European University Institute

DEPARTMENT OF ECONOMICS

Essays in Applied Game Theory

Irina Kirysheva

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 16 March 2015

European University Institute
Department of Economics

Essays in Applied Game Theory

Irina Kirysheva

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Fernando Vega-Redondo, Supervisor, Bocconi University
Prof. Piero Gottardi, EUI
Prof. Paolo Pin, Università degli Studi di Siena
Prof. Giovanni Ponti, Universidad de Alicante

“What is our life? A game!”

P.I.Tchaikovsky, opera ”Queen of Spades”

Acknowledgements

In this chapter I would like to thank all those who helped me during these five years. Without you I would never be able to do it.

Most importantly, I want to thank my supervisors, Fernando Vega-Redondo and Pierro Gottardi. It was a great journey during which I've learned a lot. I am grateful for your patience, for encouragement, for challenge, for trust. I want to thank you for all the valuable advice, for teaching how to ask questions and what's the best way to tackle them. I also want to thank Prof. Levine, Prof. Matozzi, Prof. Morelli, Prof. Rahman, Prof. Sutter for devoting your time when I came for your opinion. I wish I had more opportunities to discuss research with you.

I want also to thank Italian Ministry of Foreign Affairs for giving me this opportunity to do my PhD. Four yours I've spent in Italy and EUI were absolutely amazing. I am extremely grateful for this experience.

I especially want to thank Svetlana Pivovarova and Larisa Korableva for your help on my research. Hours, days and nights I've spent talking to you, guys, discussing ideas and models are priceless. I hope our research paths will continue to go together.

At this point I would also like to thank Arevik Mkrtchan, Hinnerk Gnutzman, Piotrek Spiewanoski, Kenan Huremovic, Ilya Archakov, Anastasya Antsygina and all participants of Network and Game Theory Working Group. Thank you folks for always being there for me both personally and research-wise.

I especially want to thank my coauthor, colleague, ex-flatmate, and dear friend Chara Papioti. I am so happy our roads crossed several times. Chara, without you this last year would be so much harder (as if it was not already tough enough). I hope our collaboration and friendship will continue despite all the kilometers that lie between us now.

Dear Jessica Spataro, Lucia Vigna, Marcia Gastaldo, Francoise Thauvin, Thomas Bourke, I want to thank you so much for your amazing work! I appreciate the way you made our life much easier.

I can not even try to name all my colleagues and friends from EUI who supported me during this way. Thank you all guys so much!

I am also grateful to ESC-Rennes school of business in France that gave me job for my fifth year of PhD. As economists we know that everything comes at a cost and there's no free lunch. I am grateful for experience you gave me, thought it came at the expense of my research work.

I also want to thank my mom. I know she worries a lot about all my undertakings. She was very helpful when sometimes I felt lost on the track. I know she's very proud of me and I hope to meet her expectations.

Finally, I am eternally grateful to my husband. His support was absolutely outstanding. He was always there for me in the minutes of doubt, helped me with some proofs, and constantly believed in me. This was a great journey for both of us and I am absolutely sure I would not be able to do it without him.

Abstract

My thesis covers different aspects of applied game theory.
The first paper looks at a two-sided asymmetric information game where agents make a collaborative decision not knowing each other’s types. In the model, an intermediary has full knowledge about the types of agents and can make a decision that brings information to some types. However, once he puts the information on the table the agents are not obliged to pay him, which undermines his incentive to participate in the first place. I find that, nevertheless, the intermediary is still welfare-improving.

In my second chapter I search for the optimal prize schemes in contests with sabotage. In the presence of sabotage, a standard prize scheme where the entire prize is given to the winner is no longer optimal as it creates very high incentives for sabotage. I show that in that case, an optimal prize structure may also assume a positive reward for contestants that are behind. With a higher number of contestants sabotage becomes a public good and therefore it is a lesser concern for the designer. In that case, when sabotage is expensive, the designer can achieve the first best by giving the whole sum to the winner. When I extend the problem to the continuous case the solution crucially depends on the cost of sabotage. When sabotage is expensive, the principal wants to give all of the prize to the winner, while when it is cheap he does not want to make a contest at all, and distributes all prizes equally.

In the third paper we analyze to what extent knowing game theory alters a persons' behavior. Our experiment showed a huge difference in results before and after the course. However results suggest that players behave less cooperatively not because of the knowledge of game theory per se, but due to changed expectations. We have also found that a course on game theory increases the level of reasoning.

Contents

List of Tables

f

Chapter 1

Introduction

This thesis covers different aspects of applied game theory. Two papers have in common the game-theoretical approach to model a real-life phenomenon. The third paper questions the influence game theory itself have on students performance and attitude in a classroom experiment.

The first paper *Informed Middlemen and Asymmetric Information* looks at two-sided asymmetric information game where agents make a collaborative decision not knowing types of each other. This structure describes the market for booking agents (intermediaries between bands and promoters), and generally for talent agents. In the model an intermediary has full knowledge about the types of agents and can make a decision that brings information to some types. However, once he puts the information on the table agents are not obliged to pay him, which undermines his incentive to participate in the first place. To obtain my results I develop the concept of PBE stable to bilateral deviations. I find that, nevertheless, the intermediary is still welfare-improving and restores efficiency. He either brings information to the most vulnerable type or to nobody. The situation is drastically different when I look at two informed intermediaries that compete in prices. In this case there is no equilibrium in pure strategies. Nonexistence in this sense is similar to that of screening models, although standard ways of dealing with it (e.g. reactive equilibrium concept) do not work here. Once the competition between intermediaries increases sufficiently, equilibrium reoccurs. In this equilibrium there is partial specialization between intermediaries - every pair of intermediaries sets a different wage and concentrates on a particular match.

In my second paper *Optimal Prize Allocation in Contests with Sabotage* I search for the optimal prize schemes in contests with sabotage. Contests are powerful mechanisms to induce the right incentives from the agents. In a contest with multiple participants particular prize distribution can allow a principal to maximize the expected effort he can get. In literature it is shown that if principal allocates positive prizes it is optimal to give all the sum to the leader. However, in real-life we see various contests that have multiple prizes. I consider possibility of sabotage in contests as a possible explanation. Under the presence of sabotage standard prize scheme is no longer optimal as it creates very high incentives to sabotage. I show that in that case optimal prize structure may also assume positive rewards for contestants that are behind. This result always holds in the case of two contestants. With higher number of contestants I differentiate between two sabotage protocols. With individual protocol sabotage becomes a public good and therefore it is a lesser concern for designer. In that case when sabotage is expensive designer can achieve the first best by giving the whole sum to the winner. Then I extend the model to continuous case. Here, the solution crucially depends on the cost of sabotage. When sabotage is expensive, principal wants to give all prize to the winner, while when it is cheap it does not want to make a contest at all, and distributes all prizes equally. This result explains one of the reasons why companies like Microsoft have given up "forced ranking" schemes.

My third paper *Teaching to be Selfish: Classroom Experiment on Prisoners Dilemma* is a join work with Chara Papioti. We have written this work during the year we have spent teaching at ESC-Rennes, France. In our paper we analyze to what extent knowing game theory alternates persons' behavior. We have conducted classroom experiments on Prisoners Dilemma during the course on Microeconomics that we were teaching. To identify the source of changed behavior we have conducted different treatments. Our experiment showed that indeed there is a huge difference in results before and after the course. However results suggest that players behave less cooperatively not because of the knowledge of game theory per se, but due to the change in expectations. We've also found that course on game theory increases the level of reasoning. Our last result also shows gender differences in expectations after the course. While female participants were expecting their partner to defect, males expected only cooperation.

Chapter 2

Informed middleman and asymmetric information

I.Kirysheva[1]

[1] irina.kirysheva@nu.edu.kz

2.1 Introduction

Economic literature long has been interested in the problems that include the presence of a third party in the transaction. This research applies to trade, flows of goods in networks, referrals in the job market and much more. However, despite the vast coverage, little has been said on the role of an intermediary as a bridge to overcome asymmetric information.

There are situations when a third party possesses information that is valuable for other agents. For example, a football agent has good knowledge both about the ability of potential players and the perspectives of different teams; booking agents (intermediaries between musicians and performance promoters) know both the potentials of the band and the peculiarities of different promoters. At the same time, players themselves cannot acquire this information or it might be extremely costly, and absence of information leads to inefficiencies. In this case the informed third party might have an incentive to participate in the ongoing transaction and eventually reveal some part of his knowledge. Obviously, an intermediary would like to use his informational advantage in the most valuable way. The paradox is that once he takes an action that brings something new to the agent the latter might use the already revealed information on his own and decide not pay to the intermediary in order to economize. This in turn changes the incentives of an intermediary to participate in the first place.

My main question in this setting is whether, given the above, the intermediary can still solve the inefficiency and how much he can gain from the informational advantage if the agents are not committed to using his service. Will agents always operate through an intermediary if he provides the most efficient way to collaborate? Finally, what is the optimal structure of the market of intermediaries - do agents benefit from competition between intermediaries?

In my model a pair of individuals think about collaborating with each other, making a partnership. As I want to model an informational part, I do not consider the moral hazard aspect of the partnership problem but assume that the success of the partnership crucially depends on types of both agents but not on their efforts. Suppose we have a partnership of two songwriters - one of them is a composer who writes music, and another is a lyricist. Songwriters might differ substantially in their types (it might be their ability, talent or experience in writing songs).The song is successful only in the case the both music and lyrics are good.

2.2 Literature review

The literature related to our model can be divided into two main groups according to different aspects it is capturing that are addressed in our work: there is literature on intermediaries in different economic situations, and then there is literature on informed principal.

In their paper on structural holes Goyal and Vega-Redondon in [10] explore motives for link formation when agents can either pay or extract rents from intermediation depending on their position in the network. They show that without capacity constraints a star network emerges, where the central player acts as an intermediary and enjoys significant rents from his position. On the contrary, with the presence of capacity constraints, cycle is the equilibrium network, with no one agent being an intermediary (as no one is essential for connecting any two others) and all getting the same payoffs.

A valuable group is literature on referrals. Montgomery in [23] examines the role of employee referrals on the labor market with adverse selection. The model is two-period, and workers can be of two types - high and low. Workers that are employed in the first period recommend those linked to them. The ability of connected workers is exogenously correlated. Montgomery shows that in equilibrium companies will hire only those second-period workers that were introduced by high ability employees. In the second period, workers that were hired through an acquaintance receive wages below their expected productivity, therefore a firm gets a positive expected profit. As firms are competitive, the wage they are paying to the first-period workers exceeds their expected productivity because first-period workers also have an optional value that can lead to a positive expected profit in the second period. Montgomery also investigates the impact of network structure on wage dispersion: an increase in either network density or correlation between the productivity of connected agents increases wage dispersion. However, in this model the reference decision of first-period workers is non-strategic - they always give references to any agents with whom they are connected.

Saloner, in [28] considers a dynamic model of references with reputation. There is more than one competing referee that has received a signal about the abilities of his candidates. On the one hand, each referee wants more of his candidates to be hired; one the other hand, he also cares about their average quality. Thus, there is a trade-off between recommending more friends or recommending fewer of them but of higher quality. Each referee uses a cut-off strategy. As a result of the model, although the referees act strategically, resulting equilibrium is efficient (so the result is the same as if the firm itself got the signals that the referees had).

Rubinstein and Wolinsky in [27] incorporate intermediaries in a bargaining and matching framework. In their work, a market consists of three types of agents - buyers, sellers and intermediaries. This model doesn't address the advantages of intermediation, as buyers and sellers meet a middleman by some exogenous process.

A large strand of literature considers the role of the intermediary in trade. Gehrig in [9] looks at the intermediary in the market with costly search. Buyers and sellers choose between direct trade or trade through an intermediary who purchases and sells products. The intermediary offers the service of immediacy by posting the bid and ask prices directly, thus allowing agents to avoid the costly search. The author finds that traders with low gains from trade are not willing to pay intermediary costs and will go for a direct trade. Monopolist intermediary will charge positive spread, while in case of competition the classical Bertrand result applies.

Stahl in [30] and Yanelle in [31] look at two-sided price competition. They find that non-Walrasian equilibria with positive bid-ask spreads may emerge, even when intermediation technology is costless because intermediaries offer attractive bid prices, and obtain a monopoly position towards buyers. Moreover, the existence of equilibrium may be problematic.

Garella in [8] looks at trade with asymmetric information, and finds that intermediation may complete the market system when asymmetric information causes failure without one. This result is obtained under the hypothesis that the intermediary randomizes the price offers to the seller.

As intermediary is the person who at the same time possesses information and proposes a contract this is related to the problem of informed principal (e.g.Mayerson in [25]; Maskin and Tirole in [19] and [20]; Severinov in [29]). However, in my model the main question is not the one of the contract design, and in fact intermediary is restricted to a very particular type of contracts that I am interested in.

Ivashina in [15] looks at the information intermediary role of banks in fortifying acquisitions. Banks through long-term relational contracts with their clients get access to their private information which they later might pass to potential acquires. The most plausible motive for this is that the banks try to transfer debt from weaker clients to a more stronger ones. Unlike their case I am looking at two-sided asymmetric information.

2.3 Basic Model

In the baseline model I look at the situation when the reward intermediary can ask for (I call it wage hereafter) is fixed exogenously. This is the first step on the way to a more realistic model where intermediary can decide on the amount of wage.

2.3.1 Description of the game

I am looking at the Bayesian game with three players: agent 1 (Ag_1), agent 2 (Ag_2), and an intermediary (Int). Two agents are thinking about collaborating with each other.

In the game potential collaborators can be either of high or low types ($\Theta = \{H,\, L\}$), probability of being high type is p. Types are assigned independently. If agents decide to collaborate, an individual success realizes for each of them independently with p_h or p_l depending on their type ($p_l < p_h < 1$). Agents know their own type but they do not know the type of each other which is the source of uncertainty for them. The collaboration is successful only in case both agents succeeded on their parts.

Intermediary has full information about players types. Observing this, he may decide to send what I call "an intermediation offer" (the content of it is specified later). At the next stage agents decide how they would like to collaborate with each other: either through an intermediary if he had offered such an option, or directly, or don't want at all.

The timing of the game is following:

Stage 0. Nature determines the agents types.

Stage 1. Intermediary observes agents types and decides whether he wants to send intermediation offers or not.

Stage 2. Agents receive or do not receive an offer and based on this decide how they want to be connected.

Stage 3. Links are formed.

For convenience I introduce the following notation: H+ denotes an information set where a high type agent receives an intermediation offer; L+ describes an information set where a low type individual have received an intermediation offer; H- for the information set where a high type individual hasn't received an offer; finally, L- stands for the information set in which a low type agent hasn't received any intermediation offer. Therefore, information sets $\mathcal{I}_{Ag_1} = \mathcal{I}_{Ag_2} = \mathcal{I} = \{H+,\, H-,\, L+,\, L-\}$.

I look for Perfect Bayesian Equilibrium that is stable to bilateral deviations.

Here I expanded the standard concept of PBE to allow also for stability with respect to bilateral deviations. Stability to bilateral deviation lies in the hart of the paradox described above that once agents are offered an intermediation offer they might reevaluate their decision and prefer to collaborate without any intermediary.

I consider bilateral deviation in the *ex ante* fashion so that agents would like to bilaterally deviate if it gives them higher expected payoff before knowing which type they are.

Definition 2.1. There is a profitable bilateral deviation from the strategy profile $(s^*_{Ag_1}, s^*_{Ag_2}, s^*_{Int})$ when for some $(i,\, j) \subset (Ag_1,\, Ag_2,\, Int)$ $\exists\, (s_i,\, s_j)$ ($s_i \in S_i,\, s_j \in S_j$) such that $EU_i(s_i, s_j, s^*_{-\{i,\,j\}}) \geq EU_i(s^*_i, s^*_j, s^*_{-\{i,\,j\}})$ and $EU_j(s_i, s_j, s^*_{-\{i,\,j\}}) \geq EU_j(s^*_i, s^*_j, s^*_{-\{i,\,j\}})$ with at least one of the inequalities being strict.

Definition 2.2. A strategy profile $s^* = (s^*_{Ag_1}, s^*_{Ag_2}, s^*_{Int})$ and a belief system $\mu^* = (\mu^*_{Ag_1}(\mathcal{I}), \mu^*_{Ag_2}(\mathcal{I}))$ constitute a symmetric Perfect Bayesian Equilibrium that is stable to bilateral deviations if

i) s^* is sequentially rational under μ^*

ii) there is no profitable bilateral deviation under μ^*

iii) μ^* is derived by a bayes rule whenever possible

and *iv*) $s^*_{Ag_1} = s^*_{Ag_2}$ and $\mu^*_{Ag_1}(\cdot) = \mu^*_{Ag_2}(\cdot)$.

2.3.2 Payoffs

There are three types of payoffs in this model: payoff agent gets when collaborates directly with another agent without any third party; what they can get when they accept the offer of an intermediary and collaborate through him; and finally, the payoff of an intermediary

Direct collaboration is always costly - agents have to invest in establishing new relationships, so they have to pay c irrespective of whether collaboration is successful or not. On the other hand, the gain of connection is random and depends on types of both agents. If two agents decided to collaborate directly the probability their collaboration is successful is $p_{Ag_1} p_{Ag_2}$. In case their collaboration turns out to succeed it brings participants the value of 2δ which they divide equally.

In the case agents chose to accept the offer of an intermediary, the latter offers a particular contract to the agents. Here, I restrict attention to the contract I am interested in. There an intermediary promises an agent to compensate him in case collaboration failed due to the fault of another agent. So, he makes a payment that is conditional on personal successes of the agents in their part of the project and not conditional on the success of collaboration as a whole. In case he was successful on his part of the project agent gets $\delta - w$, where w is what I denoted by a wage of an intermediary. The rest of the value generated goes to the intermediary. So, if the agent was successful (which happens with p_{Ag_1}), independently of what has happened to his partner this period, he gets $(\delta - w)$, and if he wasn't (with $1 - p_{Ag_1}$) he gets nothing.

Now, let's look at what an intermediary can get from his service. When agents accepted his service and both were successful (which happens with $p_{Ag_1}p_{Ag_2}$), intermediary gets $2w$ from them (w from each agent). However, if exactly one of the agents has failed intermediary pays to another $(\delta - w)$.

So, intermediary gets paid when the collaboration he is bridging is successful. In case collaboration failed but there was one agent who succeeded on his part of the project, intermediary compensates this agent. And if the project has failed because both agents failed with their parts an intermediary gets zero. This specific form of contract, though artificial at the first sight, describes what is going on in some areas of business where intermediary play crucial roles. According to the article of D. Starosta in "Clubbing space" magazine booking agents "...[G]uarantee the payment of forfeit to the artist, in case the concert didn't take place, and at the same time booker guarantees the performance to promoters, and if anything should provide another artist of the same style".

2.4 Results

There are many different possibilities for equilibria for different values of c and w. In some equilibria intermediary is inactive while in others he participates in the collaboration.

The region of special interest is area where $c \in [p_l\,(pp_h + (1-p)\,p_l)\,\delta;\; p_l p_h \delta]$. Here a cost of direct collaboration is quite high so that low type agents don't want to collaboration with ex ante unknowns type. At the same time it's not sufficiently high to rule out the temptation of low types to collaboration with a high type. This conflict results in no pure strategies equilibrium under such c in case there is no intermediary present.

2.4.1 No intermediary baseline

Theorem 2.3. *When $c \in [p_l(pp_h + (1-p)p_l)\delta, p_h p_l \delta]$ in the absence of an intermediary the only PBE stable to bilateral deviations is a mixed strategies PBE where high type agent always goes for direct link, and low type agent goes for a link with probability $\beta = \frac{p_h p_l \delta - c}{c - p_l^2 \delta}\frac{p}{1-p}$[2].*

In mixed strategy equilibrium case with probability β^2 an inefficient match of two low agents is formed, moreover with probability $(1-\beta)$ an efficient match of high and low agent is not formed - so, there are two sources of inefficiency that can not be avoided in this setting due to lack of information. Here, low type agent is the vulnerable one who needs more information about his potential partner.

2.4.2 Presence of an intermediary

Compared to a baseline game without an intermediary, intermediary game can lead to a drastically different result.

First, I look at different possibilities for bilateral deviations and conditions that bilateral deviation proofness require.

Suppose that in some strategy profile $\{s_1, s_2, s_{Int}\}$ intermediary makes an offer to matches $\mathcal{M} = \left\{\tilde{\Theta}_i \times \tilde{\Theta}_j\right\} \in \Theta \times \Theta$ and agents accept this offer. Then, for this strategy profile to be stable to bilateral deviations by two agents I have to check that there are no s_1', s_2' : $s_i'\left(\tilde{\Theta}_i, \text{no offer}\right) = s_i\left(\tilde{\Theta}_i, \text{no offer}\right)$, $s_i'\left(\Theta\backslash\tilde{\Theta}_i, \cdot\right) = s_i\left(\Theta\backslash\tilde{\Theta}_i, \cdot\right)$ but $s_i'\left(\tilde{\Theta}_i, \text{offer}\right) = \{\text{direct collaboration}\}$ for $i = \{1, 2\}$. There are four possibilities for $(\tilde{\Theta}_1,$

[2]Most of the profs are in Appendix

$\tilde{\Theta}_2$) that an intermediary can offer his service to: 1) $(\tilde{\Theta}_1 \times \tilde{\Theta}_2) = \{HH\}$, 2) $(\tilde{\Theta}_1 \times \tilde{\Theta}_2) = \{HH, HL, LH\}$, 3) $(\tilde{\Theta}_1 \times \tilde{\Theta}_2) = \{LL\}$, and 4) $(\tilde{\Theta}_1 \times \tilde{\Theta}_2) = \{HH, HL, LH, LL\}$. In the description of equilibria below I am covering all of these possibilities for bilateral deviations.

Another possibility for a bilateral deviation is that an intermediary and one of the agents deviate. For this to be a profitable deviation in must be that actions of two agents do not coincide in the information sets where they get the offer of an intermediary. It happens only when intermediary provides service to a match of LL, while all others are collaborating directly. In this equilibrium it should be that in $H+$ agent chooses to collaborate directly, otherwise the intermediary would have been happy to provide an offer to matches containing high type agents either. In case the cost of direct collaboration is sufficiently high the agent might prefer to change his action in $H+$ information set, and intermediary would then prefer to send an offer to more matches. The condition for no bilateral deviation is $2\left(p_h p_l \delta - c\right) > (\delta - w)\left(p_h + p_l\right)$.

Now I go through different regions for the wage of an intermediary and look at equilibria in these regions.

Theorem 2.4. *When $w \leq (1 - p_h)\,\delta$ the only PBE stable to bilateral deviations of an intermediary game is the same as that of a no-intermediary game where high type agent always goes for direct link, and low type agent goes for a link with $\beta = \frac{p_h p_l \delta - c}{c - p_l^2 \delta} \frac{p}{1-p}$.*

Obviously, when the wage of an intermediary is so low that he is not interested in participating even when he faces the best match possible, the situation is just the same as in the case of no intermediary.

Theorem 2.5. *When $w \in \left[(1 - p_h)\,\delta;\ \left(1 - \frac{2p_h p_l}{p_h + p_l}\right)\delta\right]$ and $c \geq p_h\left(w - (1 - p_h)\,\delta\right)$ in addition to a no intermediary equilibrium there is another one where a pair of high type agents HH collaborate through an intermediary, HL collaborate directly with probability $(1 - \beta)$, and LL do it with probability β^2.*

The result above states that when the wage of an intermediary is a bit higher (such that he is interested in working only with the best possible match) but not too high compared to the cost (so that even knowing for sure the type of their potential partner agents still want to make deal with an intermediary) there is another equilibrium possible where a pair of high type agents are collaborating indirectly, and all others behaving just as they did without an intermediary.

Theorem 2.6. *When $w \in \left[\left(1 - \frac{2p_h p_l}{p_h + p_l}\right)\delta;\ (1 - p_l)\,\delta\right]$ and $c > \frac{\delta p_h (p p_h + 2(1-p) p_l) - (\delta - w)(p_h + (1-p) p_l)}{p + 2(1-p)}$ in addition to no intermediary equilibrium there is another one where pairs of HH and HL collaborate through intermediary, while pair of LL does not collaborate at all.*

When the wage of an intermediary is even higher so that he is willing to work with any match except the worst one, there is an equilibrium when good and average matches (HH and HL) work through intermediary, while the bad match of LL does not collaborate at all which is actually good thing to do for the agents. In this equilibrium absence of the offer of an intermediary provides full information to low type agents about their partner and helps them to stay away from the inefficient match.

Theorem 2.7. *When* $w \geq (1-p_l)\,\delta$ *and* $(pp_h + (1-p)\,p_l)\,(\delta - w) \geq p^2\left(p_h^2\delta - c\right) + (1-p)^2\,p_l\,(\delta - w)$ *in addition to no intermediary equilibrium there is another one where every pair of agents collaborate indirectly through an intermediary.*

Theorem 2.8. *When* $w \geq (1-p_l)\,\delta$, $2\,(p_h p_l \delta - c) > (\delta - w)\,(p_h + p_l)$ *and* $c > p_l\,(w - (1-p_l)\,\delta)$ *in addition to no intermediary equilibrium there is another one where the pair of low type agents* LL *collaborates indirectly, while* HL *and* HH *choose direct collaboration.*

There are several equilibria with intermediary that are possible in the highest wage region - either all matches collaborate through an intermediary, or only the worst match is intermediated and all others are working directly.

These equilibria are more efficient than the no intermediary one. Situation when an intermediary works with the LL match allows to overcome both inefficiencies that arise without intermediary. First, bad matches of LL never work directly with each other, they only collaborate through an intermediary which is efficient. From the economy point of view collaboration through an intermediary is free, because neither agents nor intermediary do have to pay the cost of collaboration. So from the social planner perspective the best thing is if all matches work through an intermediary. Second, in this equilibrium the average match of HL agents always collaborate. This is again an improvement with respect to the case of no intermediary as in the latter case agents in HL match used to collaborate with probability less than one.

The next claim shows that equilibrium where everyone collaborate through an intermediary is more efficient (the total welfare is higher) than no intermediary equilibrium.

Claim 2.9. The total welfare is higher in the equilibrium where every match collaborates through an intermediary that in equilibrium with no intermediary at all.

Let me sum up what we've learned from this section: the higher is the wage of an intermediary, the more willing he is to work with worse matches. If his wage is low he works only with the best match HH; when the wage is higher he also works with average match HL; finally when his wage is really high he starts working with the bad match LL. The case of the highest wage results in a branching of equilibria into two - we might have the situation that intermediary works with all matches, or he works only with bad match.

2.5 Endogenous wage

Now when we have discussed what can happen in the game with predetermined wage I allow an intermediary to set his wage. Intermediary does it before knowing types of agents which seems to be a reasonable assumption - like a booking agent who has set wage for his service before having any knowledge of which are the bands and promoters he's going to work with.

2.5.1 Monopoly intermediary

The timing of the new game is following:

Stage 1. Intermediary decides on the wage w^* he is willing to ask for

Stage 2. Nature defines the combination of types of agents that is realized

Stage 3. Intermediary observes agents types and decides whether he wants to send intermediation offers or not.

Stage 4. Agents receive or do not receive an offer and based on this decide on how they want to collaborate. Even not haven received the offer agents see the wage w^* an intermediary has set.

Stage 5. Collaborations are/are not formed.

There are four possibility for an intermediary with respect to what matches he can aim for at the equilibrium: he can either work only with HH, or only with LL , or everybody except LL, and finally he can intermediate everyone. The wage he can set in each of the variants is defined by the bilateral deviation condition.

When he works only with HH he may set the wage as high as $w^*_{HH} = \frac{c}{p_h} + (1-p_h)\,\delta$. His expected payoff under this wage is $\pi_{HH} = 2cp^2$.

When he works only with LL he may set the wage as high as $w^*_{LL} = \frac{c}{p_l} + (1-p_l)\,\delta$. His expected payoff under this wage is $\pi_{LL} = 2c\,(1-p)^2$.

When an intermediary aims for everybody except LL pair (I denote this case by $\overline{LL}$) he can set the wage as high as $w^*_{\overline{LL}} = \frac{(2-p)}{p_h+(1-p)p_l}c + \frac{((1-p)p_l+p_h(1-2p_l(1-p)-pp_h))\delta}{p_h+(1-p)p_l}$. At this wage he gets $\pi_{\overline{LL}} = 2c(2-p)p$.

Finally, an intermediary may work with every type of agents. He sets the wage $w^* = \delta - \frac{p\left(p_h^2\delta-c\right)}{p_h+(1-p)p_l}$ and gets $\pi^* = 2pc\frac{\bar{p}}{p_h+(1-p)p_l} + \frac{2\bar{p}\delta(1-p)p_l(p_h+pp_h+p_l-pp_l)}{p_h+p_l-pp_l}$.

Even if intermediary knows for sure that agents are going to accept the offer he makes, he still might have different preferences over equilibria for various parameter values. The general pattern looks like this: when p is low, intermediary receives highest expected payoff in LL equilibrium[3]; when p is average, he gets the highest payoff in equilibrium where he works with all matches; finally, when p is high, he prefers $\overline{LL}$ equilibrium[4].

When intermediary sets some wage w he receives $\pi(w) = \pi(w,\, CE(w,\, c))$, where $CE(w,\, c) \in \mathcal{CE}(w,\, c)$ is some continuation equilibrium played in a game with the cost of direct connection c after an intermediary sets w; and $\mathcal{CE}(w,\, c)$ is the set of all possible continuation equilibria when cost is c and wage w. I denote by NIE a no-intermediary continuation equilibrium - an equilibrium played after wage w where agents do not use an intermediary and are not collaborating through him.

Claim 2.10. Under any c there is a unique equilibrium where intermediary 1) works with LL matches only and sets $\tilde{w}^* = w^*_{LL} = \frac{c}{p_l} + (1-p_l)\,\delta$ (when p is low); 2) works with all combinations of agents and sets $\tilde{w}^* = w^* = \delta - \frac{p\left(p_h^2\delta - c\right)}{p_h+(1-p)p_l}$ (when p is average); 3) works with all matches except LL and sets $\tilde{w}^* = w^*_{\overline{LL}} = \frac{(2-p)}{p_h+(1-p)p_l}c + \frac{((1-p)p_l+p_h(1-2p_l(1-p)-pp_h))\delta}{p_h+(1-p)p_l}$ (when p is high) iff $NIE \notin CE(\tilde{w}^*,\, c)$.

Proof. First, suppose that the strategy profile where intermediary sets $\tilde{w}^*$ is indeed a unique equilibrium in the game. Then it should be that $NIE \notin CE(\tilde{w}^*,\, c)$. Indeed, suppose that NIE was a part of continuation equilibrium when wage $\tilde{w}^*$ is set. In that case another wage (corresponding to the second largest expected payoff for the intermediary) would be also an equilibrium wage, as after deviation to $\tilde{w}^*$ NIE might be played.

Other way, suppose that $NIE \notin CE(\tilde{w}^*,\, c)$. Then indeed the only equilibrium is the one where intermediary sets $\tilde{w}^*$. It is obvious that $\tilde{w}^*$ is an equilibrium wage, as any deviation from it brings intermediary lower expected payoff. No other wage w' could be an equilibrium wage as there would be a profitable deviation to $\tilde{w}^*$ for an intermediary. □

Unfortunately, when $c \leq p_l p_h \delta$ $NIE \in \mathcal{CE}(\tilde{w}^*,\, c)$, which means that under this c equilibrium in monopoly intermediary case is not unique. As for $\forall w$ in this region $NIE \in \mathcal{CE}(w,\, c)$ any wage can be supported as an equilibrium.

[3] If $p_h \in \left[p_l;\, p_l\frac{3+\sqrt{5}}{2}\right]$, $p <$Root$[p_h^2 - p_l^2 + \left(-3p_h^2 - 4p_hp_l + 3p_l^2\right)x + \left(7p_hp_l - 3p_l^2\right)x^2 + (p_h^2 - 3p_hp_l + p_l^2)x^3,\, 1]$

If $p > p_l\frac{3+\sqrt{5}}{2}$, $p <$Root$[p_h^2 - p_l^2 + \left(-3p_h^2 - 4p_hp_l + 3p_l^2\right)x + \left(7p_hp_l - 3p_l^2\right)x^2 + (p_h^2 - 3p_hp_l + p_l^2)x^3,\, 2]$

[4] $p > \frac{p_hp_l + p_l^2}{p_h^2 - p_hp_l + p_l^2}$

Due to the multiplicity not much can be said about the outcome of the game in this case, and therefore some refinement is needed. I assume that if there is a Pareto inferior equilibrium it is never played. It turns out that NIE is a Pareto inferior equilibrium here, as for any wage that an intermediary sets there is an intermediary equilibrium that Pareto dominates NIE.

Theorem 2.11. *Under the refinement specified above the only equilibrium in the game with a monopoly intermediary is the one where intermediary 1) works with LL matches only and sets $\tilde{w}^* = w^*_{LL}$ when p is low; 2) works with all combinations of agents and sets $\tilde{w}^* = w^*$ when p is intermediate; 3) works with all matches except LL and sets $\tilde{w}^* = w^*_{\overline{LL}}$ when p is high*

When p is low intermediary makes an offer to a pair of low type agents. The offer provides full information to the low type agent about their potential partner - in both cases if they have or have not received the offer. This is important as low type agents that has not received an offer knows that they face a high type agent and happily go for direct partnership. In this equilibrium LL collaborate through intermediary, and all others do it directly.

When p is average intermediary makes an offer to all types of matches. This equilibrium is efficient as collaboration through an intermediary maximizes the total welfare. Here intermediary does not provide any additional information to the agent in equilibrium, and agents accept it as given the wage of the offer their no-bilateral-deviation constraint is still satisfied.

When p is high intermediary makes an offer to any combination of agents except the pair of low types. Agents that get an offer accept it, while agents that do not get the offer stay not connected. Here the lack of an offer provides full information for the low type agent about their potential partner, and helps to stay away of the inefficient matches with another low type agent. It is a welfare improvement compared to no intermediary case.

2.5.2 Two intermediaries

Suppose now that there are two potential intermediaries $\mathcal{N}_{Int} = \{\mathcal{I}^1, \mathcal{I}^2\}$. Before observing the realization of types each of them sets a wage w^*_i. After that intermediaries observe realization of the match, the wage their opponent has set and decide whether they want to propose their service or not. Then the game continues just as described for the case of exogenous w with the difference that now each agent has also to decide on which intermediary they would like to work with if any.

The timing of the game is following.

1. Both intermediaries simultaneously set their wages w_1 and w_2.

2. Nature realizes uncertainty about the match.

3. Intermediaries observe types of agents and the opponents wage, and decide whether they want to send offer or not ($\mathcal{A}_{Int_1}(\theta_{ag_1}, \theta_{ag2}, w_1, w_2)$, $\mathcal{A}_{Int_2}(\theta_{ag_1}, \theta_{ag2}, w_1, w_2) \in \{0, 1\}$).

4. Agents receive offers from two intermediaries or only from one, or from none of them. They decide on the offer if any. $\mathcal{A}_{ag_i}(\theta_i, \mathcal{A}_{Int_1}, \mathcal{A}_{Int_2}, \ w_1, \ w_2) \in$ {accept offer of an intermediary, continue without an intermediary, do not continue}.

5. Uncertainty realizes and the payments are made.

The strategy of each agent here is $s_i : \Theta \times [0, \delta]^2 \times \mathcal{A}_{Int_1} \times \mathcal{A}_{Int_2}$ $\rightarrow$ {Intermediary 1, Intermediary 2, not connected, direct}, and strategies of intermediaries are $s_{Int} : \{w; \ [0, \delta]^2 \times \mathcal{M} \rightarrow$ {send; do not send}.

Lemma 2.12. *Let $s' = \left(s'_1, \, s'_2, \, s'_{Int_1}, \, s'_{Int_2}\right)$ and $s'' = \left(s''_1, \, s''_2, \, s''_{Int_1}, \, s''_{Int_2}\right)$ be two strategy profiles where $w_1 > w_2$, and let $M = \{m_1, \, ...m_k\}$ be the matches for which the intermediaries both make an offer[5], where $(s'_1, \, s'_2)$ applied to $[w_1, \, w_2,$ offer by Int_1, offer by $Int_2]$ at any $m \in M$ has the acceptance of w_1, and $(s''_1, \, s''_2)$ applied to $[w_1, \, w_2,$ offer by Int_1, offer by $Int_2]$ at any $m \in M$ has the acceptance of w_2. Then, both agents get higher expected payoff in s'' than in s'.*

Proof. Indeed, for any $i \in \{1, \, 2\} : E\pi_i(s''_1, s''_2, \, s''_{Int_1}, s''_{Int_2}) = \sum_{\{\theta_1, \, \theta_2\} = m \subseteq M} p\left(\theta_1\right) p\left(\theta_2\right) p_i \left(\delta - w_2\right) +$

$\left(1 - \sum_{\{\theta_1, \, \theta_2\} = m \subseteq M} p\left(\theta_1\right) p\left(\theta_2\right)\right) \overline{E\pi}_i > \sum_{\{\theta_1, \, \theta_2\} = m \subseteq M} p\left(\theta_1\right) p\left(\theta_2\right) p_i \left(\delta - w_1\right) +$
$\left(1 - \sum_{\{\theta_1, \, \theta_2\} = m \subseteq M} p\left(\theta_1\right) p\left(\theta_2\right)\right) \overline{E\pi}_i = E\pi_i(s'_1, \, s'_2, \, s'_{Int_1}, \, s'_{Int_2})$. The first term $\sum_{\{\theta_1, \, \theta_2\} = m \subseteq M} p\left(\theta_1\right) p\left(\theta_2\right) p_i \left(\delta - w_2\right)$ is expected payoff from choosing the service of intermediary 2 when any of the matches in M was realized. The second term is expected payoff when any other match $\tilde{m} \in \mathcal{M} \backslash M$ was realized. As $w_2 < w_1$, $E\pi_i(s''_1, s''_2, \, s''_{Int_1}, s''_{Int_2}) >$ $E\pi_i(s'_1, s'_2, \, s'_{Int_1}, s'_{Int_2})$. □

Theorem 2.13. *There does not exist a PBE stable to bilateral deviations where both intermediaries set the same wage $w_1 = w_2 = w$.*

Proof. Suppose that there is a PBE stable to bilateral deviations where $w_1 = w_2 = w$. In this case I assume that agents just go with equal probability for the offer of each of

[5] where $M \subseteq \mathcal{M}$, and $\mathcal{M} = \{HH, \, HL, \, LH, \, LL\}$ is the set of all possible matches

the intermediaries. Under w intermediaries provide service to some subset of matches $M \subseteq \mathcal{M}$. First, I show that in equilibrium of this form both intermediaries should get zero expected payoff; second, I show that there would always be a profitable deviation for some intermediary.

Let's assume that intermediaries get positive expected payoff. This means that $w > \underline{w}$, where $\underline{w}$ is the lowest wage when it is still profitable for intermediary to provide service to M. M can be either $\{HH,\, HL,\, LH,\, LL\}$, $\{HH\}$, $\{HH,\, HL,\, LH\}$, or $\{LL\}$. Here a strategy of an intermediary $s_{Int} : \{w;\ w \times \mathcal{M} \rightarrow \{\text{send; do not send}\}\}$. First, let's consider cases when $M \neq \{LL\}$. In that case any intermediary (without loss of generality intermediary 1) has an incentive to deviate set $w' = w - \epsilon$. Due to the above Lemma **??** agents would accept the offer of an intermediary 1, and his deviation from $s^*_{Int_1} = (w, \text{send if } M)$ to $s_{Int_1} = (w', \text{send if } M)$ would be profitable. Expected payoff intermediary gets after the deviation is $E\pi_{Int_1}\left(s_{Int_1},\, s^*_{Int_2},\, s^*_{Ag_1},\, s^*_{Ag_2}\right) = \sum_{\{\theta_i,\, \theta_j\} \in M} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i,\, \theta_j,\, w')$, where $p(\theta_i) = \left\{ \begin{array}{c} p, \text{ if } \theta_i = H \\ 1 - p, \text{ if } \theta_i = L \end{array} \right\}$, is just probability of the type θ_i; and $\pi(\theta_i,\, \theta_j,\, w')$ is the payoff intermediary gets when he provides service to the match $\{\theta_i,\, \theta_j\}$ under w'. When ϵ is sufficiently small (so, that w' is very close to w) $E\pi_{Int_1}\left(s_{Int_1},\, s^*_{Int_2},\, s^*_{Ag_1},\, s^*_{Ag_2}\right) > E\pi_{Int_1}\left(s^*_{Int_1},\, s^*_{Int_2},\, s^*_{Ag_1},\, s^*_{Ag_2}\right) = \sum_{\{\theta_i,\, \theta_j\} \in M} \frac{1}{2} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i,\, \theta_j,\, w)$.

So, when $M \neq \{LL\}$ there can not be an equilibrium where intermediaries get positive expected payoff. Suppose that intermediaries set $w = \underline{w}$, such that they get zero expected payoff. Then, any intermediary has an incentive to deviate to some wage w' where he provides service to some M', such that $w' > \underline{w}'$, where $\underline{w}'$ is zero-profit wage for a match M'. Without loss of generality, I assume that intermediary 1 deviates from $s^*_{Int_1} = (w, \text{send if } M)$ to $s_{Int_1} = (w', \text{send if } M')$. After the deviation $E\pi_{Int_1}\left(s_{Int_1},\, s^*_{Int_2},\, s^*_{Ag_1},\, s^*_{Ag_2}\right) = \sum_{\{\theta_i,\, \theta_j\} \in M' \cap \{\theta_i,\, \theta_j\} \notin M} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i,\, \theta_j,\, w') > 0$ if $M' \cap M^c \neq \emptyset$ and $w' > w$ (where M^c is an complement set of M), or $E\pi_{Int_1}\left(s_{Int_1},\, s^*_{Int_2},\, s^*_{Ag_1},\, s^*_{Ag_2}\right) = \sum_{\{\theta_i,\, \theta_j\} \in M'} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i,\, \theta_j,\, w') > 0$. Therefore, the deviation would be profitable.

To complete the analysis I consider the case of $M = \{LL\}$. Equilibrium where intermediary connects LL exists only when no bilateral-deviation conditions are satisfied $w \leq \delta$, and $w \geq \frac{2c + (p_h + p_l - 2p_h p_l)\delta}{p_h + p_l}$. So, if intermediaries in equilibrium target $M = \{LL\}$ they set the wage $w > \underline{w}$ and always get positive profit. Just in similar fashion as I have shown that there can not be equilibrium where intermediaries get positive profit, I can show that there can not be equilibrium where $w > \frac{2c + (p_h + p_l - 2p_h p_l)\delta}{p_h + p_l}$ which is the minimum wage where equilibrium of intermediary working with LL match exists. Therefore, if there is an equilibrium of this form it should be that $w = \frac{2c+(p_h+p_l-2p_hp_l)\delta}{p_h+p_l}$. However, expected payoff from targeting LL match under wage $w = \frac{2c+(p_h+p_l-2p_hp_l)\delta}{p_h+p_l}$ is lower than targeting all matches and setting $w = \delta - \frac{p\left(p_h^2\delta-c\right)}{p_h+(1-p)p_l}$. Therefore, some intermediary may want to deviate from $s^*_{Int_1} = \left(\frac{2c+(p_h+p_l-2p_hp_l)\delta}{p_h+p_l}, \text{send if } LL\right)$ to $s_{Int_1} = \left(\delta - \frac{p\left(p_h^2\delta-c\right)}{p_h+(1-p)p_l}, \text{send all the time}\right)$, and this deviation is profitable. Therefore, there can not be a symmetric equilibrium where intermediaries work with LL. □

The intuition behind this theorem is very straightforward. When two intermediaries set some wage where they get positive profit they would like to undercut one another and get the whole market. While in case they set wage that gives them zero expected profit every each of them would want to set the wage targeting some different types of agents where he gets positive profit.

Lemma 2.14. *Suppose that intermediaries set $w_1 \neq w_2$ and target M_1 and M_2. Then, in PBE stable to bilateral deviations it can not happen that under both w_1 and w_2 it is profitable to work with both M_1 and M_2.*

Proof. Suppose that in PBE stable to bilateral deviations under both w_1 and w_2 it is profitable to work with both M_1 and M_2. Then intermediary 2 can deviate from $s^*_{Int_2} = (w_2, \text{send if } M_2)$ to $s_{Int_2} = (w_2, \text{send if } M_1 \cup M_2)$. His expected payoff after deviation is $E\pi\left(s^*_{Int_1}, s_{Int_2}, s^*_{Ag_1}, s^*_{Ag_2}\right) = \sum_{\{\theta_i,\, \theta_j\}\in M_1\cup M_2} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i, \theta_j, w_2) > \sum_{\{\theta_i,\, \theta_j\}\in M_2} p(\theta_i)\, p(\theta_j)\, \pi(\theta_i, \theta_j, w_2) = E\pi\left(s^*_{Int_1}, s^*_{Int_2}, s^*_{Ag_1}, s^*_{Ag_2}\right)$, so the deviation is profitable. □

This Lemma 2.14 rules out situations when, for example, both $w_1,\, w_2 > (1-p_l)\,\delta$, and Intermediary 1 targets $M_1 = \{HH, HL\}$, Intermediary 2 targets $M_2 = \{HL, LL\}$. Under this wages working with any match is profitable, therefore intermediary offering lowest wage (Intermediary 2) can switch to targeting a larger subset of matches $\{HH, HL, LL\}$.

Theorem 2.15. *There does not exist a PBE stable to bilateral deviations where $w_1 \neq w_2$.*

Proof. First, I show that intermediaries target by their wages different matches - that is each intermediary is an ex post monopolist in his wage region; then I prove that each intermediary should get the maximum possible payoff in this wage region; finally, I should that (for not degenerate parameter values) there is always a profitable deviation.

Suppose that there exists a PBE stable to bilateral deviations where $w_1 \neq w_2$. Intermediary 1 provides service in equilibrium for some subset of matches $M_1 \subseteq \mathcal{M}$, intermediary 2 provides service to $M_2 \subseteq \mathcal{M}$. Without loss of generality I assume $w_1 > w_2$.

First, I want to show that intermediaries are ex post monopolists in their wage regions - this means that when intermediary i targets some match $m^* \in M_i$, then he is the only one who targets m^*. Suppose this was not the case, and the match m^* is targeted by both intermediaries. Due to the Lemma 2.14 I can concentrate on the case where $M = m^*$ is the only match that both intermediaries are targeting. Then, intermediary 1 has a profitable deviation from $s^*_{Int_1} = (w_1, \text{ send if } M)$ to $\left\{ \begin{array}{c} s_{Int1}{=}(w_2{-}\epsilon, \text{ send if } M), \text{ if } w_2{>}\underline{w}_2 \\ s_{Int1}{=}(w_2, \text{ send if } M), \text{ if } w_2{=}\underline{w}_2 \end{array} \right\}$. This deviation is profitable.

Next I want to show that there can not be a PBE stable to bilateral deviations where $\forall\ i,\ \ w_i^* \in [\underline{w}_i,\ w_i)$. Indeed, suppose that there is such an equilibrium, and in this equilibrium intermediary i provides service to some $M_i \subset \mathcal{M}$. As he is ex post monopolist to matches in M_i intermediary i could deviate from the strategy $s^*_{Int_i}\,(w_i^*, \text{ send if } M_i)$ to the strategy $s_{Int_i}\,(w_i, \text{ send if } Mi)$ where $w_i > w_i^*$ and the deviation is profitable.

Therefore, the only equilibrium candidate is the one when any intermediary i sets $w_i = \bar{w}_i$, where $\bar{w}_i$ is the largest possible wage under which one can work with the matches M_i.Then intermediary i that gets lower expected payoff of two has an incentive to deviate from the strategy $s^*_{Int_i} = (\bar{w}_i,\ M_i)$ to $s_{Int_i} = (\bar{w}_j - \epsilon,\ M_j)$ and this deviation is profitable. □

The only possibility for an equilibrium to exist is when both intermediaries get the same expected payoff by setting wages $\bar{w}_1$, $\bar{w}_2$, and targeting M_1, M_2; and at the same time they can not get higher expected payoff by targeting some other match $m_3 \in \mathcal{M}\setminus\{M_1 \cup M_2\}$ and setting $\bar{w}_3$.

So, it turns out that in the game with two competing intermediaries there is no PBE in pure strategies stable to bilateral deviations. This happens due to the fact that there are too many matches and too few intermediaries that aim at these matches. Even if intermediaries set different wages and target different matches there would always be a profitable deviation for some of them. This potential deviation has three components - if intermediary is unsatisfied with expected profit he gets he might either aim at another match; or he can rise the wage still aiming for the same combination of types (and he can do it if he's a monopolist in this region - so, if another intermediary aims for different matches), or he might undercut the second intermediary. All these result in nonexistence of an equilibrium in this kind of a game. There is a need for a tighter competition between intermediaries to equilibrate the system. Even though here intermediaries are not monopolists ex ante they turn out to be monopolists ex post on the market of services for the specific match they are working with. However, being a monopolist ex post in their particular wage region ruins the equilibrium of a price setting game ex ante.

This nonexistence result is similar to nonexistence in screening models. Just like there intermediaries here aim their offers at different types, as they know that in subsequent equilibrium this particular types would accept it. One of the ways to cope with nonexistence problem in screening is the concept of reactive equilibrium. It allows the obedient agent to react to the deviation of another agent. Only deviations that can not bring losses given the reaction are allowed.

In my model an intermediary can never suffer losses in expected terms, the worst situation for him is that he gets zero. Therefore, the exact type of argument that works in reaction equilibrium is not applicable here.

In case we modify reaction equilibrium concept and look only at deviations that bring positive payoffs given the reaction of an obedient agent we still have problems as this a too harsh way to tackle nonexistence problem. Instead of nonexistence we get multiple equilibrium where every combination of wages will be an equilibrium. Indeed, when intermediary deviates to w' obedient agent can always react with $w' - \epsilon$, and leave the deviator with zero expected surplus. If I allow for these type of reactions any combination of wages can be part of an equilibrium.

So, we see that though monopoly intermediary does restore efficiency and improves the welfare, competition between intermediaries ruins the equilibrium and it's not even possible to predict outcome of the game.

On the other hand, it is interesting to get nonexistence similar to Rothschild-Stiglitz in a different type of game. As reactive equilibrium does not work in this setting there is a need for a different idea about tackling the problem. It might also be useful to run an experiment and see if there is any pattern in participants behavior.

2.5.3 More than two intermediaries

We've seen in previous section that ex post competition between intermediaries is insufficient to provide the existence of equilibrium. A natural question here is what is the minimal number of intermediaries needed for equilibrium to exist and what is the properties of this equilibrium.

Theorem 2.16. *Suppose we have N intermediaries. If there exist a PBE equilibrium stable to bilateral deviations then in this equilibrium every match $m \in \mathcal{M}$ is approached at least by two intermediaries.*

Proof. First, I want to show that no match $m \in \mathcal{M}$ can be targeted only by one intermediary. Then, I show that there can not be a match that is not targeted at least by some intermediary.

It is easy to show that if some match m is approached by only one intermediary i there can not be an equilibrium where intermediary sets $w_i < \bar{w}_i$. Then, if there exist some other match m' that is approached by more than one intermediary it is easy to show that those intermediaries will set the lowest possible wage where they get zero expected payoff from working with m'. Therefore, any of intermediaries working with m' can deviate to $s_{Int_j}\,(\bar{w}_i - \epsilon,\, m)$ and get positive expected payoff.

If there is no such match m' targeted by two intermediaries, then any other intermediary j is also an ex post monopolist for some match m_j, therefore there again can not be an equilibrium where they set $w_j < \bar{w}_j$. Then if $\exists j \in \{1,...,i-1,\, i+1,\, N\} : E\pi_j(\bar{w}_j) < E\pi_i(\bar{w}_i)$, j would like to deviate from $s^*_{Int_j}\,(\bar{w}_j,\, M_j)$ to $s_{Int_j}(\bar{w}_i - \epsilon,\, M_i)$ and enjoy higher payoff. In case there's no j that gets lower payoff than intermediary i but $\exists j \in \{1,...,i-1,\, i+1,\, N\} : \pi_j(\bar{w}_j) > \pi_i(\bar{w}_i)$, i would like to deviate to $s_{Int_i}(\bar{w}_j - \epsilon,\, M_j)$ and enjoy higher payoff. The only way there can be such an equilibrium is when for $\forall j \in \{1,...,i-1,\, i+1,\, N\} : \pi_j(\bar{w}_j) = \pi_i(\bar{w}_i)$ which requires very specific parameter and is a very fragile condition for existence of equilibrium. Moreover, for the specific game of mine it is just generally not true.

Now I show that there can not be a match that is not addressed by some intermediary. Suppose indeed $\exists\ \tilde{m} \in \mathcal{M}$ that does not receive any offer of an intermediary. According to what is said above all intermediaries are no ex post monopolists for the match they are targeting, therefore they receive zero expected payoff. Then, any intermediary i can deviate to $s_{Int_i}\,(w'_i,\, \tilde{m})$ where w'_i is such wage where he targets $\tilde{m}$ and gets positive expected payoff. Therefore, we have a contradiction. $\square$

These theorem tells us that in equilibrium every match should be coordinated by some intermediary, and that no match can be coordinated by only one intermediary. So, there should be an least twice as many intermediaries as there are different matches. For the game I study at least 6 intermediaries are needed.

In that case there is both competition ex ante and ex post. In equilibrium all agents interact only through intermediaries - there is no direct collaboration. Intermediaries specialize on some particular type of matches that they work with - but they always have competitor/s that specialize on the same matches.

It is worth talking about the information property of the equilibrium. The information set of the agent is now richer because it also includes the wage. The wage itself does not bring any additional information as intermediary decides on it ex ante. However, when $N \geq 6$ intermediaries perfectly specialize, then the wage gives an agent full information about the type of the partner. Still, this does not change agents behavior as no bilateral deviation conditions (a bit different but very similar[6]) are always satisfied.

This equilibrium is efficient as all matches collaborate through an intermediary which is a cheaper technology. All agents have perfect knowledge about the type of their partner but they still prefer to use the offer of an intermediary because competition between those has pushed the wage down. At the same time it is precisely the credit of competition between intermediaries and subsequent specialization on particular matches that lead to full information disclosure in equilibrium.

[6]They are the same for LL and HH matches. Condition for HL match is not $p_h(\delta - w) > p_h p_l - c$, and it's always satisfied.

2.6 Conclusion

In this paper I study how informed intermediary restores efficiency when agents face imperfect information. The paradox is that agents are not committed to pay the intermediary once he brings information to the table. This changes the incentive of an intermediary to participate. I find that despite the lack of commitment monopoly intermediary still restores the efficiency. The outcome of the game depends on the prevalence of the high type because it determines the best equilibrium for the intermediary. In the equilibrium intermediary either provides full information (by the fact of an offer in one equilibrium or by the lack of it in another) to low type agents that are the most vulnerable to incompleteness of information , or does not disclose any information at all. In all equilibria presence of intermediary is welfare improving.

The situation changes when we add competition between intermediaries. If competition is insufficient the equilibrium breaks and fails to exist (in a manner similar to Rothschild-Stiglitz). Unfortunately, nonexistence here can not be tackled with the help of reactive equilibrium concept and the question of appropriate solution concept is still open.

When the number of intermediaries rises sufficiently to guarantee competition ex post, the only equilibrium is the one where intermediaries specialize on different matches. Agents in every match receive offers from at least two intermediaries. The outcome of the game is again efficient. However, unlike the monopoly case, the superior knowledge of an intermediary is not unique (as there are many of them), and therefore, intermediaries can not benefit from it.

2.7 Appendix

Proof. Theorem 2.3 . Suppose that H believes that another H chooses direct collaboration with μ_h and an L type agent chooses direct collaboration with μ_l. An L agent believes that H chooses direct collaboration with ν_h and L chooses direct collaboration with ν_l. Given this when H decides to participate in collaboration he expects to get $p\mu_h\left(p_h^2\delta - c\right) + (1-p)\mu_l\left(p_h p_l \delta - c\right)$. When a low type agent chooses to go for a direct collaboration he expects to get $p\nu_h\left(p_h p_l \delta - c\right) + (1-p)\,\nu_l\left(p_l^2\delta - c\right)$. First let us consider four possibilities of pure strategies equilibrium: 1) both types of agents always choose to participate in collaboration; 2) both types of agents do not want to participate in collaboration; 3) high type agents always want to participate in collaboration, while low types prefer not to do it; 4) high type agent prefer not to participate in collaboration, while low type agents do want to participate.

1) In this case belief consistency requires $\mu_h = \nu_h = \mu_l = \nu_l = 1$. Then low type expects to get from participating in collaboration $p_l\left(pp_h + (1-p)\,p_l\right)\delta - c < 0$. Therefore this can not be an equilibrium.

2) In this case belief consistency requires $\mu_h = \nu_h = \mu_l = \nu_l = 0$. Then both types of agents expect to get 0 from participating in collaboration, and this can be a part of equilibrium. However, this can not be part of equilibrium that is stable to bilateral deviations a agents might bilaterally switch to a strategy profile where they participate in collaboration in case the high type is realized to them. Under such new strategy profile each of them gets expected payoff of $p^2\left(p_h^2\delta - c\right) > 0$, so the bilateral deviation is profitable.

3) In this case belief consistency requires $\mu_h = \nu_h = 1$ and $\mu_l = \nu_l = 0$. Then payoff of a low type agent from participating is $p\left(p_h p_l \delta - c\right) > 0$, which means that he would like to take part in collaboration, and therefore initial construction can not be a part of equilibrium.

4) In this case belief consistency requires $\mu_h = \nu_h = 0$ and $\mu_l = \nu_l = 1$. Given this both types of agents would prefer to deviate as for high type agent payoff from participating in collaboration is $p\left(p_h p_l \delta - c\right) > 0$ and for low type agents it is $(1-p)\left(p_l^2\delta - c\right) < 0$.

Therefore, there can not be an PBE stable to bilateral deviations in pure strategies. Now let us look at possible equilibria in mixed strategies.

First of all let's see if high type agent may participate in collaboration with some probability $0 < \alpha < 1$. Belief consistency requires that $\mu_h = \nu_h = \alpha$. Then, payoff from going for direct connection for a high type agent is $p\alpha\left(p_h p_l \delta - c\right) + (1-p)\mu_l\left(p_h p_l \delta - c\right) > 0$ for $\forall \mu_l$. Therefore, in any mixed equilibrium high type agents always go for collaboration.

It means that $\nu_h = 1$. Then, expected payoff of a low type agent from going to direct connection is $p\left(p_h p_l \delta - c\right) + (1-p)\,\nu_l\left(p_l^2\delta - c\right)$. The only ν_l that can support this mixed strategy equilibrium is the one where low type agents are indifferent between going for direct connection or not being connected at all: $\nu_l = \frac{p}{1-p}\frac{p_h p_l \delta - c}{c - p_l^2\delta}$. So, there is a mixed strategy equilibrium where high type agents always go for direct connection, while low type agents go for direct connection with $\beta = \frac{p}{1-p}\frac{p_h p_l \delta - c}{c - p_l^2\delta}$. □

Proof. Theorem 2.4 . When $w < (1-p_h)\,\delta$ intermediary does not want to propose his offer for any combination of agents. Even if the best possible match of HH has realized intermediary expects to get $2p^2 p_h\left(w - (1-p_h)\,\delta\right) < 0$ for $\forall w < (1-p_h)\,\delta$. Then the situation is just the same as in case of no intermediary, therefore, equilibrium is also the same. □

Proof. Theorem 2.5 . When $c \geq p_h\left(w - (1-p_h)\,\delta\right)$ two agents don't want to deviate bilaterally to another strategy profile where in case they are of high type they prefer to participate in collaboration directly[7]. Under such wage intermediary would want to provide his service only to a match of two high types, as even in case one of the agent is of low type, expected payoff of an intermediary is $(p_h + p_l)w - [p_h(1-p_l) + p_l\,(1-p_h)]\delta < 0$ for $\forall w \in \left[(1-p_h)\,\delta;\ \left(1 - \frac{2p_h p_l}{p_h + p_l}\right)\delta\right]$.

First, let's see that reported strategy profile is indeed a PBE stable to bilateral deviations. As was already discussed two agents don't want to deviate and change their behavior in the $H+$ information set. As intermediary sends an offer only to HH bilaterally changing to a strategy profile with different $s_i\,(H+)$, $s_j\,(L+)$, or $s_i\,(L+)$, $s_j\,(L+)$ does not bring any additional payoff as $L+$ information set is always out of equilibrium. Situation in information sets that correspond to no offer of an intermediary is similar to that of no-intermediary (with only difference that possible matches here are HL and LL, as HH match always receives the offer), therefore, the same behavior is part of equilibrium here too.

Now, let's see that there is no other equilibrium. The only other possibility is that there might be an equilibrium where something different happens to HL and/or LL match. It can't be that HL and/or LL would collaborate indirectly as intermediary under such wage does not want to deal with this matches. And it can not happen either that either HL or LL would turn out to be collaborating directly as this might not be a part of equilibrium due to the arguments of Theorem 2.3 . □

[7] Deviation is unprofitable when $p\left(p\left(p_h^2\delta - c\right) + (1-p)\,\beta\left(p_h p_l \delta - c\right)\right) + (1-p)\left(p\beta\left(p_h p_l \delta - c\right) + (1-p)\,\beta^2\left(p_l^2\delta - c\right)\right) < p^2 p_h\left(\delta - w\right) + p\,(1-p)\,\beta\left(p_h p_l \delta - c\right) + (1-p)\,p\beta\left(p_h p_l \delta - c\right) + (1-p)^2\beta^2\left(p_l^2\delta - c\right)$

Proof. Theorem 2.6 . Condition $c > \frac{\delta p_h p p_h + 2(1-p)p_l) - (\delta - w)(p_h + (1-p)p_l)}{p + 2(1-p)}$ specifies parameter values when there is no bilateral deviation[8] by agents. Condition on wage of an intermediary specifies a situation when intermediary is willing to work with matches of HH and HL, but not with a match of LL.

Reported strategy profile is indeed an equilibrium. As discussed above agents do not wont to change their behavior in informations sets that corresponds to the offer of an intermediary. Neither do they want to behave differently in case they do not receive an offer because $H-$ information set is out of equilibrium and in $L-$ information set they behave optimally as they are able to avoid collaboration in an inefficient LL match. Intermediary also behaves optimally as he sends an offer to the largest subset of matches when it is still valuable for him to do so.

There is no other PBE stable to bilateral deviations except the two described as agents behave optimally given that they receive the offer (as they do not want to bilaterally deviate), and they will always receive the offer because this is optimal for an intermediary. □

Proof. Theorem 2.7 . When $w \geq (1 - p_l)\,\delta$ the wage of an intermediary is so high that he is willing to intermediate any match, even the bad match of LL agents. The condition on the cost of direct collaboration eliminate the bilateral deviations by agents[9]. So, neither agents nor intermediary has any incentive to deviate. □

Proof. Theorem 2.8 . As it was already said above in this wage region intermediary is willing to provide his service to any combination of agents. The second condition is a condition for no bilateral deviation by an intermediary and an agent. When the above inequality holds agents do not want to change his behavior in $H+$ information set provided that intermediary would also change his behavior and send an offer both to HL and LL matches. The third condition on the cost of direct connection eliminates bilateral deviation by two agents. In the region of my main interest where $c \in [p_l\,(pp_h + (1-p)\,p_l)\,\delta;\, p_l p_h \delta]$ this last condition is always satisfied as the cost of direct connection is already sufficiently high such that agents prefer to collaborate indirectly in case the worst possible type has realized for them and their partners.

This is indeed an equilibrium as given the strategy of an intermediary all agents behave optimally - both those that that have received an offer and have not. Intermediary also

[8]Bilateral deviation is unprofitable when $pp_h\,(\delta - w) + (1-p)\,pp_l\,(\delta - w) > p\left[p\left(p_h^2\delta - c\right) + (1-p)\,(p_h p_l \delta - c)\right] + (1-p)\,p\,(p_h p_l \delta - c)$

[9]It seems that that agents can switch either to a strategy profile where they change behaviors only in $H+$ information set (and get $p^2\left(p_h^2\delta - c\right) + (1-p)^2 p_l\,(\delta - w)$), or in both $H+$ and $L+$ information sets (and get $p^2\left(p_h^2\delta - c\right) + p\,(1-p)\,\beta\,(p_h p_l \delta - c)$). However, it is easy to show that the first expression is always larger in the region of interest where $c \in [p_l \bar{p} \delta,\, p_l p_h \delta]$

behaves optimally as it is not profitable for him to deviate and send an offer to different matches even if one of agents deviates with him. □

Proof. Claim 2.9 . In the first equilibrium where all agents collaborate indirectly the total welfare is the sum of expected payoffs of two agents and an intermediary. $W_1 = 2\bar{p}(\delta - w) + 2\bar{p}(w - \delta(1-\bar{p})) = 2\bar{p}^2\delta$[10]. As it is stated in Theorem2.3 a no intermediary equilibrium is the one where high types always want to collaborate and low types want to do it with positive probability. The total welfare here is the sum of expected payoff of both agents. $W2 = \frac{2p2(ph-pl)2c\delta}{(c-pl2\delta)}$. As $\frac{\partial W2}{\partial c} < 0$ and $c \in [pl\bar{p}\delta; phpl\delta]$ $W2 \leq 2p(ph-pl)\delta\bar{p} < W1$. □

Proof. Theorem2.11.It is clear that indeed setting $\tilde{w}^*$ and working with corresponding matches is an equilibrium. Here intermediary gets the highest expected payoff. If he deviates to another wage, even in case the intermediary equilibrium is played after the deviation, his expected payoff is lower.

Now, let me show that there is no other equilibrium[11]. If there is some different equilibrium if intermediary deviates to $\tilde{w}^*$ he knows that there can be only two continuation equilibria one of them being No Intermediary Equilibrium, which is Pareto inferior. Therefore, if he works with different subset of matches, he can deviate and set $\tilde{w}^*$ and get higher expected payoff, as this wage corresponds to the equilibrium where he gets the highest expected payoff. If he already works with preferred subset of types but sets $w < \tilde{w}^*$ he can increase his wage and get higher payoff, as his expected payoff is an increasing function of wage. □

[10] Where $\bar{p} = pp_h + (1-p)p_l$

[11] It is important that even though both w^*, $w^*_{LL} \in [1-p_l, \delta]$ regions for existence of these two different equilibrium do not intersect ($LL \notin CE(w^*, c)$ and $All \notin CE(w^*_{LL}, c)$).

Chapter 3

Optimal Prize Allocation in Contests with Sabotage

I.Kirysheva[1]

[1] irina.kirysheva@nu.edu.kz

3.1 Introduction

An important issue in economic theory is how to align incentives of workers in organizations with incentives of the principals. Due to simplicity of implementation tournaments (a payment scheme that is based on the rank of the agents rather on their performance) are extensively used[2]. Performance evaluation schemes (such as tournaments) are widely used by the companies. Some corporations (for example Microsoft) use "sticky tanking" to evaluate their employers. In this system their performance is allocated to one of the groups, e.g. "excellent", "good", "average", "poor", "very poor" etc. Sometimes this scheme provides the base for bonus payouts (and actually links them to individual performance). Sometimes the reward people are competing for is the promotion, sometimes it can mix of the benefits, and it might be even recognition[3].

Tournaments have been widely studied in economic literature. While majority of research concentrated on tournaments with a single prize, there is also important work done on multiple-prize tournament. For example,Moldovanu et. al. [21, 22] have papers on optimal prize allocation in the contests. They find that with linear cost function it is optimal for designer to allocate the whole prize sum to the winner.

One concern that is important for the tournaments is that they can be subject to sabotage. There are several papers that look at tournaments with sabotage. For example, Chen in [4] looks at one-period model where players can make productive efforts, or can make destructive efforts towards their colleagues. He finds that able agents are more likely to be subject to sabotage attacks. Also, due to sabotage activities the most talented agents might not have the highest chance to be promoted. Similar problem is analyzed by Munster in [24]. He also finds that talented agents are sabotaged more heavily, and that sabotage equalizes the probability of promotion for agents of different characteristics. Gurtler in [11] looks at the dynamic sabotage game with psychic cost of being sabotaged. Due to these costs it might be optimal for talented agents in the first period to actually help others and sabotage themselves.

Unlike the previous works mentioned Amegashie et. al. [1] consider dynamic contests with sabotage where players can sabotage not the current rivals but those they might meet in the future. Contestants are divided into two semi-finals, and they can help player from another semifinals. They find that there is an equilibrium where only the most able player engages in sabotage, which is a surprising result, as usually it is the most talented agent who suffers most from being sabotaged by others.

[2]See for example Lazear and Rosen [16] for discussion.
[3]http://vimeo.com/97045946

Another strand of literature on contests with sabotage is experimental one. For example, Carpenter et. al. [3] conducts a real-effort experiment, where participants could sabotage their contestants. They found that players actually provided less effort because they anticipated to be victims of sabotage. In other works by Harbring et.al. [13],[12] authors divide players into three types according to the cost of the effort: favorites, normals, underdogs. They find that sabotage behavior varies with the composition of types of players - for example, underdogs sabotaged favorites less in the contest with more favorites. Another finding is that sabotage decreases if saboteurs identity is revealed.

Given this concern about the tournaments it is worth looking how presence of sabotage changes the result of optimal prize allocation for the tournament designer. Even intuitively it seems that giving all the prize to the winner will create high incentives to sabotage. So, it might be optimal for the principal to provide positive prizes also to those who have lost in the tournament. The result also should depend on the number of participants in a contest. With more than two players sabotage becomes a public good, which can undermine incentives to sabotage, and make it a lesser concern for the designer.

Many of the real-world contests have multiple prizes, where not exclusively the winner gets positive prize. For example, majority of sports contests award gold, silver and bronze; in tender contests the second ranked firm can still be used as a back-up supplier; in labor market several workers can be promoted.

I find that in continuous case with multiple prizes the optimal solution crucially depends on the cost of sabotage. If sabotage is expensive designer wants to give all the prize to the winner just like in classical no-sabotage case. However, when sabotage is cheap he does not want to make a contest at all. Remarkably, we observe companies giving away the "forced-ranking" reward scheme. For example, Microsoft got rid of "stack ranking" in 2013[4].

[4] www.businessinsider.com/microsoft-just-killed-its-controversial-stack-ranking-employee-review-system-2013-11

3.2 Example

I start with the simple example that provides the main intuition. Principal decides on two non negative prizes V_1 and V_2 that he distributes between two contestants. The winner of the contest gets V_1, another player gets V_2, in case of the same result each one gets $\frac{V_1+V_2}{2}$. Agents can choose effort level $e \in \{0, 1\}$. However, unlike standard model, they can also sabotage the outcome of their opponent by making a distructive effort $d \in \{0, 1\}$. The principal does not observe neither the productive effort, nor the sabotage agents have made. Instead for each agent i he observes the effort net of sabotage $\tilde{e}_i = e_i - d_j$, and he distributes prizes according to $\tilde{e}_i$ - the agent with highest $\tilde{e}_i$ gets the first prize V_1, another agent gets V_2.

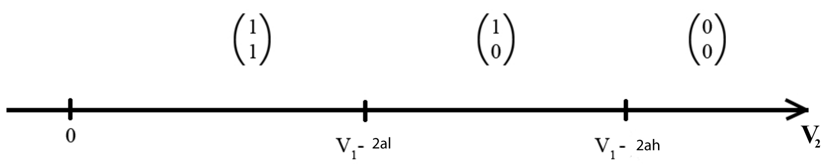
Agents can be of two types - either high or low ($\Theta = \{H, L\}$). The type of the agent determines the costs of productive effort - for high types it is a_h, and for low types it is $a_l > a_h$ (it is easier for high type agents to produce). The sabotage activity is also costly and costs a_s, where $a_h < a_s < a_l$. Therefore, it is cheaper for low type to sabotage, while for high type it is cheaper to make a production effort. The cost of sabotage does not depend on the type.

The principal wants to maximize the sum of expected net efforts ($E(\tilde{e}_1 + \tilde{e}_2)$) given that the sum of prizes equals to the budget available for the principal, $V_1 + V_2 = P$.

3.2.1 No sabotage baseline

In case there is no sabotage possibility the optimal prize allocation result is just as predicted by the classical result on optimal prize structure.

I am looking for PBE $s : \Theta \to e$. Here $\begin{pmatrix} e_h \\ e_l \end{pmatrix}$ means that high type agent chooses effort level e_h, while low type agent chooses e_l.



The picture shows which equilibrium will be played for different relationships between V_1 and V_2.

Claim 3.1. Without sabotage the optimal prize scheme assumes that the principal gives not more that $V_1 - 2a_l$ to the loosing individual.

Proof. The first best outcome for the principal is to guarantee that both types make efforts. In that case the expected payoff for high type agent is $\frac{V_1+V_2}{2} - a_h$, and the expected payoff for low type agent is $\frac{V_1+V_2}{2} - a_l$. The first best outcome can be supported as an equilibrium if low type does not want to deviate and get V_2 instead. Therefore, the following inequality should be satisfied: $\frac{V_1+V_2}{2} - a_l \geq V_2$. This brings us the restriction on V_2 that $V_2 \leq V_1 - 2a_l$. □

We see that for the principal it's optimal to make V_2 really low compared to V_1 ($V_2 \leq V_1 - 2a_l$) which is just in line with the result of [21]. In this case solutions contains the case $V_1 = P$, and $V_2 = 0$. So, in the absence of sabotage it is optimal for the principal to give the whole prize sum to the winner, as this will allow principal to achieve the first-best outcome, where both types make effort.

So, here I get the standard optimal prize allocation results but applied to discrete case. The first best equilibrium can be supported by a set of values of V_1 and V_2 that imply the winner getting high reward, including the case where $V_1 = P$, and $V_2 = 0$. This is intuitive as in the absence of sabotage there is no need to make high second-prize as It will only increase incentive to shirk.

3.2.2 Sabotage case

Now I assume that players can make both productive and distructive efforts. I look for the equilibrium of the form $s : \Theta \to e \times d$. I denote a strategy by $\begin{pmatrix} e_h \\ d_h \\ e_l \\ d_l \end{pmatrix}$ where e_h represents the productive effort of a high type agent, d_h is the destructive effort of high type agent, e_l is the productive effort of a low type agent, and d_l is a destructive effort of low type agent.

Claim 3.2. When sabotage is possible principal can not achieve first-best outcome where both types make productive effort and do not sabotage.

Proof. The first-best result for the principal is when agents are playing the strategies $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. However, this strategy profile can not be an equilibrium because there exists a profitable deviation. When playing $s_i = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, agents receive expected payoff $EU_i = \frac{V_1+V_2}{2} - pa_h - (1-p)a_l$. If an agent deviates to $s_i' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ he gets expected payoff of $EU_i' = \frac{V_1+V_2}{2} - pa_h - (1-p)a_s > EU_i = \frac{V_1+V_2}{2} - pa_h - (1-p)a_l$ as $a_s < a_l$. Therefore, it's not possible achieve a first-best when sabotage is possible.

□

Now, the principal does not want to make V_2 very low with respect to V_1 as this induces high incentives to sabotage from both types of agents. While first-best outcome is not possible in this circumstances, the principal still can guarantee the second-best outcome.

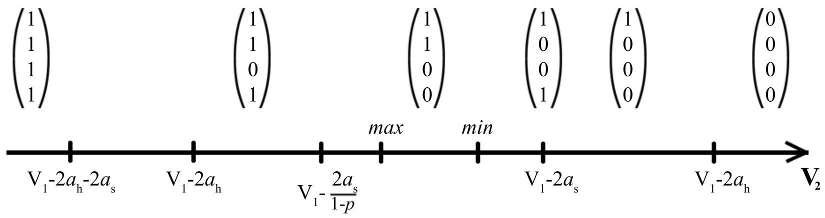
Claim 3.3. The principal can achieve the second-best outcome where high type makes productive effort while low type at least does not sabotages when $V_2 \in [V_1 - 2a_s;\ V_1 - 2a_h]$

Proof. The strategy profile $s_i^* = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ constitutes an equilibrium when there's no profitable deviation. The expected payoff is $EU_i(s_i^*, s_{-i}^*) = p[p\frac{V_1+V_2}{2} + (1-p)V_1 - a_h] + (1-p)[pV_2 + (1-p)\frac{V_1+V_2}{2}]$. This strategy profile is an equilibrium if there are no profitable deviations. Potential strategies, where one can deviate are $s_i' = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $s_i'' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $s_i''' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. $EU_i(s_i', s_{-i}) = p[V_1 - a_h - a_s] + (1-p)[pV_2 + (1-p)\frac{V_1+V_2}{2}]$, this deviation is not profitable when $V_2 \geq V_1 - \frac{2a_s}{p}$. $EU_i(s_i'', s_{-i}) = p[p\frac{V_1+V_2}{2} + (1-p)V_1 - a_h] + (1 -$

$p)[p\frac{V_1+V_2}{2}+(1-p)V_1-a_s]$, this deviation is non profitable when $V_2 \geq V_1 - 2a_s$. Finally, $EU_i(s_i''', s_{-i}) = p[pV_2+(1-p)\frac{V_1+V_2}{2}]+(1-p)[pV_2+(1-p)\frac{V_1+V_2}{2}]$, this is unprofitable when $V_2 \leq V_1 - 2a_h$. Combining all inequalities we get that the principal can reach the second-best when $V_2 \in [V_1 - 2a_s; V_1 - 2a_h]$.

□

The relationship between the V_1 and V_2 defines the equilibria played in a following manner:



So, we see that for the principal it is optimal to set $V_2 \in [V_1 - 2a_s; V_1 - 2a_h]$. This prize distribution results in equilibrium where only high type agent makes an effort. It is impossible for the principal to induce the low type agent to exert a productive effort without any destructive effort.

This example shows that given the presence of sabotage it is optimal for the principal to provide also the positive price for the second comer so that low type agents does not have incentive to sabotage their partners.

3.2.3 N-player case with private information

Now suppose that there are $N \geq 3$ players that compete in a contest. Principal distributes N prizes V_1, V_2... V_N. First I consider the private information case - where each contestant only knows his type and does not know types of other contestants. Principal has no information about the types. With this information structure for each contestants others look exactly the same because their types are not known and only a priori distribution is known.

Now there is a possibility for different sabotage protocols: either by $d = 1$ an agent can sabotage all other players at the same time simultaneously (I call this mass sabotage protocol) or he has to choose one individual and sabotage this particular individual without harming others (this is individual sabotage protocol).

With individual sabotage protocol destructive effort acts as a public good (or in this case more "public bad"), therefore one could expect to see less sabotage in equilibrium as agents would want to free-ride on others. Therefore, a principal may be less worried about the sabotage possibility and may again want to give majority of the prize sum to the winner (to go back to the prize scheme that is optimal in the absence of sabotage).

Now the principal has to distribute N prizes V_1, V_2 ... V_N, where V_1 goes to the agent with the highest observed outcome $\tilde{e}_i$, V_2 to the agent with the second outcome, and so on. Finally, V_N goes to the agent with the poorest performance.

3.2.3.1 Mass sabotage protocol

In case of mass sabotage individual can harm all his colleagues simultaneously. In this case sabotage takes more the form of cheating. By making a sabotage decision of $d = 1$ harms all other players at the same time. Here the result will be similar to two-player case. The principal can not achieve the first-best equilibrium, but he can guarantee the second-best one where high type makes a productive effort and the low type at least does not sabotages.

Claim 3.4. In a contest with $N \geq 3$ players and mass sabotage protocol principal can not achieve first best where both types work and nobody sabotages.

Proof. Suppose there exists an equilibrium where $s_i^* = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$(so that $e_h^* = e_l^* = 1$, and $d_h^* = d_l^* = 0$). The expected payoff in this equilibrium would be

$$EU(s_i^*, s_{-i}^*) = p[\tfrac{\sum_{i=1}^N V_i}{N} - a_h] + (1-p)[\tfrac{\sum_{i=1}^N V_i}{N} - a_l].$$

Then a player can switch to a strategy $s' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, where $e_l = 0$, and $d_l = 1$, and get

$$EU\left(s_i', s_{-i}^*\right) = p[\tfrac{\sum_{i=1}^N V_i}{N} - a_h] + (1-p)[\tfrac{\sum_{i=1}^N V_i}{N} - a_s] > EU_L(s_i^*, s_{-i}^*)$$

as $a_s < a_l$. □

Claim 3.5. The principal can achieve the second-best outcome, where high type makes an effort, and low type does not sabotages. He does so by giving at least some part of his budget to second and third prizes. The restriction on the relationship between V_1 and V_2, $V_3...V_N$ are of the form $\alpha_1 V_1 \leq \alpha_0 + \sum_{i=2}^{N} \alpha_i V_i$, where $\alpha_1 > 0$.

Proof. The strategy profile $s_i^* = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ constitutes an equilibrium when there are no profitable deviations. Expected payoff from playing this strategy profile is

$$EU(s_i^*, s_{-i}^*) = pEU_H(s_i^*, s_{-i}^*) + (1-p)EU_L(s_i^*, s_{-i}^*) = p[\sum_{k=0}^{N-1} \begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k \frac{\sum_{i=1}^{N-k} V_i}{N-k} - a_h] + (1-p)[\sum_{k=0}^{N-1} \begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k \frac{\sum_{i=N-k}^{N} V_i}{k+1}],$$

where the first part is the expected payoff of being a high type while the second part is the expected payoff of being the low type. Again I am checking for three alternative strategies that can be potentially profitable deviations: $s_i' = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $s_i'' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $s_i''' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. The intuition for considering these particular strategies is that high type has two potential deviations - he can either sabotage additionally to working or he can neither work nor sabotage. It is not profitable for a high type to switch to sabotage instead of productive effort as sabotage is more expensive for him. Respectively, for low type there is only one potential deviation - where he sabotages. Low type will not deviate to just working as working is more expensive for him. Neither should we consider possible deviation where he works and sabotages at the same time as these two activities for low type are more expensive than for high type, therefore, we should get respective restrictions from no-deviation condition for high type. Out of these potential three deviations two involve one type sabotaging (high type in s' and low type in s'') and precisely from these two deviations I will get restrictions on V_1.

$$EU(s_i', s_{-i}^*) = pEU_H(s_i', s_{-i}^*) + (1-p)EU_L(s_i', s_{-i}^*) =$$
$$p[V_1 - a_h - a_s] + (1-p)[\sum_{k=0}^{N-1} \begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k \frac{\sum_{N-k}^{N} V_i}{k+1}].$$

This deviation is non-profitable when

$$V_1\Big(1-\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k}\Big) \le a_s + \sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k\sum_{i=2}^{N-k}V_i}{N-k}.$$

This is the first inequality that gives us the restriction on V_1.

Coefficient in front of $V_1$is greater than zero. Indeed, $(1-\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k}) > (1-\sum_{k=0}^{N-1}\binom{N-1}{k}p^{N-1-k}(1-p)^k) = (1-(p+(1-p))^N = 0$.

$EU(s_i'', s_{-i}^*) = pEU_H(s_i'', s_{-i}^*) + (1-p)EU_L(s_i'', s_{-i}^*) = p[\sum_{k=0}^{N-1}\binom{N-1}{k}p^{N-1-k}(1-p)^k\frac{\sum_{i=1}^{N-k}V_i}{N-k} - a_h] + (1-p)[\sum_{k=0}^{N-1}\binom{N-1}{k}p^{N-1-k}(1-p)^k\frac{\sum_{i=1}^{N-k}V_i}{N-k} - a_s]$. This is not profitable when $\sum_{k=0}^{N-1}\binom{N-1}{k}p^{N-1-k}(1-p)^k\frac{\sum_{i=1}^{N-k}V_i}{N-k} - a_s \le \sum_{k=0}^{N-1}\binom{N-1}{k}p^{N-1-k}(1-p)^k\frac{\sum_{i=N-k}^{N}V_i}{k+1}$. Rearranging the terms we get limitation on V_1: $V_1(\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k} - \frac{(1-p)^{(N-1)}}{N}) \le a_s + \sum_{i=2}^{N}V_i(\sum_{k=N-i}^{N-2}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{k+1} + \frac{(1-p)^{N-1}}{N} - \sum_{k=0}^{N-i}\binom{N-1}{k}p^{N-1-k}(1-p)^k)$. This is the second inequality that gives us restrictions on V_1. Coefficient in front of V_1 is larger than zero, as $\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k} - \frac{(1-p)^{(N-1)}}{N} \ge \frac{(p+(1-p))^{(N-1)}}{N} - \frac{(1-p)^{(N-1)}}{N} > 0$. Sign in front of any $V_i \in \{V_2, V_3 ... V_N\}$ is determined by

$\beta_i = \sum_{k=N-i}^{N-2}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{k+1} + \frac{(1-p)^{N-1}}{N} - \sum_{k=0}^{N-i}\binom{N-1}{k}p^{N-1-k}(1-p)^k$. Here, the first two pars represent the probability of winning respective V_i when being a low type, while the last part represents the probability of winning V_i when being a high type.

Therefore, we get two restrictions on V_1that should be satisfied simultaneously - $V_1|$

$$
\begin{cases}
V_1\left(1-\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k}\right) \leq a_s + \sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k\sum_{i=2}^{N-k}V_i}{N-k} \\
V_1\left(\sum_{k=0}^{N-1}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{N-k} - \frac{(1-p)^{(N-1)}}{N}\right) \leq a_s + \sum_{i=2}^{N} V_i\left(\sum_{k=N-i}^{N-2}\frac{\binom{N-1}{k}p^{N-1-k}(1-p)^k}{k+1} + \right. \\
\left. \frac{(1-p)^{N-1}}{N} - \sum_{k=0}^{N-i}\binom{N-1}{k}p^{N-1-k}(1-p)^k\right)
\end{cases}
$$

So, the final restriction on V_1 indeed has the form $\alpha_1 V_1 \leq \alpha_0 + \sum_{i=2}^{N}\alpha_i V_i$, where $\alpha_1 > 0$.

□

3.2.3.2 Individual sabotage protocol

In the case of individual sabotage protocol N agents I should also specify how players sabotage. I assume that a player sabotages one opponent by choosing $d \in \{0,\, 1\}$. I look for the symmetric PBE of the same form as before. Each agent chooses a productive effort $e \in \{0,\, 1\}$ and destructive effort $d \in \{0,\, 1\}$. When a player decides to make a destructive effort he chooses at random which one of opponents will be sabotaged.

First I look under which conditions there exist such a combination of $\{V_1,\, V_2...V_N\}$ where a principal can guarantee first-best outcome (both types make only productive efforts and do not sabotage at all). I again assume that in case of equal outcomes agents share the sum of relevant prizes. For example if all agents ended with the same $\tilde{e}$ they share the sum of N prizes among each other, while if $N-1$ agents have the same observed outcome they share either $(V_1 + V_2 + ...+V_{N-1})$ or $(V_2 + V_3 + ... + V_N)$depending on weather they are leaders or followers.

When this first-best strategy profile is an equilibrium the expected payoff of players is $EU\,(\boldsymbol{s}^*) = p(\frac{\sum_{i=1}^{N}V_i}{N} - a_h) + (1-p)(\frac{\sum_{i=1}^{N}V_i}{N} - a_l)$. As all agents make the same amount of effort in equilibrium they are correctly expecting to end up with the same $\tilde{e}_i$ and therefore to share the prize sum.

Claim 3.6. The strategy profile $\boldsymbol{s}^* = \{e_H = 1,\, d_H = 0,\, e_L = 1,\, d_L = 0\}^3$ constitutes a symmetric PBE in the game described above if$\{V_1,\, V_2,\, V_3\}$ satisfy the following inequalities:

(1) $\sum_{i=1}^{N-1}\frac{V_i}{(N-1)} - V_N \leq Na_s$

(2) $\sum_{i=1}^{N} V_i - \frac{(N-2)(V_N+V_{N-1})}{2} \geq N(a_L - a_s)$

(3) $\frac{\sum_{i=1}^{N} V_i}{N} - V_N \geq a_L$

Proof. In order for $\boldsymbol{s}^* = \{e_H = 1,\ d_H = 0,\ e_L = 1,\ d_L = 0\}^N$ to be a symmetric PBE we should ensure that there are no profitable deviations. There are three possible deviations from $\boldsymbol{s}^*$ for each type (to make effort and sabotage; just to sabotage; and not to make neither effort nor sabotage). Combination of prizes where none of these deviations is profitable determines the set of $\{V_1,\ V_2,...,\ V_N\}$ where $\boldsymbol{s}^*$ is an equilibrium strategy profile.

I argue that only three deviations are binding in determining the domain where equilibrium exists: 1) any type deviates to make both productive effort and sabotage; 2) the low type thinks about switching to $e_L = 0$ and $d_L,= 1$; and 3) the low type wants to switch to $e_L = 0$, and $d_L = 0$.

First, the deviation to both productive effort and sabotage brings the same no-deviation condition for both types: $\frac{\sum_{i=1}^{N} V_i}{N} \geq \frac{\sum_{i=1}^{N-1} V_i}{N-1} - a_s$. Rearranging, we get condition $\sum_{i=1}^{N-1} \frac{V_i}{N(N-1)} - \frac{V_N}{N} \leq a_s$. This gives us the restriction on $\{V_1, V_2...V_{N-1}\}$.

High type never wants to deviate to $e_H = 0$, $d_L = 1$ as sabotage for him is more costly and less efficient. However, low type can be tempted to deviate. In that case he will get $\frac{V_N + V_{N-1}}{2} - a_s$. No-deviation condition requires that $\frac{V_N + V_{N-1}}{2} - a_s \leq \frac{\sum_{i=1}^{N} V_i}{N} - a_L$, which is equal to $\frac{\sum_{i=1}^{N} V_i}{N} - \frac{(N-2)(V_N + V_{N-1})}{2N} \geq a_L - a_s$. This gives restriction on V_N and V_{N-1}.

Finally, as effort is more expensive for low type he has more incentive to deviate to no work/no sabotage profile This deviation brings to him the payoff of V_N, therefore, the binding condition for this deviation not to be profitable is $\frac{\sum_{i=1}^{N} V_i}{N} - a_L \geq V_N$. This condition gives us restriction on V_N.

□

Rearranging the first and third inequalities we get:

(1) $V_N \geq \sum_{i=1}^{N-1} \frac{V_i}{(N-1)} - N a_s$

(3) $V_N \leq \sum_{i=1}^{N-1} \frac{V_i}{(N-1)} - \frac{N a_L}{N-1}$

This inequalities immediately show that $\boldsymbol{s}^*$ can be an equilibrium profile only if $(N-1)a_s \geq a_L$. It is a necessary condition for $\boldsymbol{s}^*$ to be an equilibrium.

The next question I am addressing is when this equilibrium can be supported by giving all the prize sum to the winner ($V_1 = S$, $V_2 = ... = V_N = 0$). It turns out that this is possible when $S \leq N(N-1)a_s$, $S \geq N(a_L - a_s)$, and $S \geq N a_L$, where S is the prize sum. So the first-best equilibrium can be supported by giving all the prize to the winner when $S \in [Na_L;\ N(N-1)a_s]$. If the prize sum does not lie in this region the first-best equilibrium assumes that the winner does not get the whole prize sum.

If the first-best outcome can not be achieved, for example when sabotage is relatively cheap ($(N-1)a_s \geq a_L$) the best option for the principal is the second-best outcome. As I've mentioned before the second-best outcome is when high type agent makes only productive effort, while the low type agent makes no efforts at all - neither productive nor distructive. I want to find for which combination of prizes this strategy profile constitutes an equilibrium.

In this case the high type agent has an expected payoff of $EU_H = \sum_{k=0}^{N-1}\begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k \frac{\sum_{i=1}^{N-k} V_i}{N-k} - a_h$. The low type agent has $EU_L = \sum_{k=0}^{N-1}\begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k \frac{\sum_{i=N-k}^{N} V_i}{k+1}]$. High type can deviate to a strategy profile where he makes both productive and sabotage effort, or to a strategy profile where he does not make any effort at all. Low type agent can deviate to a strategy profile where he can either make a sabotage effort, or to a strategy profile where he makes both types of efforts. These four conditions make the restriction on V_1, V_2,...,V_N space where the desirable strategy profile is supported as an equilibrium.

Claim 3.7. When $(N-1)a_s \geq a_L$ the principal can only guarantee second-best outcome where high type agent works, while low-type agent does not work and does not sabotage. He can guarantee this outcome by setting $V_1, V_2 ... V_N$ that satisfy following inequalities:

Proof. If a low type deviates to a profile where he sabotages he gets the following expected payoff: $EU_L = \sum_{k=0}^{N-1}\begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k\left(\frac{N-k-1}{N-1}\frac{\sum_{i=N-k-1}^{N} V_i}{k+2} + \frac{k}{N-1}\frac{\sum_{i=N-k}^{N-1} V_i}{k}\right) - a_s$. The deviation is not profitable when $\sum_{k=0}^{N-1}\begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k\left(\frac{N-k-1}{N-1}\frac{\sum_{i=N-k-1}^{N} V_i}{k+2} + \frac{k}{N-1}\frac{\sum_{i=N-k}^{N-1} V_i}{k}\right) - a_s \leq \sum_{k=0}^{N-1}\begin{pmatrix} N-1 \\ k \end{pmatrix} p^{N-1-k}(1-p)^k\frac{\sum_{i=N-k}^{N} V_i}{k+1}$. Rearranging I get $\sum_{i=1}^{N-1}\alpha_i V_i + \alpha_N V_N \leq a_s$, where $a_i = (C(N-1, N-i-1)p^i(1-p)^{N-i-1}\frac{i}{(N-1)(N-i+1)} + \sum_{k=N-i}^{N-1} C(N-1,k)p^{N-1-k}(1-p)^k(\frac{N+1}{(N-1)(k+2)} - \frac{1}{k+1}))$, and $a_N = \sum_{k=0}^{N-1} C(N-1,k)p^{N-1-k}(1-p)^k(\frac{N-k-1}{(N-1)(k+2)} - \frac{1}{k+1})$. We see that $\alpha_1 > 0$, so that this inequality always restricts V_1. For other coefficients it's also possible to make several statements. The coefficient $\alpha_i > 0$ if $2k > N-3$. This will always be the case when $i < \frac{N+3}{2}$. So, this inequality restricts at least first $\frac{N+3}{2}$ prizes. For example, when $N = 3$ it actually means restriction on both $V_1$and V_2, while when N rises at least $\frac{N}{2}$ prizes should be restricted from above.

If a low type deviates to a profile where he both sabotages and makes productive effort he gets the following expected payoff:

$EU_L = \sum_{k=0}^{N-1} \binom{N-1}{k} p^{N-1-k}(1-p)^k \left(\frac{N-k-1}{N-1} \frac{\sum_{i=1}^{N-k-1} V_i}{N-k-1} + \frac{k}{N-1} \frac{\sum_{i=1}^{N-k} V_i}{N-k}\right) - a_s - a_l$. This deviation is not profitable when $\sum_{i=1}^{N-1} \gamma_i V_i - \gamma_N V_N \leq a_s + a_l$, where $\gamma_i = \sum_{k=0}^{N-i} C(N-1,k)p^{N-1-k}(1-p)^k\left(\frac{k}{(N-1)(N-k)} + \frac{1}{N-1}\right) - \sum_{k=N-i}^{N-1} C(N-1,k)p^{N-1-k}(1-p)^k \frac{1}{k+1}$, and $\gamma_N = \sum_{k=0}^{N-1} C(N-1,k)p^{N-1-k}\frac{(1-p)^k}{k+1}$. We can see easily that $\gamma_1 > 0$ therefore this inequality restricts V_1.

If a high type deviates to a profile where he both works and sabotages then we get again the inequality of the form $\sum_{i=1}^{N-1} \gamma_i V_i - \gamma'_N V_N \leq a_s$, where γ_i are the same as in previous case, while $\gamma'_N = \frac{p^{N-1}}{N}$. As we have the same coefficient in front of V_1 this inequality also will restrict the first prize from the above.

Finally, if high type deviates to the profile where he neither makes productive effort nor sabotages he we get the restriction of $\sum_{i=1}^{N} \delta_i V_i \geq a_h$, where $\delta_i = \sum_{k=0}^{N-i} C(N-1,k)p^{N-k-1}(1-p)^k \frac{1}{N-k} - \sum_{k=N-i}^{N-1} C(N-1,k)p^{N-k-1}(1-p)^k \frac{1}{k+1}$. Here $\delta_1 > 0$, while $\delta_N < 0$, so this inequality restricts V_N and (possibly) other prizes for loosers. $\square$

3.3 Continuous cost case

3.3.1 Two agents

Now instead of looking at the case of discrete types I consider the case where costs are distributed according to some distribution $F\,[0,\,1]$, and $a_s \in [0,\,1]$. To simplify the solution I assume now that the cost of productive effort e_i for the type a_i is $\frac{e_i}{a_i}$ so that for the agents with lowest a productive effort is the most costly. The cost of distructive activity is $\frac{d_i}{a_s}$ where a_s is the cost of sabotage that is unique for all agents. I am looking for equilibrium bidding strategies $e\,(a)$, and $d\,(a)$. Moreover, I look for equilibrium where the sum of two $\gamma\,(a) = e\,(a) + d\,(a)$ is monotone in a. I find that the optimal bidding function is piecewise function:

$$\gamma^*\,(a) = e^*\,(a)\,, \text{ when } a \geq a_s$$
$$\gamma^*\,(a) = d^*\,(a)\,, \text{ when } a < a_s$$

where $d^*\,(a) = (V_1 - V_2)\,a_s \int_0^a f\,(a)\,da$, and $e^*\,(a) = (V_1 - V_2)\int_{a_s}^a af\,(a)\,da + d(a_s)$.

The designer chooses V_1 and V_2 in order to maximize $-2\int_0^{a_s} d^*\,(a)\,f\,(a)\,da + 2\int_{a_s}^1 e^*\,(a)\,f\,(a)\,da$.

Claim 3.8. When $a_s > \bar{a}_s$ the optimal solution of the principal implies that principal gives all the prize to the winner and when $a_s \leq \bar{a}_s$ the optimal solution implies that principal sets $V_1 = V_2$.

Proof. The objective function of the principal is

$$W\,(a_s) = -2a_s\,(V_1 - V_2)\int_0^{a_s}\left(\int_0^a f\,(x)\,dx\right) f\,(a)\,da + 2\,(V_1 - V_2)\int_{a_s}^1\left(\int_{a_s}^a xf\,(x)\,dx\right) f\,(a)\,da.$$

We can rewrite it as

$$W\,(a_s) = (V_1 - V_2)\,k\,(a_s, a)$$

If $a_s = 0$ then $W\,(a_s) = 2\,(V_1 - V_2)\int_0^1\left(\int_0^a af\,(a)\,da\right) f\,(a)\,da = (V_1 - V_2)\,k\,(0,\,a)$, where $k\,(0,\,a) > 0$. Then principal would then want to set $V_1 = S$, and $V_2 = 0$, where S is the total prize sum that principal wants to distribute. If $a_s = 1$ then $W\,(a_s) = -2a_s\,(V_1 - V_2)\int_0^1\left(\int_0^a f\,(a)\,da\right) f\,(a)\,da = -\,(V_1 - V_2)\,k\,(1,\;a)$, where $k\,(1,\,a) > 0$. Then principal would like to set $V_1 = V_2$, not making any contest at all but instead distributing prizes equally. I argue that $W\,(a_s)$ is monotone in a_s. The $k\,(a_s,\,a)$ is monotonically decreasing in a_s $\frac{dk}{da_s} = -2(\int_0^{a_s}\left(\int_0^a f\,(a)\,da\right) f\,(a)\,da) + a_sF\,(a_s)\,f\,(a_s)) - a_sf\,(a_s)\,(1 - F\,(a_s)) < 0$. Therefore, $k\,(a_s,\,a)$ is monotonically decreasing in a_s, and $k\,(0,\,a) > 0$, while $k\,(1,\,a) < 0$. So, by intermediate value theorem $\exists$ some $\bar{a}_s$ s.t.

$k(\bar{a}_s, a) = 0$, while for $a_s < \bar{a}_s$ $k(a_s, a) > 0$ and in order to maximize W the principal will set $V_1 = S$, and $V_2 = 0$. $\square$

3.3.2 Multiple agents

Now I assume that there are N agents who compete for three prizes V_1, V_2... V_N. I also assume that cost is distributes uniformly: $a \sim U[0, 1]$.

3.3.2.1 Mass sabotage protocol

If sabotage follows mass protocol then when one agents chooses a sabotage level $d(a)$ he harms all other agents by this amount at the same time. The objective function of an agent in this case is

$$\sum_{i=0}^{N-1} V_i C(N-1, i) \left(\gamma^{-1}(e+d)\right)^{N-1-i} \left(1 - \gamma^{-1}(e+d)\right)^i - \frac{e}{a} - \frac{d}{a_s} \to \max_{e \geq 0,\, d \geq 0}.$$

Again optimal decisions of the agents take a form

$$\gamma^*(a) = e^*(a), \text{ when } a \geq a_s$$
$$\gamma^*(a) = d^*(a), \text{ when } a < a_s$$

where

$d^*(a) = V_1 a_s a^{N-1} + a_s \sum_{i=1}^{N-2} V_{i+1} C(N-1, i)[(N-1-i)B(a, N-1-i, i+1) - iB(a, N-i, i)] + V_N a_s (1-a)^{N-1}$
$e^*(a) = e_0^* + c = V_1 a^N \frac{N-1}{N} + \sum_{i=1}^{N-2} V_{i+1} C(N-1, i)[(N-1-i)B(a, N-i, i+1) - iB(a, N-i+1, i)] - V_N(N-1)B(a, 2, N-1) + c,$

where $B(x, k, q)$ is a partial Beta-function.

For the case of $N = 3$ agents the optimal sabotage and effort functions will take the following form:

$$d^*(a) = a_s a \left[V_1 a + 2V_2(1-a) - 2V_3\left(1 - \frac{a}{2}\right)\right]$$
$$e^*(a) = a^2 \left[\frac{2}{3} V_1 a + 2V_2\left(\frac{1}{2} - \frac{2}{3}a\right) - 2V_3\left(\frac{1}{2} - \frac{a}{3}\right)\right] + c.$$

We can see that V_1, and V_2 have positive influence on both productive effort and sabotage, while V_3, has negative influence on both types of efforts.

The constant c is found from $d^*(a_s) = e^*(a_s)$ and it is equal to
$c = a_s^2 \left(\frac{1}{3}V_1 a_s + 2V_2\left(\frac{1}{2} - \frac{1}{3}a_s\right) - 2V_3\left(\frac{1}{2} - \frac{a_s}{6}\right)\right)$.

The objective function of the principal is

$$W(a_s) = -3\int_0^{a_s} d^*(a)\, da + 3\int_{a_s}^1 e^*(a)\, da \rightarrow \max_{V_1,\, V_2,\, V_3}$$

We can rewrite it as $W(a_s) = V_1k_1 + V_2k_2 + V_3k_3$, where $k_1 = \frac{1}{6} + \frac{a_s^3}{3} - \frac{5}{6}a_s^4$, $k_2 = a_s^2\left(1 - 3a_s + \frac{5}{3}a_s^2\right)$, $k_3 = -\frac{1}{6} - a_s^2 + \frac{8}{3}a_s^3 - \frac{5}{6}a_s^4$, so the objective function of the principal is linear in prizes (just like in [21]). If $a_s = 0$, which corresponds to sabotage being infinitely costly, $k_1 = \frac{1}{6}$, $k_2 = 0$, $k_3 = -\frac{1}{6}$, and the principal should give the whole prize sum to the winner. In contrast, if $a_s = 1$ meaning that sabotage is very cheap $k_1 = -\frac{1}{3}$, $k_2 = -\frac{1}{3}$, $k_3 = \frac{2}{3}$.

We also observe that the second prize will always have lower impact on the principal objective function than the first prize, however relationships with the third prize depend on a_s.

In this case the optimal decision of the principal will also have a form of bang-bang solution:

$$V_1 = S,\ V_2 = V_3, \text{ when } a_s \leq \bar{a}_s$$
$$V_1 = V_2 = V_3, \text{ when } a_s > \bar{a}_s$$

,where $\bar{a}_s \approx 0.7$.

So, when the cost of sabotage is below $\bar{a}_s$(which means that sabotage is expensive) the principal will indeed want to distribute all the prize to the winner. However, once the cost of the sabotage surpasses $\bar{a}_s$he does not want to make the contest at all, and distributes prizes equally between participants. In that case too many types too sabotage, and the damage of sabotage would overweight for the principal the gains of effort.

3.4 Conclusions

I find that in the presence of sabotage can change substantially the optimal allocation of the prizes in a contest. The scheme where all sum is allocated to the winner may be not optimal anymore because it creates very high sabotage incentives.

This indeed happens for the case of two agents. Here, principal optimally also gives a substantial reward for the looser so that to discourage latter for sabotaging. Such prize allocation allows principal to achieve second best.

In case of more than two agents sabotage may or may not be bit concern depending on the sabotage protocol. In cases of mass sabotage the result of two-agent case still holds. However, the case of individual sabotage differs substantially as sabotage becomes a public good. Desire of the players to free-ride on cost of sabotage can result (for high sabotage cost) in first-best outcome being achieved again with a standard winner-gets-all payment scheme.

Results become more extreme for continuous type distribution. For two players and for three players with uniform distribution the decision of a principal has a form of a "bang-bang" solution - either principal gives all the prize to the winner when sabotage is expensive or he does not want to make a contest at all when sabotage is cheap.

This provides a very important insight that a contest might not be attractive at all when sabotage is very easy.

Chapter 4

Teaching to be selfish: classroom experiment on PD

I. Kirysheva[1], C. Papioti[2]

[1]irina.kirysheva@nu.edu.kz
[2]chara.papioti@esc-rennes.fr

4.1 Introduction

Game theory courses now constitute an essential part of education in both economics and business. They teach students main concepts but also introduce them to that particular way of thinking, teach to take into account strategies of other players and to view the game as it is for other counterparts.

In our paper we wanted to analyze to what extent the fact that one is accustomed with game theoretical tools alternates persons' behavior and in particular the decisions made while playing prisoners dilemma. We wondered if those who knew game theory behave systematically more selfish and less cooperative. We also asked ourselves if these potential differences in behavior were present only in some circumstances and not in others (e.g. when playing against computer, or random partner, and not when playing with a well-known people). To address these questions we have conducted classroom experiments where students were playing Prisoners Dilemma. To find the source of changed behavior we conducted different treatments.

It is important to know the influence of studying game theory on students' behavior for the purpose of teaching economics. Moreover, this is a useful insight for experimental economics as it helps to plan better the sample of participants and to know the potential differences between players with different course backgrounds.

Our experiment shows that indeed there is a huge difference in results before and after the course. However our experiment suggests that players behave less cooperatively not because of the knowledge of game theory per se, but because they know that their opponents are accustomed to game theory and this alters their expectations.

We've also found that before the course players only took into account their own information about the game while after the course they have considered both their information and that of their partner which shows that their level of reasoning has increased. We've also seen that when two partners in a group were endowed with different information (one had knowledge on the partners identity while another didn't have) those who had more information were especially cooperative, while those who had less information behaved in especially individualistic manner.

Finally, after the course we have also found gender differences in expectations in random treatment, which in turn caused the difference in played strategies between male and female players. While female participants were expecting their partner to defect, males expected only cooperation.

To sum up our results, we've found that after the course of game theory students increase their level of reasoning and behave significatly less cooperatively because they expect their opponents to defect; this is especially female students when playing with a random partner.

4.2 Literature

There are several strands of literature that are relevant to our work.

Classroom experiments have been extensively used in economics due to their relative simplicity and accessibility. Mostly papers investigate the influence of classroom experiments techniques on the learning process - e.g. Holt ([14]). There are not so many classroom experiment on classical prisoners dilemma though.

There are several works on the effect of studying economics. Marwell et. al. [18] conducted a public good experiment and have found that economists contribute significantly less (around 20% for economists compared to around 40% for other disciplines). However, in their study Isaac et. al. [17] have not found any differences in contributions between economists and sociologists when the game was repeated and non-discrete with respect to contribution levels.

Frank et. al. in [6] look at the connection between being an economist and cooperation. Authors have conducted a survey on charity donation for different faculty and have found that professional economist donate less. They also perform the prisoners dilemma game between students of different backgrounds that also show that economists tend to defect more. The authors comment a bit on the self-selection issue that initially more self-interested people may be more appealed towards economics. They also point the substantial gender differences in their study showing that male participants behave less cooperatively. The authors comment on potential importance of beliefs differences between economists and non-economists but do not elicit beliefs explicitly in their experiment.

Unlike the work of Frank et. al. [6] Yezer et.al. [32] argue that while economists may behave selfishly when playing hypothetical games or responding to surveys, in real-world they still cooperate a lot. The authors create a"lost-letters" experiment where letters with cash were left in different classrooms. The results show that economist students return the letters significantly more often than students of other disciplines.

Frank et. al. [7] address the concern raised by Yezer et. al. [32] by explaining that their claim concerned mainly behavior in social dilemmas. Authors still argue that in social dilemmas economist students are marginally less cooperative. They place some doubt on the accuracy of results of Yezer et. al. [32] and outline that several social dilemma experiments still have shown significantly less cooperation by economists.

Ferraro in [5] argues that economic theories shape the day-to-day institutions, social norms and language, and therefore become self-fulfilling. The authors cite the evidence that self-interest is a learned behavior.

One of the closest papers to what we are doing is the one by Rubinstein [26]. The author has made his students play some preclass online games before the course on game theory and after the course. He played various games to test different game-theoretical concepts and have found that there does not seem to be any substantial difference in how students behave before and after the course. However unlike our experiment in the work of Rubinstein [26] the choices were hypothetical (there were no prizes associated with them), and the Prisoners Dilemma was present only in a dynamic version.

The paper of Byrnes et. al. [2] looks at the differences in risk-attitude between men and women. This is relevant for our research as we have found gender particularities in behavior and differences in risk attitude might be a potential explanation. In the paper authors conduct a meta-analysis of 150 studies where risk-attitudes of males and females were measured. They have found that for almost all types of risk behaviour there was a significant gender differences with males being more risk-taking.

In our research we concentrated on the influence of game theory on the behavior of the students. While being partially linked to economic concepts (and applied to many of them) game theory can alter performance in a particular manner. Game theory teaches students to think strategically, to see the game also through the opponents eyes and to best respond to the opponents strategies. In our experiment we were especially interested in the reasons for the change in behavior. We've conducted different treatment to distinguish between potential mechanisms of the influence of game theory course.

4.3 Experiment

We've conducted series of classroom experiments on students at ESC-Rennes (Ecole Superieure de Commerce in Rennes). The students were a part of our class on Microeconomics. This course was taught at the PGE2 program (equivalent to the end of undergraduate/beginning of master programs). Students were business students and this was the first course on economics they've ever had, neither had they before any game theory course. We've played the prisoners dilemma game in the class before and after studying the game theory. We have also conducted different treatments both before and after experiments.

We formulated our prisoners dilemma as a story and we kept the context. Our idea was that for those who are not accustomed with either game theory or even mathematical way of thinking it is easier to follow the game in a context. In every class the winner (or a random draw from the set of winners when there were several of them) was rewarded with the prize. Students applauded to the prizes and were very enthusiastic which confirms that those prizes provided enough incentives for them. In the treatment with belief elicitation each correct belief guess was awarded with a chocolate.

We have 199 students who played the game in the beginning of the course who were divided into 9 groups. Out of them 3 groups were of stranger matching protocol where each player was randomly matched with someone and didn't know with whom he was playing. Two groups were of partner matching protocol, where both participants in the pair knew exactly with whom they were playing. Three groups were of half stranger/half partner matching: in each pair one player knew his partner, while another player in the pair didn't know the partner. Finally, one group played against computer.

There were 226 students in the experiment after game theory course who were divided into 10 groups[3]. Out of them three groups were playing random treatment (paired with strangers), two groups were playing partner matching, three groups were playing half random/half partner with only one in the pair having full knowledge, and two groups were playing against the computer. Moreover, one group in each of the treatments (random, partner, half random/half partner, computer) was playing with a modified payoff table and modified story. While it was still prisoners dilemma in its nature, it was presented now as a duopoly story. We also had one random treatment where experiment was not performed by a class teacher but by other person. This treatment addressed the concern that in classroom experiments students might just want to please

[3]This were the same students are before the course. There were some students who only participated before or after due to their presence in the class

the teacher and behave accordingly. Finally, in one of the random treatments we've elicited beliefs after the main decision was made.

We have also included the data on the gender of the players and on the final course score. Gender variable allows us to check if indeed males are less cooperative, while score is a proxy for student ability and dilience.

4.4 Results

4.4.1 Before studying GT

We have four main treatments in our experiment that differ in the matchingprotocol: random, partner, half random/half partner, and computer. Though, we were assigning each treatment in a completely random fashion to different groups, we still want to check that there was no potential selection of subjects, for example that one particular group contains smarter students than other groups.

First we look at the average exam scores and gender for different treatments (see Figure 4.1). As we can see there does not seem to be any significant difference neither in the exam grades nor in the gender between various treatment groups. We also perform a test to confirm that treatment assignment was completely random.

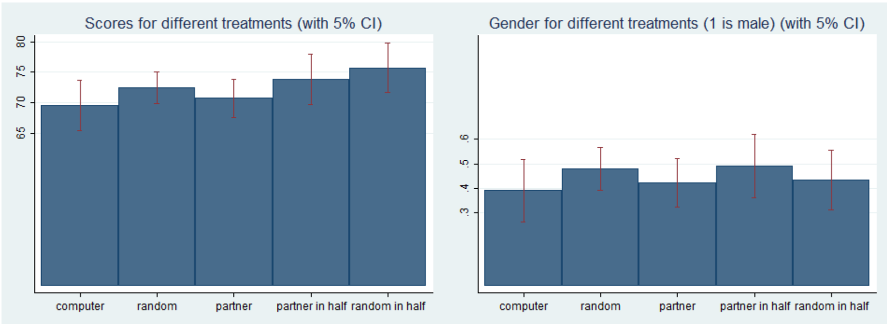


FIGURE 4.1:

To compare our four treatments and show that they are similar with respect to the students who participated we perform the one-way ANOVA test[4]. We compare if either final score or gender is differently distributed in any of our treatments. For both variables one-way ANOVA test rejects the hypothesis that the distribution is different between treatments (see Appendix).

As our treatment assignments are indeed random we can proceed to the treatment effects estimation. First, we investigate what happens before students learn game theory. The results of a probit regression are presented in Table 4.1. We've checked several specifications and get robust results. We estimate the probability of defection (e.g. playing Nash strategy) for four treatments, while computer treatment serves as a baseline. For interpretational purposes we've divided half random/half partner treatment into two: those who knew their partner (half partner), and those who didn't (half random).

From here we already see that probability of defection is lower for all four treatment in comparison to the computer one. Moreover, for partner and partner in half it is

[4]It allows to check that samples in multiple groups are drawn from populations with the same means

Table 4.1: Before the Game Theory

	(1) defection	(2) defection	(3) defection	(4) defection
defection				
random	-0.337	-0.335	-0.359	-0.356
	(0.28)	(0.28)	(0.28)	(0.28)
partner	-0.666**	-0.685**	-0.704**	-0.705**
	(0.30)	(0.30)	(0.30)	(0.30)
partner in half	-0.849**	-0.709**	-0.767**	-0.755**
	(0.34)	(0.35)	(0.35)	(0.35)
random in half	-0.253	-0.198	-0.206	-0.194
	(0.33)	(0.33)	(0.33)	(0.33)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	199	192	192	192

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

significantly lower. To get more detailed picture on the influence of different treatments on defection we look at the marginal effects (Table 4.2).

Table 4.2: Marginal effects of treatment before the Game Theory

	(1) defection	(2) defection	(3) defection	(4) defection
computer	0.600	0.598	0.606	0.604
	(0.09)	(0.09)	(0.09)	(0.09)
random	0.467	0.466	0.465	0.464
	(0.06)	(0.06)	(0.06)	(0.06)
partner	0.340	0.332	0.333	0.331
	(0.07)	(0.07)	(0.07)	(0.07)
partner in half	0.276	0.32	0.310	0.313
	(0.08)	(0.09)	(0.09)	(0.09)
random in half	0.500	0.521	0.525	0.528
	(0.09)	(0.09)	(0.09)	(0.09)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	199	192	192	192

Standard errors in parentheses

We can see that we get robust marginal effects for different model specifications. The highest probability of playing Nash strategy is found for the computer treatment (around 0.6). The second highest probability is for those not knowing their partner in half random/half partner treatment (0.52), while for completely random treatment it's a bit smaller (0.46). Finally, the most cooperative behavior is reported by partner treatment and those playing a partner part in half random/half partner treatment (around 0.33 and 0.31 respectively). Based on this results we can divide our treatments in three blocks - 1) computer, 2) where players didn't know the partner (completely random and half random), and 3) where players knew the partners (partner and half partner). We see that playing against the real person (compared to playing against computer) matters and creates higher incentives to cooperate. Knowing the partner ensures even more cooperation. We can also observe that in the second block when participant is informationally disadvantaged (in the half random treatment he does not know the opponent, while opponent knows him) he tends to compensate it by slightly higher probability of defection.

One important result we get is that before learning game theory only personal information matters. Outcome of completely random and half random treatments are practically identical; the same is true for complete partner or half partner treatments.

We also look at the marginal effects of the gender (Table 4.3). Male participants have probability of 0.48 of defection, while for females it's slightly lower (0.41). There is indeed a difference but it's not that big as was mentioned in some previous literature.

Table 4.3: Marginal effects of gender before the Game Theory

	(1) defection	(2) defection
female	0.411	0.411
	(0.05)	(0.05)
male	0.480	0.480
	(0.05)	(0.05)
score	No	Yes
Observations	192	192

Standard errors in parentheses

4.4.2 After studying GT

It is not surprising that after having a course on Game Theory probability of defection among students increased significantly. Figure 4.2 presents the defection probabilities before and after.

FIGURE 4.2:

Before only in 0.43 percents of the cases students played confess; after the figure has risen to 0.7. As we are particularly interested in what stands behind it we look at the dynamics of defection for every treatment.

Table 4.4 shows the probit regression for probability of defection after the course.

TABLE 4.4: After the Game Theory

	(1) defection	(2) defection	(3) defection	(4) defection
defection				
random	0.111	0.085	0.128	0.100
	(0.28)	(0.29)	(0.29)	(0.29)
partner	0.318	0.278	0.287	0.273
	(0.30)	(0.31)	(0.31)	(0.31)
partner in half	0.034	0.078	0.112	0.080
	(0.31)	(0.33)	(0.32)	(0.33)
random in half	0.261	0.186	0.248	0.198
	(0.32)	(0.33)	(0.32)	(0.33)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	226	221	221	221

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We see that after the course for all the treatments the probability to play Nash strategy is (insignificantly) higher compared to computer treatment. This is exactly the opposite of what was going on before the course. So, whether you know your partner or are playing with a random person the sole fact that your opponent is accustomed with game theory increases your probability of defection.

In Table 4.5 we see the marginal effects of treatments.

Table 4.5: Marginal effects of treatment after the Game Theory

	(1) defection	(2) defection	(3) defection	(4) defection
computer	0.64	0.665	0.653	0.66
	(0.09)	(0.09)	(0.09)	(0.09)
random	0.686	0.695	0.699	0.698
	(0.06)	(0.06)	(0.06)	(0.06)
partner	0.755	0.758	0.752	0.755
	(0.06)	(0.06)	(0.06)	(0.06)
partner in half	0.658	0.692	0.693	0.691
	(0.08)	(0.08)	(0.08)	(0.08)
random in half	0.73	0.729	0.739	0.731
	(0.07)	(0.07)	(0.07)	(0.07)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	226	221	221	221

Standard errors in parentheses

So, we see that there is a substantial difference in how people play before and after game theory. We want to distinguish to what extend this difference is attributed to the fact that players are already accustomed with game and play it for the second time (the game they've already studied during the course), and to what extend the difference is due to knowing the concepts of game theory in general. Additionally, we want to trace the framing effect we might have due to using PD story.

4.4.3 Modified story treatment

After the course we've had five groups out of eleven to play modified game. The new game was essentially a PD but framed as an oligopoly story. In Table 4.6 we see the result of probit regression for both modified games and original ones.

We can see immediately that results are different for modified and not modified game. For not modified game the probability of defection is (insignificantly) higher for all treatments (compared to computer baseline), while for modified game it is lower for half partner treatment. In this sense it is a bit similar to before game theory situation for the original game where half partner treatment had the lowest probability of defection.

In Table 4.7 we have marginal effects for modified and original games after the GT course.

Table 4.6: After the Game Theory

	modified				not modified			
defection								
random	0.000	0.209	0.231	0.401	0.131	0.059	0.081	0.042
	(0.53)	(0.57)	(0.58)	(0.60)	(0.33)	(0.35)	(0.34)	(0.35)
partner	0.000	0.217	0.175	0.344	0.415	0.314	0.316	0.286
	(0.55)	(0.58)	(0.58)	(0.60)	(0.37)	(0.38)	(0.38)	(0.38)
partner in half	-0.727	-0.224	-0.222	-0.082	0.355	0.246	0.264	0.225
	(0.59)	(0.65)	(0.65)	(0.67)	(0.37)	(0.39)	(0.38)	(0.39)
random in half	0.067	0.204	0.350	0.440	0.355	0.238	0.262	0.214
	(0.63)	(0.65)	(0.69)	(0.70)	(0.37)	(0.39)	(0.38)	(0.39)
gender	No	No	Yes	Yes	No	No	Yes	Yes
score	No	Yes	No	Yes	No	Yes	No	Yes
Observations	77	74	74	74	149	147	147	147

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.7: Marginal effects after the Game Theory

	modified				not modified			
computer	0.800	0.746	0.740	0.698	0.571	0.612	0.606	0.619
	(0.13)	(0.15)	(0.16)	(0.17)	(0.11)	(0.11)	(0.11)	(0.11)
random	0.800	0.807	0.808	0.817	0.622	0.634	0.637	0.635
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
partner	0.800	0.809	0.793	0.802	0.724	0.725	0.721	0.722
	(0.09)	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
partner in half	0.545	0.671	0.664	0.669	0.704	0.702	0.703	0.701
	(0.15)	(0.15)	(0.16)	(0.15)	(0.09)	(0.09)	(0.09)	(0.09)
random in half	0.818	0.806	0.838	0.827	0.704	0.699	0.702	0.697
	(0.12)	(0.12)	(0.11)	(0.11)	(0.09)	(0.09)	(0.09)	(0.09)
gender	No	No	Yes	Yes	No	No	Yes	Yes
score	No	Yes	No	Yes	No	Yes	No	Yes
Observations	77	74	74	74	149	147	147	147

We see again a difference between modified game and a not modified one. First of all, defection probabilities in modified game are higher than in original game for all treatments but half partner one. Therefore, we see that framing effect is indeed present, and players indeed defect more when the game is framed as oligopoly rather than prisoners dilemma. In modified game the defection probabilities for all treatments but half partner are higher than in computer treatment (0.81 for random, 0.8 for partner, 0.82 for half random compared to only 0.69 in computer). So, we observe that playing with a real person matters, and it surprisingly leads to higher probability of playing Nash strategy and less cooperation. So, the sole fact that the partner is accustomed to game theory rises expectations about defection and leads indeed to less cooperation.

In modified game there is substantially less defection by partners in half partner/half random treatment. It can not be explained by social feelings (e.g. trust) because in partner treatment (where both players know exactly with whom they are playing) players defect much more (80% for both partner treatment vs. only 67% in half partner). We have already observed the slight differences in behavior between informational advantageous partners and disadvantageous randoms in half partner/half random treatment when we analyzed the PD before game theory. There half partners were more cooperative than partners in both partner treatment, and half randoms were less cooperative than randoms in both random treatment. We have observed the same pattern in both modified and unmodified games after the course, though with different magnitude. Those who knew with whom they are playing while their partner did not were more cooperative than those whose opponent shares their full knowledge (half partner vs. both partner). Similarly, those who didn't know their partner while the partner had this information were more individualistic than those whose partner did not have any information either (half random vs. both random).

Finally, there are some patterns in behavior when playing the game for the first time (PD before game theory and modified game after) and when plying the game they've already played before (PD after the course). In the games played for the first time defection in random treatments (both random and randoms in half random/half partner) is higher than in partner treatments (both partner and partners in half partner/half random). While for the game already played this becomes reversed. The highest defection is seen in both partner treatment where both players know with whom they are playing. The magnitude of this differences in defection rate between random and partner treatments also changes with players experience. When they had no knowledge about game theory and about this particular game they were playing in random treatment players were playing Nash 13 percent points more than in partner treatment. When playing a new game (modified treatment) after learning game theory the difference in defection between random and partner treatment became very small - only 1 percent point. However, when playing a game they've already played after the course the whole situation becomes reversed, and now in the partner treatment participants were playing Nash strategy 8 percent points more ofter than in random treatment. So, while before game theory social feelings and trust played huge role for participants, after the course they were not as important, and even playing with a known partner resulted in a higher defection when the game was already played before.

We also see that after the course for those who play PD defection probabilities in half random and half partner treatments are practically the same (0.7) and this is a very robust finding. Before the course these two probabilities were not so close, with half random probability being closer to both random, while half partner performance was closer to both partner. So, before GT behavior of the players just depended on the information they had about their partner and not on what the opponent knew. However, after the course the probabilities of half random and half partner treatments diverge from subsequent probabilities of full treatment and converge to each other. So, after the course when playing the familiar game both personal information and information available to the opponent become equally relevant.

4.4.4 Other teacher treatment

Now we address the concern for classroom experiments that students might unconsciously try to please their teacher and play accordingly. We have conducted one random treatment where the experiment was performed by another person. Table 4.8 presents the results of probit regression of the influence of the other teacher on the results.

TABLE 4.8: Other teacher treatment

	(1) defection	(2) defection	(3) defection	(4) defection
defection				
other_teacher=1	0.161	0.175	0.070	0.201
	(0.38)	(0.40)	(0.41)	(0.44)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	45	44	44	44

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We see that the influence of the other teacher on defection probability is (insignificantly) positive. If our initial concern was correct this would suggest that with the usual teacher students were playing more Nash while with the new teacher players should defect less. However, the the data does not support this and suggests quite the opposite.

The marginal effects of other teacher are presented in the Table 4.9.

With their usual teacher the players defected in 60% of the cases, while with another person conducting experiment they are doing it in 66%. So, we should not be worried that our results are driven by the fact that students want to please the teacher and play as they think they are expected to play.

Table 4.9: Other teacher treatment marginal effects

	(1) defection	(2) defection	(3) defection	(4) defection
other_teacher=0	0.591	0.602	0.625	0.602
	(0.10)	(0.11)	(0.09)	(0.10)
other_teacher=1	0.652	0.667	0.647	0.665
	(0.10)	(0.10)	(0.09)	(0.09)
gender	No	No	Yes	Yes
score	No	Yes	No	Yes
Observations	45	44	44	44

Standard errors in parentheses

4.4.5 Gender influence

Finally, in Table 4.10 we also look at the marginal effects of gender on the performance after the course.

Table 4.10: Marginal effects of gender after the Game Theory

	(1) defection	(2) defection
female	0.733	0.734
	(0.04)	(0.04)
male	0.681	0.681
	(0.05)	(0.05)
score	No	Yes
Observations	221	221

Standard errors in parentheses

While for both male and female players the probability of defection have risen, the figure has increased much more for women (from 0.41 to 0.73 for female, and from 0.48 to 0.68 for male). One might think that this can reflect that girls are usually more diligent students. However, there are no differences in final scores between female and male students with even boys having slightly higher scores (72.61 for male and 72.19 for female). We turn to players' beliefs for the possible explanation of this phenomena. In one random treatment we've elicited players beliefs. In table 4.11 we see the distribution of beliefs in opponents actions by sex.

We see that remarkably all male players expected their opponents to play cooperatively. On the contrary, our of 14 female players only 3 expected cooperation, while 11 expected defection. So, it turns out that man and women have completely different expectations

TABLE 4.11: Expectations by gender

	N
expect cooperation	
female	3
male	9
Total	12
expect defection	
female	11
male	0
Total	11
Observations	23

about their opponents in the random treatment - majority (precisely all) of male players expect their opponent to cooperate, while majority (78%) of female players expect their opponent to defect. In Table 4.12 we see what actions males and females were actually choosing for different expectations.

TABLE 4.12: Outcomes by expectation

	males	females
expect cooperation		
cooperate	6	1
defect	3	2
Total	9	3
expect defection		
cooperate		1
defect		10
Total		11
Observations	9	14

As we've already seen all male players expect their partners to cooperate. Though we do not have enough observation to make a strong prediction, we still can see that out of them 66% choose to cooperate themselves, and 33% choose to defect. On the contrary, only 3 women out of 14 expected her opponent to cooperate, and only 1 out of these 3 chose to cooperate herself. All the rest female players expected defection with all but one woman choosing to defect themselves.

So, we see that in random treatment female behavior is affected more by the course than male behavior. This is largely attributed to differences in expectations between men and women in random treatment, with women expecting their opponents to defect much more often, while men expecting cooperation.

4.5 Conclusions

In our experiment we have analyzed the effects of learning game theory on students behavior in a Prisoners Dilemma game. We have seen the substantial changes in overall performance and have addressed the possible cause of these changes.

Before the course students behave cooperatively mostly because social feelings they had towards their opponents (like friendship) lead to more optimistic belief about opponents actions. They were especially cooperative when they knew exactly with whom they were playing. We have also observed in the treatments where two players in a group were endowed with different information that players only take into account their private information and not that of the opponent.

However, after the course players' behavior became more selfish and they were playing Nash much more often. After the course participants are less cooperative when playing with a real person instead of computer. So, the sole fact that opponent himself knows game theory leads participants to be more selfish (whereas there was only small change in behavior for those playing against computer). Contrary to what might have been expected, after the course students defected especially a lot when playing with a known partner. In the treatments with different information distribution after the course players took into account both their own information and information of their opponent.

When playing a new game in the treatment with different information distribution between players, those who were informationally advantaged behave especially cooperatively, while those who were informationally disadvantaged behave in a particularly individualistic manner.

We've found substantial gender differences in the way students respond to learning game theory. Contrary to what have been suggested before, we have found that after the course when playing with a random partner girls were defecting much more often than boys. These differences are brought by the differences in their expectations. After the course in the random treatment majority of female players expected their partner to defect, while all male players expected cooperation.

4.6 Appendix

One-way ANOVA test for final score.

Source	SS	dF	MS	F	Prob>F
Between groups	1637.45349	4	409.363371	3*1.67	3*0.1568
Within groups	100208.39	408	245.608799		
Total	101845.844	412	247.19865		

For final score

One-was ANOVA test for gender.

Source	SS	dF	MS	F	Prob>F
Between groups	0.531535986	4	0.132883997	3*0.53	3*0.7111
Within groups	101.599215	408	0.249017683		
Total	101.599215	412	0.247890171		

For gender

Bibliography

[1] J Atsu Amegashie and Marco Runkel. Sabotaging potential rivals. *Social Choice and Welfare*, 28(1):143–162, 2007.

[2] James P Byrnes, David C Miller, and William D Schafer. Gender differences in risk taking: A meta-analysis. *Psychological bulletin*, 125(3):367, 1999.

[3] Jeffrey Carpenter, Peter Hans Matthews, and John Schirm. Tournaments and office politics: Evidence from a real effort experiment. *The American Economic Review*, 100(1):504–517, 2010.

[4] Kong-Pin Chen. Sabotage in promotion tournaments. *Journal of Law, Economics, and Organization*, 19(1):119–140, 2003.

[5] Fabrizio Ferraro, Jeffrey Pfeffer, and Robert I Sutton. Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review*, 30(1):8–24, 2005.

[6] Robert H Frank, Thomas Gilovich, and Dennis T Regan. Does studying economics inhibit cooperation? *The Journal of Economic Perspectives*, pages 159–171, 1993.

[7] Robert H Frank, Thomas D Gilovich, and Dennis T Regan. Do economists make bad citizens? *The Journal of Economic Perspectives*, pages 187–192, 1996.

[8] Paolo Garella. Adverse selection and the middleman. *Economica*, pages 395–400, 1989.

[9] Thomas Gehrig. Intermediation in search markets. *Journal of Economics & Management Strategy*, 2(1):97–120, 1993.

[10] Sanjeev Goyal and Fernando Vega-Redondo. Structural holes in social networks. *Journal of Economic Theory*, 137(1):460–492, 2007.

[11] Oliver Gürtler and Johannes Münster. Rational self-sabotage. *Mathematical Social Sciences*, 65(1):1–4, 2013.

[12] Christine Harbring and Bernd Irlenbusch. Sabotage in tournaments: Evidence from a laboratory experiment. *Management Science*, 57(4):611–627, 2011.

[13] Christine Harbring, Bernd Irlenbusch, Matthias Kräkel, and Reinhard Selten. Sabotage in corporate contests–an experimental analysis. *International Journal of the Economics of Business*, 14(3):367–392, 2007.

[14] Charles A Holt. Teaching economics with classroom experiments: A symposium. *Southern Economic Journal*, pages 603–610, 1999.

[15] Victoria Ivashina, Vinay B Nair, Anthony Saunders, Nadia Massoud, and Roger Stover. Bank debt and corporate governance. *Review of Financial Studies*, 22(1):41–77, 2009.

[16] Edward P Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts, 1979.

[17] R Mark Isaac, Kenneth F McCue, and Charles R Plott. Public goods provision in an experimental environment. *Journal of Public Economics*, 26(1):51–74, 1985.

[18] Gerald Marwell and Ruth E Ames. Economists free ride, does anyone else?: Experiments on the provision of public goods, iv. *journal of Public Economics*, 15(3):295–310, 1981.

[19] Eric Maskin and Jean Tirole. The principal-agent relationship with an informed principal: The case of private values. *Econometrica: Journal of the Econometric Society*, pages 379–409, 1990.

[20] Eric Maskin and Jean Tirole. The principal-agent relationship with an informed principal, ii: Common values. *Econometrica: Journal of the Econometric Society*, pages 1–42, 1992.

[21] Benny Moldovanu and Aner Sela. The optimal allocation of prizes in contests. *American Economic Review*, pages 542–558, 2001.

[22] Benny Moldovanu, Aner Sela, and Xianwen Shi. Carrots and sticks: prizes and punishments in contests. *Economic Inquiry*, 50(2):453–462, 2012.

[23] James D Montgomery. Social networks and labor-market outcomes: Toward an economic analysis. *The American economic review*, pages 1408–1418, 1991.

[24] Johannes Münster. Selection tournaments, sabotage, and participation. *Journal of Economics & Management Strategy*, 16(4):943–970, 2007.

[25] Roger B Myerson. Mechanism design by an informed principal. *Econometrica: Journal of the Econometric Society*, pages 1767–1797, 1983.

[26] Ariel Rubinstein. Experience from a course in game theory: pre-and postclass problem sets as a didactic device. *Games and Economic Behavior*, 28(1):155–170, 1999.

[27] Ariel Rubinstein and Asher Wolinsky. Middlemen. *The Quarterly Journal of Economics*, pages 581–594, 1987.

[28] Garth Saloner. Old boy networks as screening mechanisms. *Journal of Labor Economics*, pages 255–267, 1985.

[29] Sergei Severinov. An efficient solution to the informed principal problem. *Journal of Economic Theory*, 141(1):114–133, 2008.

[30] Dale O Stahl. Bertrand competition for inputs and walrasian outcomes. *The American Economic Review*, pages 189–201, 1988.

[31] Marie-Odile Yanelle. The strategic analysis of intermediation. *European Economic Review*, 33(2):294–301, 1989.

[32] Anthony M Yezer, Robert S Goldfarb, and Paul J Poppen. Does studying economics discourage cooperation? watch what we do, not what we say or how we play. *The Journal of Economic Perspectives*, pages 177–186, 1996.