used to achieve such equilibria are *trigger* strategies and make use of a punishment threat, which is triggered when a player deviates from the agreed plan of action. Intuitively, multigame contact help to achieve such equilibria because interacting in multiple games with the same partner usually expands, but never decreases, a player's ability to punish. Facing a stronger punishment, therefore, means that deviations become less attractive and players can reach cooperative equilibria at lower discount factors.

In this experiment, we pair two indefinitely repeated games: a prisoner's dilemma and a stag hunt game. Our treatment of interest is the difference in subjects' behavior between situations with and without multigame contact. Multigame contact occurs when a subject plays the two indefinitely repeated games with the same partner and there is no multigame contact if a subject plays each indefinitely repeated game with a different partner. By its nature, the stag hunt game has the potential to help us better understand whether subjects can benefit from this enhanced ability to punish. Indeed, the stag hunt game has two obvious Nash equilibria in pure strategies, one leading to high payoffs and the other one to low payoffs, which can be used to discipline others' behaviors. We also compare situations in which the two games are played simultaneously within each round with situations in which the prisoner's dilemma is played before the stag hunt game within each round. This variation should—in theory—have implications in situations of multigame contact but no implications without.

In our data, we do not observe that multigame contact increases efficiency, measured by the frequencies of cooperation in the prisoner's dilemma and coordination on the payoff-dominant equilibrium in the stag hunt game. When the two games are played simultaneously within a round, multigame contact is even detrimental. It does not help to increase cooperation in the prisoner's dilemma while it increases the frequency of coordination on the low payoff equilibrium in the stag hunt game. When games are played sequentially within a round—the treatment that should lead to the most cooperation with multigame contact—we replicate the "double-edged sword" found in Laferrière, Montez, et al. (2022); that is, reaching the efficient outcome in the prisoner's dilemma and the stag hunt game at the same time or in neither of those are more frequent with multigame contact. In this case, the beneficial and detrimental effects of multigame contact somewhat average out. Overall, there is evidence that at least some subjects punish across games—i.e., subjects link the two games—with multigame contact, but this does not help to increase cooperation in the prisoner's dilemma.

It is difficult to explain why cooperation fails to increase with multigame contact in our experiment, although, at least some, subjects seem to understand that they could benefit from it. Without multigame contact, we observe low frequencies of cooperation in the prisoner's dilemma and high frequencies of coordination on the high payoff equilibrium in the stag hunt game. Therefore, a subject in a situation of multigame contact may have a lot to lose in the stag hunt game if she decides to use it as a punishment mechanism for the prisoner's dilemma. In theory, this is ideal as the more there is to lose in the stag hunt game—i.e., the stronger the punishment is— the less attractive becomes a deviation in the prisoner's dilemma. However, it only works if subjects think that punishment across games is likely. Since subjects do not always punish across games with multigame contact, such strong punishments become less credible, therefore, reducing the beneficial effects of multigame contact. When games are played simultaneously within a round, we found that multigame contact is even detrimental to cooperation. If we assume that subjects find punishment across games unlikely and that some subjects still punish, then multigame contact does not increase cooperation and reduces coordination on the payoff dominant equilibrium in the stag hunt game. With simultaneous games, the chances of being punished is not only function of the partner's behavior, but also of the likelihood of moving to a next round. When games are played sequentially, the chances of being punished only depends on the partner's behavior as the stag hunt game is played with certainty after the prisoner's dilemma. Therefore, the chances of being punished in the sequential case increase compared with the simultaneous case, making deviations less attractive. This could explain why we see some beneficial effects of multigame contact in the sequential variant.

For games with symmetric payoffs, a subject who punishes her partner reduces not only the partner's future payoffs but also her own; the stronger the punishment, the more future payoffs she has to sacrifice. At some point, the punishment could even be so strong that it becomes not credible. Therefore, there may be a trade-off between a punishment's strength and its credibility. If true, then situations with moderate punishments may lead to more cooperation than situations with strong possible punishments.

3.A Appendix

3.A.1 Additional tables

Table 3.A.1: Summary of the sessions

	Simultaneous		*Sequential*	
	1Partner	*2Partner*	*1Partner*	*2Partner*
Sessions	3	6	3	6
Matching groups	6	6	6	6
Subjects	52	118	58	112
Decisions	4,290	9,682	4,778	9,182

Notes: Number of sessions, matching groups, subjects, and decisions by treatment. We count one decision each time a subject has to take a decision in both games, i.e. this is equivalent to the total number of rounds played by all the subjects.

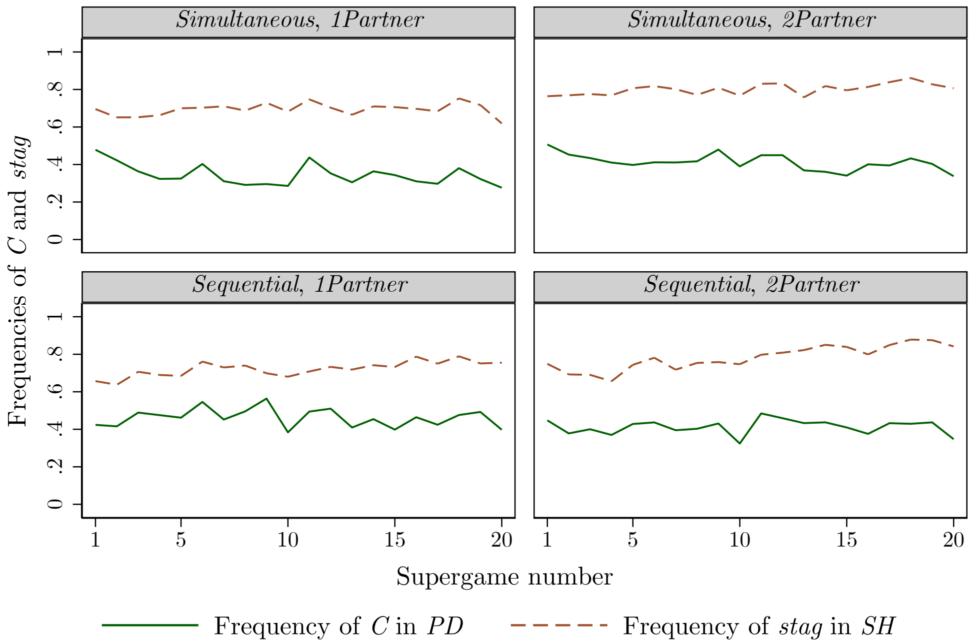


Figure 3.A.1: Frequencies of C and $stag$ over time by treatment. The mean frequencies are computed using the matching group averages for each supergame. We show the first 20 supergames, which is the minimum total number of supergames played by each matching group.

Table 3.A.2: Supergames' duration by matching group

Treat.		Match. gr.		Supergames duration						Statistics supergames					
Sim.	*2P*	ID	Subj.	3	4	5	6	7	8+	N	Mean	Min	Max	Total	Finite
1	0	31	8	15	5	2	1	0	0	23	3.5	3	6	81	0
		32	8	9	6	2	2	1	0	20	4.0	3	7	80	1
		41	10	10	5	4	2	0	0	21	3.9	3	6	82	1
		42	8	9	4	5	1	0	1	20	4.2	3	9	83	0
		91	10	15	3	3	0	0	1	22	3.8	3	11	83	3
		92	8	12	2	3	2	1	1	21	4.1	3	8	86	0
	1	11	16	15	5	2	1	0	0	23	3.5	3	6	81	0
		21	20	9	6	2	2	1	0	20	4.0	3	7	80	0
		51	20	10	5	4	2	0	0	21	3.9	3	6	82	0
		61	22	9	4	5	1	0	1	20	4.2	3	9	83	0
		71	20	14	3	3	0	0	1	21	3.8	3	11	80	0
		81	20	12	2	3	2	1	1	21	4.1	3	8	86	0
0	0	131	10	15	5	2	1	0	0	23	3.5	3	6	81	0
		132	10	9	6	2	2	1	0	20	4.0	3	7	80	1
		141	10	10	5	4	2	0	0	21	3.9	3	6	82	1
		142	10	9	4	5	1	0	1	20	4.2	3	9	83	0
		171	10	15	3	3	0	0	1	22	3.8	3	11	83	3
		172	8	12	2	3	2	1	1	21	4.1	3	8	86	0
	1	101	18	15	5	2	1	0	0	23	3.5	3	6	81	0
		111	20	9	6	2	2	1	0	20	4.0	3	7	80	0
		121	18	10	5	4	2	0	0	21	3.9	3	6	82	0
		151	20	9	4	5	1	0	1	20	4.2	3	9	83	0
		161	18	14	3	3	0	0	1	21	3.8	3	11	80	0
		181	18	12	2	3	2	1	1	21	4.1	3	8	86	0

Notes. Summary of the supergames duration for each matching group. The first two columns show the treatment played. *2P* is a binary variable taking the value 1 for *2Partner* and 0 for *1Partner*. Sim. is a binary variable taking the value 1 if the *PD* and *SH* games were played simultaneously and 0 if they were played sequentially. The last digit of a matching group ID is its number in the session and the other digits form the session ID (e.g., 31 refers to matching group 1 in session 3 and 141 refers to matching group 1 in session 14). There are up to 2 matching groups within a session. Subj. is the number of subjects in the matching group. Columns 5 to 10 show frequencies of supergames duration for six selected intervals. Columns 11 to 15 show summary statistics for supergame duration.

3.A.2 Experimental instructions

The instructions were originally written in French. Depending on the treatment, minimum changes were made to the instructions. Below you can see the translated instructions for the treatment *2Partner* when both games are played simultaneously.

Instructions

General information

You are going to participate to a study financed by the Swiss National Science Foundation (SNSF). Depending on your decisions, you will have the opportunity to earn a substantial amount of money. Please read the following instructions carefully.

These instructions are exclusively reserved for your usage. You are not allowed to communicate with the other participants. If you violate this rule, you will be banned from the experiment and receive no payment.

Throughout the study, we will not speak in CHF but in points. At the end of the study, your gains will be converted to CHF. The exchange rate between CHF and points is CHF 1 = 1000 points. Once the study is finished, you will receive your gains in cash plus a show-up fee of CHF 15.

The study is divided into matches. For each match, you are paired with two other randomly drawn participants in the room. These participants are called your partners. You will interact with these same two partners for several rounds. We will see later what determines the length of a match. Once a match is over, two new partners are randomly drawn. The figure below shows the difference between matches and rounds:

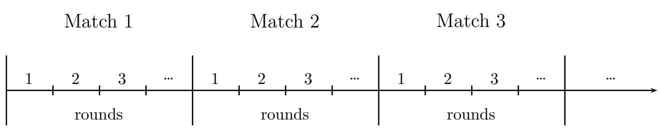


Figure 1: Matches and rounds

Your identity will never be revealed and you will never receive information about your partners.

You will play several matches and your partners will change between each match. You do not know how many matches you will play.

Rules of the game

Below, you can see the decisions screen. The header shows the current match number as well as the round number in the current match. Here you have the example of round 1 in match 2.

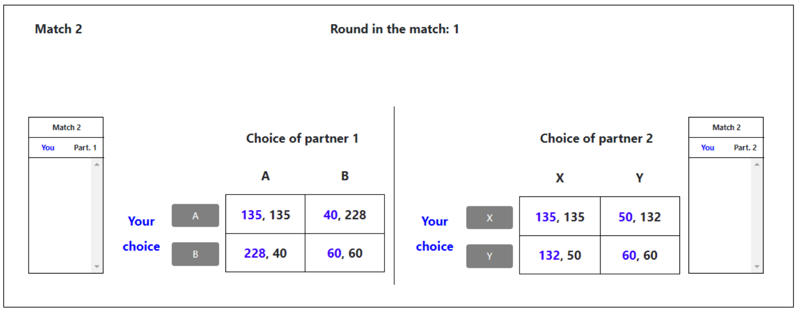


Figure 2: Decisions screen

The body of the screen is divided into two parts by a vertical line. At each round, you have two decisions to take. Specifically, you have to take one decision for the left part of the screen and one for the right part. The two tables in the middle show the possible gains for you and your partners. The decision for each table consists of choosing between the first line and the second line. In the left table, click on the gray button **A** or **B** to choose the first or second line. The decision is similar for the right part clicking the gray button **X** or **Y**. Your partner 1 does the same either choosing column **A** or column **B** on the left part and your partner 2 either chooses column **X** or column **Y** on the right part.

Each table contains four cells. The first number in blue of each cell is your gain for the round if this cell is the result of your decision and the one of your partner. The second number in black of each cell is your partner's gain. The following lines shows the four possible cases for the table on the left.

- You - **A** / Partner 1 - **A** $\rightarrow$ You - 135 Points / Partner 1 - 135 Points
- You - **A** / Partner 1 - **B** $\rightarrow$ You - 40 Points / Partner 1 - 228 Points
- You - **B** / Partner 1 - **A** $\rightarrow$ You - 228 Points / Partner 1 - 40 Points
- You - **B** / Partner 1 - **B** $\rightarrow$ You - 60 Points / Partner 1 - 60 Points

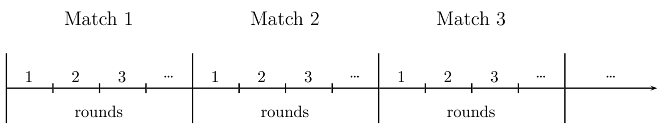
The reasoning is similar for the table on the right.

The tables on the very left and right parts of the screen remind you of your decisions and those of your respective partner for each half of the screen. Only the decisions of the current match are shown. Since the example is for round 1, the summary tables are still empty.

To take your decisions, you have to click on the gray buttons for each of the two tables in the middle. By clicking on a button, it becomes blue. Once you have taken your two decisions, a green button "Validate" appears in the lower right corner. By clicking this button, you move to the results screen. This screen will inform you about the choice of each of your partners. The results will be highlighted and your gain for each part will be displayed. If the match continues, you move to the next decisions screen and play the same game with the same two partners. If the match ends, a new screen will appear and inform you that two new partners will be randomly drawn.

Length of a match

Finally, we are going to look what determines the length of a match.



A match lasts at least 3 rounds. That means you are going to interact at least three times in a row with the same two partners.

From round 3 on, the match will stop randomly. More precisely, the match can stop at the end of round 3 with a probability of 1 chance out of 2. If the game does not stop, you move to a round 4 and there is again 1 chance out of 2 the match will stop at the end of round 4. The reasoning is identical for rounds 5 on, the match stopping at the end of each round with 1 chance out of 2. The computer randomly determines the stopping of a match.

To summarize, a match lasts at least 3 rounds. Starting from the end of round 3, the match stops at the end of each round with a probability of 1 chance out of 2.

The results screen will inform you whether the match continues or stops. Once a match is over, you move to the next one. As a reminder, two new partners are randomly drawn for the next match.

Make sure you understand the instructions. If something is not clear, please raise your hand and the organizer will come to help.

References

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science*, *46*(11), 1497–1512.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école americaine. *Econometrica*, *21*(4), 503–546.

Amaldoss, W., & Rapoport, A. (2009). Excessive expenditure in two-stage contests: Theory and experimental evidence. In I. N. Haugen & A. S. Nilsen (Eds.), *Game theory: Strategies, equilibra and theorems* (pp. 241–266). Nova Science Publishers.

Aoyagi, M., Bhaskar, V., & Fréchette, G. R. (2019). The impact of monitoring in infinitely repeated games: Perfect, public, and private. *American Economic Journal: Microeconomics*, *11*(1), 1–43.

Armantier, O., & Treich, N. (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, *62*, 17–40.

Åstebro, T., Herz, H., Nanda, R., & Weber, R. A. (2014). Seeking the roots of entrepreneurship: Insights from behavioral economics. *Journal of Economic Perspectives*, *28*(3), 49–70.

Åstebro, T., Mata, J., & Santos-Pinto, L. (2015). Skewness seeking: Risk loving, optimism or overweighting of small probabilities? *Theory and Decision*, *78*(2), 189–208.

Aumann, R. (1990). Nash equilibria are not enforceable. *Economic Decision Making (JJ Gabszewitz et. al. Eds.)*, 201–206.

Barrett, S. (1994). Strategic environmental policy and international trade. *Journal of Public Economics*, *54*(3), 325–338.

Bednar, J., Chen, Y., Liu, T. X., & Page, S. (2012). Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games and Economic Behavior*, *74*(1), 12–31.

Bernheim, B. D., & Whinston, M. D. (1990). Multimarket Contact and Collusive Behavior. *RAND Journal of Economics*, *21*(1), 1–26.

Bigoni, M., Casari, M., Skrzypacz, A., & Spagnolo, G. (2015). Time horizon and cooperation in continuous time. *Econometrica*, *83*(2), 587–616.

Blackhurst, R., & Subramanian, A. (1992). Promoting multilateral cooperation on the environment. *The Greening of World Trade Issues*, 247–268.

Blonski, M., Ockenfels, P., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, *3*(3), 164–92.

Blonski, M., & Spagnolo, G. (2015). Prisoners' other dilemma. *International Journal of Game Theory*, *44*(1), 61–81.

Brilon, S., Grassi, S., Grieder, M., & Schulz, J. (2019). Strategic competition and self-confidence. *GMU Working Paper in Economics No. 19-40*.

Brookins, P., & Ryvkin, D. (2014). An experimental study of bidding in contests of incomplete information. *Experimental Economics*, *17*(2), 245–261.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010a). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, *78*(4), 1375–1412.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010b). Supplement to "Risk and rationality: Uncovering heterogeneity in probability distortion". *Econometrica*, *78*(4), 1375–1412.

Bruhin, A., Santos-Pinto, L., & Staubli, D. (2018). How do beliefs about skill affect risky decisions? *Journal of Economic Behavior & Organization*, *150*, 350–371.

Brünner, T., Reiner, J., Natter, M., & Skiera, B. (2019). Prospect theory in a dynamic game: Theory and evidence from online pay-per-bid auctions. *Journal of Economic Behavior & Organization*, *164*, 215–234.

Bruttel, L. (2009). The critical discount factor as a measure for cartel stability? *Journal of Economics*, *96*(2), 113–136.

Busse, M. R. (2000). Multimarket contact and price coordination in the cellular telephone industry. *Journal of Economics & Management Strategy*, *9*(3), 287–320.

Cain, D. M., Moore, D. A., & Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, *36*(1), 1–18.

Camera, G., & Casari, M. (2009). Cooperation among strangers under the shadow of the future. *American Economic Review*, *99*(3), 979–1005.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*(1), 306–318.

Carpenter, J. P. (2007). The demand for punishment. *Journal of Economic Behavior & Organization*, *62*(4), 522–542.

Cason, T. N., Masters, W. A., & Sheremeta, R. M. (2020). Winner-take-all and proportional-prize contests: Theory and experimental results. *Journal of Economic Behavior & Organization*, *175*, 314–327.

Chen, D. L., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Ciliberto, F., & Williams, J. W. (2014). Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry. *RAND Journal of Economics*, *45*(4), 764–791.

Coccorese, P., & Pellecchia, A. (2009). Multimarket contact and profitability in banking: Evidence from Italy. *Journal of Financial Services Research*, *35*(3), 245–271.

Cooper, D. J., & Kühn, K.-U. (2014). Communication, renegotiation, and the scope for collusion. *American Economic Journal: Microeconomics*, *6*(2), 247–278.

Dal Bó, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review*, *95*(5), 1591–1604.

Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games. *American Economic Review*, *101*(1), 411–429.

Dal Bó, P., & Fréchette, G. R. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, *56*(1), 60–114.

Dal Bó, P., Fréchette, G. R., & Kim, J. (2021). The determinants of efficient behavior in coordination games. *Games and Economic Behavior*, *130*, 352–368.

Danz, D. (2020). Never underestimate your opponent: Hindsight bias causes overplacement and overentry into competition. *Games and Economic Behavior*, *124*, 588–603.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, *18*(4), 609–669.

Dohmen, T., Falk, A., Fliessbach, K., Sunde, U., & Weber, B. (2011). Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics*, *95*(3-4), 279–285.

Dorfman, A., Bereby-Meyer, Y., & Moran, S. (2013). When feeling skillful impairs coordination in a lottery selection task. *PLoS ONE*, *8*(6), 6–11.

Dreber, A., Rand, D. G., Fudenberg, D., & Nowak, M. A. (2008). Winners don't punish. *Nature*, *452*(7185), 348–351.

Duffy, J., & Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics*, *21*(3), 573–600.

Duffy, J., & Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, *66*(2), 785–812.

Dunne, T., Roberts, M. J., & Samuelson, L. (1988). Patterns of firm entry and exit in U.S. manufacturing industries. *RAND Journal of Economics*, *19*(4), 495–515.

Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1637), 871–878.

Ernst, C., & Thöni, C. (2013). Bimodal bidding in experimental all-pay auctions. *Games*, *4*(4), 608–623.

Evans, W. N., & Kessides, I. N. (1994). Living by the "golden rule": Multimarket contact in the us airline industry. *Quarterly Journal of Economics*, *109*(2), 341–366.

Falk, A., Fischbacher, U., & Gächter, S. (2013). Living in two neighborhoods—social interaction effects in the laboratory. *Economic Inquiry*, *51*(1), 563–578.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fehr-Duda, H., De Gennaro, M., & Schubert, R. (2006). Gender, financial risk, and probability weights. *Theory and Decision*, *60*(2-3), 283–313.

Fehr-Duda, H., & Epper, T. (2012). Probability and risk: Foundations and economic implications of probability-dependent risk preferences. *Annual Review of Economics*, *4*, 567–593.

Fehr-Duda, H., Epper, T., Bruhin, A., & Schubert, R. (2011). Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, *78*(1-2), 14–24.

Feinberg, R. M., & Sherman, R. (1985). An experimental investigation of mutual forbearance by conglomerate firms. *Industry Structure and Performance*, 139–166.

Feinberg, R. M., & Sherman, R. (1988). Mutual forbearance under experimental conditions. *Southern Economic Journal*, 985–993.

Fennema, H., & Wakker, P. (1997). Original and cumulative prospect theory: A discussion of empirical differences. *Journal of Behavioral Decision Making*, *10*(1), 53–64.

Fernandez, N., & Marin, P. L. (1998). Market power and multimarket contact: Some evidence from the Spanish hotel industry. *Journal of Industrial Economics*, *46*(3), 301–315.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Fischbacher, U., & Thöni, C. (2008). Excess entry in an experimental winner-take-all market. *Journal of Economic Behavior & Organization*, *67*(1), 150–163.
Frank, R. H., & Cook, P. J. (1995). *The winner-take-all society*. The Free Press.
Fréchette, G. R. (2012). Session-effects in the laboratory. *Experimental Economics*, *15*(3), 485–498.
Fréchette, G. R., & Yuksel, S. (2017). Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics*, *20*(2), 279–308.
Freitag, A., Roux, C., & Thöni, C. (2021). Communication and market sharing: An experiment on the exchange of soft and hard information. *International Economic Review*, *62*(1), 175–198.
Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review*, *102*(1), 337–363.
Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, *38*(1), 1–12.
Gächter, S., & Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, *13*(3), 364–377.
Ghemawat, P., & Thomas, C. (2008). Strategic interaction across countries and multinational agglomeration: An application to the cement industry. *Management Science*, *54*(12), 1980–1996.
Ghidoni, R., & Suetens, S. (2020). The effect of sequentiality on cooperation in repeated games. *American Economic Journal: Microeconomics*.
Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2002). Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory*, *104*(1), 247–272.
Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*(1), 97–113.
Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, *94*(2), 236–254.
Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*(1), 129–166.
Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.
Güth, W., Häger, K., Kirchkamp, O., & Schwalbach, J. (2016). Testing forbearance experimentally: Duopolistic competition of conglomerate firms. *International Journal of the Economics of Business*, *23*(1), 63–86.
Hamilton, B. H. (2000). Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy*, *108*(3), 604–631.
Harrison, G. W., & Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics*, *12*(2), 133–158.
Harsanyi, J. C. (1995). A new theory of equilibrium selection for games with complete information. *Games and Economic Behavior*, *8*(1), 91–122.
Harsanyi, J. C., & Selten, R. (1988). A general theory of equilibrium selection in games. *MIT Press Books*.
Heggestad, A. A., & Rhoades, S. A. (1978). Multi-market interdependence and local market competition in banking. *Review of Economics and Statistics*, 523–532.
Herbst, L. (2016). Who pays to win again? The joy of winning in contest experiments. *Working Paper of the Max Planck Institute for Tax Law and Public Finance*, *6*.

Herrmann, B., & Orzen, H. (2008). The appearance of homo rivalis: Social preferences and the nature of rent seeking. *CeDEx discussion paper series*.

Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do you put your money on it? *Economic Journal*, *115*(503), 305–318.

Jans, I., & Rosenbaum, D. I. (1997). Multimarket contact and pricing: Evidence from the us cement industry. *International Journal of Industrial Organization*, *15*(3), 391–412.

Kahneman, D. (1988). Experimental economics: A psychological perspective. In R. Tietz, W. Albers, & R. Selten (Eds.), *Bounded rational behavior in experimental games and markets. Lecture notes in economics and mathematical systems, vol 314* (pp. 11–18). Springer.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kartal, M., Müller, W., & Tremewan, J. (2021). Building trust: The costs and benefits of gradualism. *Games and Economic Behavior*, *130*, 258–275.

Keskin, K. (2016). Equilibrium notions for agents with cumulative prospect theory preferences. *Decision Analysis*, *13*(3), 192–208.

Krishna, V. et al. (1985). Finitely repeated games. *Econometrica*, *53*, 905–922.

Laferrière, V., Montez, J., Roux, C., & Thöni, C. (2022). Multigame contact: A double-edged sword for cooperation.

Laferrière, V., Staubli, D., & Thöni, C. (2022). Explaining excess entry in winner-take-all markets. *Management Science*.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization*, *17*(3), 377–400.

Mata, J., & Portugal, P. (1994). Life duration of new firms. *Journal of Industrial Economics*, *42*(3), 227–245.

Mengel, F., Orlandi, L., & Weidenholzer, S. (2022). Match length realization and cooperation in indefinitely repeated games. *Journal of Economic Theory*, *200*, 105416.

Metzger, L. P., & Rieger, M. O. (2019). Non-cooperative games with prospect theory players and dominated strategies. *Games and Economic Behavior*, *115*, 396–409.

Miller, A. R. (2010). Did the airline tariff publishing case reduce collusion? *Journal of Law and Economics*, *53*(3), 569–586.

Modak, M. (2021). Do multiple contacts matter? *Available at SSRN 3828250*.

Monroe, B. A. (2020). The statistical power of individual-level risk preference estimation. *Journal of the Economic Science Association*, *6*(2), 168–188.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes*, *103*(2), 197–213.

Moore, D. A., Oesch, J. M., & Zietsma, C. (2007). What competition? Myopic self-focus in market-entry decisions. *Organization Science*, *18*(3), 440–454.

Morgan, J., Orzen, H., Sefton, M., & Sisak, D. (2016). Strategic and natural risk in entrepreneurship: An experimental study. *Journal of Economics and Management Strategy*, *25*(2), 420–454.

Murnighan, J. K., & Roth, A. E. (1983). Expecting continued play in prisoner's dilemma games: A test of several models. *Journal of conflict resolution*, *27*(2), 279–300.

Nguyen, Q., Villeval, M. C., & Xu, H. (2016). Trust under the prospect theory and quasi-hyperbolic preferences: A field experiment in Vietnam. *Economic Development and Cultural Change*, *64*(3), 546–572.

Nikiforakis, N., & Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, *11*(4), 358–369.

Normann, H.-T., & Wallace, B. (2012). The impact of the termination rule on cooperation in a prisoner's dilemma experiment. *International Journal of Game Theory*, *41*(3), 707–718.

Parker, P. M., & Röller, L.-H. (1997). Collusive conduct in duopolies: Multimarket contact and cross-ownership in the mobile telephone industry. *RAND Journal of Economics*, 304–322.

Phillips, O. R., & Mason, C. F. (1992). Mutual Forbearance in Experimental Conglomerate Markets. *RAND Journal of Economics*, *23*(3), 395–414.

Phillips, O. R., & Mason, C. F. (1996). Market regulation and multimarket rivalry. *RAND Journal of Economics*, 596–617.

Porter, M. E. (1989). How competitive forces shape strategy. *Readings in strategic management* (pp. 133–143). Springer.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*(3), 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323–343.

Rao, J. N. K., & Scott, A. J. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Annals of Statistics*, *12*(1), 46–60.

Rapoport, A., Seale, D. A., Erev, I., & Sundali, J. A. (1998). Equilibrium play in large group market entry games. *Management Science*, *44*(1), 119–141.

Roth, A. E., & Murnighan, J. K. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical psychology*, *17*(2), 189–198.

Roux, C., & Thöni, C. (2015). Collusion among many firms: The disciplinary power of targeted punishment. *Journal of Economic Behavior & Organization*, *116*, 83–93.

Savikhin, A. C., & Sheremeta, R. M. (2013). Simultaneous decision-making in competitive and cooperative environments. *Economic Inquiry*, *51*(2), 1311–1323.

Schmidt, U., & Zank, H. (2008). Risk aversion in cumulative prospect theory. *Management Science*, *54*(1), 208–216.

Schmitt, M. (2018). Multimarket contact in the hospital industry. *American Economic Journal: Economic Policy*, *10*(3), 361–387.

Schulz, J. F., & Thöni, C. (2016). Overconfidence and career choice. *PLoS ONE*, *11*(1), e0145126.

Sheremeta, R. M. (2010). Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior*, *68*(2), 731–747.

Sheremeta, R. M. (2018). Impulsive behavior in competition: Testing theories of overbidding in rent-seeking contests. *Available at SSRN 2676419*.

Singal, V. (1996). Airline mergers and multimarket contact. *Managerial and Decision Economics*, *17*(6), 559–574.

Skyrms, B. (2001). The stag hunt. *Proceedings and Addresses of the American Philosophical Association*, *75*(2), 31–41.

Spagnolo, G. (1999a). On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion. *Journal of Economic Theory*, *89*(1), 127–139.

Spagnolo, G. (1999b). Social relations and cooperation in organizations. *Journal of Economic Behavior & Organization*, *38*(1), 1–25.

Spagnolo, G. (2001). Issue linkage, credible delegation, and policy cooperation. *CEPR Discussion Papers 2778*.

Stahl, D. O. (1991). The graph of prisoners' dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior*, *3*(3), 368–384.

Sundali, J. A., Rapoport, A., & Seale, D. A. (1995). Coordination in market entry games with symmetric players. *Organizational Behavior and Human Decision Processes*, *64*(2), 203–218.

Thomas, C. J., & Willig, R. D. (2006). The risk of contagion from multimarket contact. *International Journal of Industrial Organization*, *24*(6), 1157–1184.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Vespa, E., & Wilson, A. J. (2019). Experimenting with the transition rule in dynamic games. *Quantitative Economics*, *10*(4), 1825–1849.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Waldfogel, J., & Wulf, J. (2006). Measuring the effect of multimarket contact on competition: Evidence from mergers following radio broadcast ownership deregulation. *BE Journal of Economic Analysis & Policy*, *5*(1).

Wright, W. F., & Aboul-Ezz, M. E. (1988). Effects of extrinsic incentives on the quality of frequency assessments. *Organizational Behavior and Human Decision Processes*, *41*(2), 143–152.

Wu, B., & Knott, A. M. (2006). Entrepreneurial risk and market entry. *Management Science*, *52*(9), 1315–1330.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, *42*(12), 1676–1690.

Yang, J., Kawamura, T., & Ogawa, K. (2016). Experimental multimarket contact inhibits cooperation. *Metroeconomica*, *67*(1), 21–43.